Universality for first passage percolation on sparse random graphs

Citation for published version (APA):

DOI:
10.1214/16-AOP1120

Document status and date:
Published: 01/07/2017

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain.
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.
UNIVERSALITY FOR FIRST PASSAGE PERCOLATION ON SPARSE RANDOM GRAPHS

BY SHANKAR BHAMI DI1, REMCO VAN DER HOFSTAD2 
AND GERARD HOOGHEMSTRA

University of North Carolina, Eindhoven University of Technology and Delft University of Technology

We consider first passage percolation on the configuration model with \( n \) vertices, and general independent and identically distributed edge weights assumed to have a density. Assuming that the degree distribution satisfies a uniform \( X^2 \log X \)-condition, we analyze the asymptotic distribution for the minimal weight path between a pair of typical vertices, as well the number of edges on this path namely the hopcount.

Writing \( L_n \) for the weight of the optimal path, we show that \( L_n - (\log n)/\alpha_n \) converges to a limiting random variable, for some sequence \( \alpha_n \). Furthermore, the hopcount satisfies a central limit theorem (CLT) with asymptotic mean and variance of order \( \log n \). The sequence \( \alpha_n \) and the normalizing constants for the CLT are expressible in terms of the parameters of an associated continuous-time branching process that describes the growth of neighborhoods around a uniformly chosen vertex in the random graph. The limit of \( L_n - (\log n)/\alpha_n \) equals the sum of the logarithm of the product of two independent martingale limits, and a Gumbel random variable. So far, for sparse random graph models, such results have only been shown for the special case where the edge weights have an exponential distribution, wherein the Markov property of this distribution plays a crucial role in the technical analysis of the problem.

The proofs in the paper rely on a refined coupling between shortest path trees and continuous-time branching processes, and on a Poisson point process limit for the potential closing edges of shortest-weight paths between the source and destination.

CONTENTS

1. Introduction and results ...................................... 2569
   1.1. Motivation ........................................... 2569
       Organization of this section ......................... 2570

Received May 2014; revised September 2015.

1Supported by NSF-DMS Grants 1105581, 1310002, 160683, 161307 and SES Grant 1357622.
2Supported by the Netherlands Organisation for Scientific Research (NWO) through VICI Grant 639.033.806 and the Gravitation NETWORKS Grant 024.002.003.

MSC2010 subject classifications. 60C05, 05C80, 90B15.
Key words and phrases. Central limit theorem, continuous-time branching processes, extreme value theory, first passage percolation, hopcount, Malthusian rate of growth, point process convergence, Poisson point process, stable-age distribution, random graphs.

2568
1. Introduction and results.

1.1. Motivation. First passage percolation (FPP) is an important topic in modern probability theory, motivated by questions in a number of fields including disordered systems in statistical physics, where it arises as a building block in the analysis of various interacting particle systems such as the contact process, branching random walk and various epidemic models.

Let us start by describing the basic model. Let $\mathcal{G}$ be a connected graph on $n$ vertices. Assign independent and identically distributed (i.i.d.) random edge weights or lengths to the edges of the graph. These random edge weights generate geodesics on the graph. Think of the graph as a disordered random system carrying flow between pairs of vertices in the graph via shortest paths between them. Choose two vertices in the graph uniformly at random amongst the $n$ vertices. We
will call these two vertices “typical” vertices. Two functionals of interest are the minimal weight $L_n$ of a path between the two vertices and the number of edges $H_n$ on the minimal path, often referred to as the hopcount. We assume that the common distribution of the edge weights is continuous, so that the optimal paths are a.s. unique and one can talk about objects such as the number of edges in the optimal path.

This model has been studied intensively, largely in the context of the integer lattice $[-N, N]^d$ (see, e.g., [20, 23, 30, 39]). For the power of this model to analyze more complicated interacting particle systems, see [33] and [17] and the references therein. In the modern context, FPP problems take on an added significance. Many real-world networks (such as the Internet at the router level or various road and rail networks) are entrusted with carrying flow between various parts of the network. These networks have both a graph theoretic structure as well as weights on edges, representing for example congestion. In the applied setting, understanding properties of both the hopcount and the minimal weight are crucial, since whilst routing is done via least weight paths, the actual time delay experienced by users scales like the hopcount (the number of “hops” a message has to perform in getting from the source to the destination). Simulation-based studies (see, e.g., [14]) suggest that random edge weights have a marked effect on the geometry of the network. This has been rigorously established in various works [3, 8–10], in the specific situation of exponential edge weights.

In this paper, we study the behavior of the hopcount and minimal weight in the setting of random graphs with finite variance degrees and general continuous edge weights. Since in many applications, the distribution of edge weights is unknown, the assumption of general weights is highly relevant. From a mathematical point of view, working with general instead of exponential edge weights implies that our exploration process is non-Markovian. This is the first paper that studies FPP on random graph models in this general setting. In a forthcoming paper [11], we will show that, due to the flexible choice of degree distribution, our results carry over to various other random graph models, including rank-1 inhomogeneous random graphs as introduced in [13].

**Organization of this section.** We start by introducing the configuration model in Section 1.2, where we also state our main result, Theorem 1.2. In Section 1.3, we discuss a continuous-time branching process approximation, which is necessary to identify the limiting variables in Theorem 1.2; this identification is done in Theorem 1.3. In Section 1.4, we study some examples that allow us to relate our results to existing results in the literature. We close with Section 1.5 where we present a discussion of our results and pose some open problems.

Throughout this paper, we make use of the following standard notation. We let $\xrightarrow{\mathrm{a.s.}}$ denote convergence almost surely, $\xrightarrow{L^1}$ denote convergence in mean, $\xrightarrow{d}$ denote convergence in distribution, and $\xrightarrow{P}$ convergence in probability. For a
sequence of random variables \((X_n)_{n\geq 1}\), we write \(X_n = O_p(b_n)\) when \(|X_n|/b_n\) is a tight sequence of random variables, and \(X_n = o_p(b_n)\) when \(|X_n|/b_n \overset{p}{\rightarrow} 0\), as \(n \to \infty\). We write \(D \sim F\) to denote that the random variable \(D\) has distribution function \(F\). For nonnegative functions \(n \mapsto f(n), n \mapsto g(n)\), we write \(f(n) = O(g(n))\) when \(f(n)/g(n)\) is uniformly bounded, and \(f(n) = o(g(n))\) when \(\lim_{n \to \infty} f(n)/g(n) = 0\). Furthermore, we write \(f(n) = \Theta(g(n))\) if \(f(n) = O(g(n))\) and \(g(n) = O(f(n))\). Finally, we say that a sequence of events \((\mathcal{E}_n)_{n\geq 1}\) occurs with high probability (w.h.p.) when \(\mathbb{P}(\mathcal{E}_n) \to 1\).

1.2. Configuration model and main result. The configuration model (CM) is a random graph with vertex set \([n] := \{1, 2, \ldots, n\}\) and with prescribed degrees. Let \(d = (d_1, d_2, \ldots, d_n)\) be a given degree sequence, that is, a sequence of \(n\) positive integers with total degree

\[
\ell_n = \sum_{i \in [n]} d_i,
\]

assumed to be even. The CM on \(n\) vertices with degree sequence \(d\) is constructed as follows: start with \(n\) vertices and \(d_i\) half-edges adjacent to vertex \(i \in [n]\). Randomly choose pairs of half-edges and match the chosen pairs together to form edges. Although self-loops may occur, these become rare as \(n \to \infty\) (see e.g. [12, 26]). We denote the resulting graph on \([n]\) by \(\text{CM}_n(d)\), with corresponding edge set \(\mathcal{E}_n\).

Regularity of vertex degrees. Let us now describe our regularity assumptions on the degree sequence \(d\) as \(n \to \infty\). We denote the degree of a uniformly chosen vertex \(V\) in \([n]\) by \(D_n = d_V\). The random variable \(D_n\) has distribution function \(F_n\) given by

\[
F_n(x) = \frac{1}{n} \sum_{j \in [n]} \mathbb{1}_{\{d_j \leq x\}},
\]

where \(\mathbb{1}_A\) denotes the indicator of the event \(A\). We write \(\log(x)_+ = \log(x)\) for \(x \geq 1\) and \(\log(x)_+ = 0\) for \(x \leq 1\). Then our regularity condition is as follows:

**Condition 1.1** (Regularity conditions for vertex degrees).

(a) Weak convergence of vertex degree. There exists a cumulative distribution function \(F\) of a discrete random variable \(D\), taking values in \(\mathbb{N}\) such that

\[
\lim_{n \to \infty} F_n(x) = F(x),
\]

for any continuity point \(x\) of \(F\); that is, \(D_n \overset{d}{\rightarrow} D\).
(b) Convergence of second moment.

\[ \lim_{n \to \infty} \mathbb{E}[D_n^2] = \lim_{n \to \infty} \frac{1}{n} \sum_{j \leq n} d_j^2 = \mathbb{E}[D^2], \]

where \( D_n \) and \( D \) have distribution functions \( F_n \) and \( F \), respectively, and we assume that

\[ \nu = \mathbb{E}[D(D - 1)]/\mathbb{E}[D] > 1. \]

(a) Uniform \( X^2 \log X \)-condition. For every \( K_n \to \infty \),

\[ \lim_{n \to \infty} \mathbb{E}[D_n^2 \log(D_n/K_n)_+] = \lim_{n \to \infty} \frac{1}{n} \sum_{j \leq n} d_j^2 \log(d_j/K_n)_+ = 0. \]

By Condition 1.1(c), the random degree \( D_n \) satisfies a uniform \( X^2 \log X \)-condition. The degree of a vertex incident to a half-edge that is chosen uniformly at random from all half-edges has the same distribution as the random variable \( D^*_n \) given by

\[ F^*_n(x) = \mathbb{E}[D_n \mathbb{1}_{D_n \leq x}]/\mathbb{E}[D_n], \quad x \in \mathbb{R}, \]

which is the size-biased version of \( D_n \). The latter random variable satisfies a uniform \( X \log X \)-condition if and only if \( D_n \) satisfies a uniform \( X^2 \log X \)-condition. As explained in more detail in Section 1.3 below, \( D^*_n \) is closely related to a branching-process approximation of neighborhoods of a uniform vertex, and Condition 1.1(c) implies that this branching process satisfies a uniform \( X \log X \)-condition. By uniform integrability, Condition 1.1(c) follows from the assumption that \( \lim_{n \to \infty} \mathbb{E}[D_n^2 \log(D_n/K_n)] = \mathbb{E}[D^2 \log(D)_+] \). Further, Condition 1.1(c) implies that \( \mathbb{E}[D^2 \log(D)_+] < \infty \).

Note that Conditions 1.1(a) and (c) imply that \( \mathbb{E}[D_n^i] \to \mathbb{E}[D^i], i = 1, 2. \) When the degrees are random themselves, then the distribution function \( F_n \) as well as the left-hand side of (1.4) and (1.6), are random and we assume that the convergence in (1.3), (1.4) and (1.6) to the respective (deterministic) right-hand sides holds in

\[ \text{Indeed, we claim that if } \lim_{n} \mathbb{E}[D^2 \log(D_n/D_n)_+] = 0 \text{ for every } K_n \to \infty \text{, and if } \mathbb{E}[D_n^2] \to \mathbb{E}[D^2], \text{ then } \lim_{n} \mathbb{E}[D^2 \log(D_n)] < \infty \text{ holds. To see this, note that } \mathbb{E}[D_n^2] \to \mathbb{E}[D^2] \text{ implies that for some } M, \text{ uniformly in } n, \text{ we have } \mathbb{E}[D_n^2] \leq M. \]

If \( \mathbb{E}[D_n^2 \log(D_n)] \to \infty \), then for all \( a > 0 \) : \( \mathbb{E}[D_n^2 \log(D_n)] > a \), for all \( n \geq n_0(a) \). Take \( K_n = a/M \), then for \( n \geq n_0(a) \)

\[ \mathbb{E}[D_n^2 \log(D_n/K_n)_+] \geq \mathbb{E}[D_n^2 \log(D_n/D_n)_+] \geq \mathbb{E}[D_n^2 \log(D_n)] - \mathbb{E}[D_n^2 \log(K_n)] \]

\[ > a - M \cdot \frac{a}{2M} = a/2, \]

which contradicts our assumption that \( \lim_{n} \mathbb{E}[D_n^2 \log(D_n/K_n)_+] = 0. \)
probability. Thus, in this case, we require that, with $E_n[D_n^i] = \frac{1}{n} \sum_{j \in [n]} d_j^i$ (which is now a random variable) and for every $\varepsilon > 0$ and $i \in \{1, 2\}$,
\[
\lim_{n \to \infty} P(\left| F_n(x) - F(x) \right| \geq \varepsilon) = 0, \quad \forall x \in \mathbb{R},
\]
(1.8)
\[
\lim_{n \to \infty} P(\left| E_n[D_n^i] - E[D^i] \right| \geq \varepsilon) = 0.
\]
A similar condition replaces (1.6).

Condition (1.5) is equivalent to the existence of a giant component in $\text{CM}_n(d)$; see, for example, [28, 35, 36]. Let $F$ be a distribution function of a random variable $D$, satisfying (1.5) and $E[D^2 \log(D)_+] < \infty$. We give two canonical examples in which Condition 1.1 holds. The first is when there are precisely $n_k = \lceil nF(k) \rceil - \lceil nF(k-1) \rceil$ vertices having degree $k \geq 1$. The second is when $(d_i)_{i \in [n]}$ is an i.i.d. sequence of random variables with distribution function $F$ (in the case that $\sum_{i \in [n]} d_i$ is odd, we increase $d_n$ by 1, this does not affect the results).

Edge weights and shortest paths. Once the graph has been constructed, we attach an edge weight $\xi_e$ to every edge $e$, where $(\xi_e)_{e \in E}$ are i.i.d. continuous random variables with density $f_{\xi}: [0, \infty) \to [0, \infty)$ and corresponding distribution function $F_{\xi}$. Pick two vertices $U_1$ and $U_2$ at random from $[n]$ and let $\Gamma_{12}$ denote the set of all paths in $\text{CM}_n(d)$ between these two vertices. For any path $\pi \in \Gamma_{12}$, the weight of the path is defined as
\[
\sum_{e \in \pi} \xi_e.
\]
Let
\[
L_n = \min_{\pi \in \Gamma_{12}} \sum_{e \in \pi} \xi_e,
\]
denote the weight of the optimal (i.e., minimal weight) path between $U_1$ and $U_2$ and let $H_n$ denote the number of edges or the hopcount of this path. If the two vertices are in different components of the graph, then we let $L_n, H_n = \infty$. We are ready to state our main result. Due to the complexity of the various constants and limiting random variables arising in the theorem, we defer their complete description to the next section.

**Theorem 1.2 (Joint convergence of hopcount and weight).** Consider the configuration model $\text{CM}_n(d)$ with degrees $d$ satisfying Condition 1.1, and with i.i.d. edge weights distributed according to the continuous distribution $F_{\xi}$. Then, there exist constants $\alpha, \gamma, \beta \in (0, \infty)$ and $\alpha_n, \gamma_n$ with $\alpha_n \to \alpha, \gamma_n \to \gamma$, such that the hopcount $H_n$ and weight $L_n$ of the optimal path between two uniformly chosen vertices, conditioned on being connected, satisfy
\[
\left( \frac{H_n - \gamma_n \log n}{\sqrt{\beta \log n}}, L_n - \frac{1}{\alpha_n} \log n \right) \xrightarrow{d} (Z, Q),
\]
(1.11)
as \( n \to \infty \), where \( Z \) and \( Q \) are independent and \( Z \) has a standard normal distribution, while \( Q \) has a continuous distribution.

This is the first time that FPP on sparse random graphs with general edge weights has been studied; for the particular case where the edge weights have an exponential distribution; see, for example, [9].

In Remark 1.4 below, we will state conditions that imply that we can replace \( \alpha_n \) and \( \gamma_n \) by their limits \( \alpha \) and \( \gamma \), respectively. Theorem 1.2 shows a remarkable degree of universality. For \( CM_n(d) \) satisfying Condition 1.1, the hopcount always satisfies a central limit theorem with mean and variance proportional to \( \log n \). Also, the weight of the shortest weight path between two uniformly chosen vertices always is of order \( \log n \), and the fluctuations around \( \log n / \alpha_n \) converge in distribution. We will see that even the limit \( Q \) has a large degree of universality.

For this, as well as to define the parameters \( \alpha, \alpha_n, \beta, \gamma, \gamma_n \), we first need to describe a continuous-time branching process approximation.

1.3. Continuous-time branching processes. In this section, we define the limiting continuous-time branching process (CTBP) that describes the neighborhood structure of first passage percolation on \( CM_n(d) \). Define the size-biased distribution \( F^* \) of the random variable \( D \) with distribution function \( F \) by

\[
F^*(x) = \frac{\mathbb{E}[D1_{\{D \leq x\}}]}{\mathbb{E}[D]}, \quad x \in \mathbb{R}.
\]

When Condition 1.1(a)–(b) holds, the function \( F^* \) is the distributional limit as \( n \to \infty \) of \( F^*_n \) in (1.7). Now let \( (BP^*(t))_{t \geq 0} \) denote the following CTBP: (a) At time \( t = 0 \), we start with one individual which we refer to as the original ancestor or the root of the branching process.

(b) Each individual \( v \) in the branching process lives for a random amount of time which has distribution \( F_\xi \), that is, the edge weight distribution, and then dies. At the time of death, the individual gives birth to \( D^*-1 \) children, where \( D^* \sim F^* \). Lifetimes and number of offspring across individuals are independent.

Note that in the above construction, by Condition 1.1(b), if we let \( X_v = D^*-1 \) be the number of children of an individual, then the expected number of children satisfies

\[
\mathbb{E}[X_v] = \mathbb{E}[D^*-1] = \nu > 1.
\]

Further, by Condition 1.1(c), for \( D^* \sim F^* \) (recall footnote 3),

\[
\mathbb{E}[D^* \log (D^*)_+] < \infty.
\]

The CTBP defined above is a standard Bellman–Harris process, with lifetime distribution \( F_\xi \) and offspring distributed as \( D^*-1 \) [4, 21, 24]. The Malthusian parameter \( \alpha \) of the branching process \( BP^* \) is the unique solution of the equation

\[
\hat{\mu}(\alpha) = \nu \int_0^\infty e^{-\alpha t} dF_\xi(t) = 1.
\]
Since \( \nu > 1 \), we obtain that \( \alpha \in (0, \infty) \). We also let \( \alpha_n \) be the solution to (1.15) with \( \nu \) replaced by \( \nu_n \).

(1.16) \[
\nu_n = \frac{\mathbb{E}[D_n(D_n - 1)]}{\mathbb{E}[D_n]}.
\]

Clearly, \( \alpha_n \to \alpha \) when Condition 1.1 holds, and further \( |\alpha_n - \alpha| = O(|\nu_n - \nu|) \).

Standard theory (see, e.g., [4, 21, 24]) implies that under our assumptions on the model, namely (1.13) and (1.14), there exists a random variable \( \mathcal{W}^* \) such that

(1.17) \[
e^{-\alpha t} |\text{BP}^*(t)| \xrightarrow{\text{a.s., } L^1} \mathcal{W}^*.
\]

Here, since \( \mathbb{E}[D^* \log(D^*)] < \infty \), the limiting random variable \( \mathcal{W}^* \) satisfies \( \mathcal{W}^* > 0 \) a.s. on the event of nonextinction of the branching process and is zero otherwise. Thus, \( \alpha \) measures the true rate of exponential growth of the branching process.

By (1.15), we can define the cumulative distribution function \( \tilde{F}_\xi \), often referred to as the stable-age distribution, as

(1.18) \[
\tilde{F}_\xi(x) = v \int_0^x e^{-\alpha y} dF_\xi(y).
\]

Let \( \tilde{v} \) be the mean and \( \tilde{\sigma}^2 \) the variance of \( \tilde{F}_\xi \), that is,

(1.19) \[
\tilde{v} = v \int_0^\infty x e^{-\alpha x} dF_\xi(x), \quad \tilde{\sigma}^2 = v \int_0^\infty (x - \tilde{v})^2 e^{-\alpha x} dF_\xi(x).
\]

Then \( \tilde{v}, \tilde{\sigma}^2 \in (0, \infty) \), since \( \alpha > 0 \). We also define \( \tilde{F}_{n,\xi} \) to be the distribution function \( \tilde{F}_\xi \) in (1.18) with \( v \) and \( \alpha \) replaced by \( \nu_n \) and \( \alpha_n \), and we let \( \tilde{v}_n \) and \( \tilde{\sigma}_n^2 \) be the corresponding mean and variance.

We need a small variation of the above standard CTBP, where the root of the branching process dies immediately giving birth to a \( D \) number of children where \( D \) has distribution \( F \), the original (i.e., non-size-biased) degree distribution as in Condition 1.1(a). The details for every other individual in this branching process remain unchanged from the original description, namely each individual survives for a random amount of time with distribution \( F_\xi \) giving rise to a \( D^* - 1 \) number of children where \( D^* \sim F^* \), the size-biased distribution function \( F^* \) as in (1.12). Writing \( |\text{BP}(t)| \) for the number of alive individuals at time \( t \), it is easy to see here as well that

(1.20) \[
e^{-\alpha t} |\text{BP}(t)| \xrightarrow{\text{a.s., } L^1} \tilde{\mathcal{W}}.
\]

Further, conditionally on \( D = k \),

(1.21) \[
\tilde{\mathcal{W}} \overset{d}{=} \tilde{\mathcal{W}}^{*(1)} + \cdots + \tilde{\mathcal{W}}^{*(k)},
\]

where \( D \sim F \), and \( \tilde{\mathcal{W}}^{*(i)} \) are i.i.d. with the distribution of the limiting random variable in (1.17). Let \( \mathcal{W} \) denote a random variable distributed as \( \tilde{\mathcal{W}} \) conditioned to be positive, that is, for every \( x \geq 0 \),

(1.22) \[
\mathbb{P}(\mathcal{W} \leq x) = \mathbb{P}(\tilde{\mathcal{W}} \leq x \mid \tilde{\mathcal{W}} > 0).
\]
To simplify notation in the sequel, we will use \((\text{BP}(t))_{t \geq 0}\) to denote a CTBP with the root having offspring either one (as for the standard CTBP), \(D\) or \(D^* - 1\). It will be clear from the context which setting we are in.

We are now in a position to identify the limiting random variable \(Q\) as well as the parameters \(\alpha, \beta, \gamma, \alpha_n, \gamma_n\):

**Theorem 1.3 (Identification of the limiting variables).** The parameters \(\alpha, \alpha_n, \beta, \gamma_n, \gamma\) in Theorem 1.2 satisfy that \(\alpha\) is the Malthusian rate of growth defined in (1.15) and \(\alpha_n\) is the solution to (1.15) with \(\nu_n\) replacing \(\nu\), while

\[
\gamma = \frac{1}{\bar{\nu}}, \quad \gamma_n = \frac{1}{\alpha_n \bar{\nu}_n}, \quad \beta = \frac{\bar{\sigma}^2}{\bar{\nu}^3 \alpha}.
\]

Further, \(Q\) can be identified as

\[
Q = \frac{1}{\alpha}(-\log \mathcal{W}^{(1)} - \log \mathcal{W}^{(2)} - \Lambda + c),
\]

where \(\mathbb{P}(\Lambda \leq x) = e^{-e^{-x}}\) (so that \(\Lambda\) is a standard Gumbel random variable), \(\mathcal{W}^{(1)}, \mathcal{W}^{(2)}\) are two independent copies of the variable \(\mathcal{W}\) in (1.21), also independent from \(\Lambda\), and \(c\) is the constant

\[
c = \log(\mathbb{E}[D](\nu - 1)^2 / (\nu \bar{\nu} \bar{\sigma}^2)).
\]

**Remark 1.4 (Asymptotic mean).** In (1.11), we can replace \(\alpha_n\) and \(\gamma_n\) by their limits \(\alpha\) and \(\gamma\) precisely when \(\gamma_n = \gamma + o(1/\sqrt{\log n})\) and \(\alpha_n = \alpha + o(1/\log n)\). Since \(|\alpha_n - \alpha| = O(|\nu_n - \nu|), |\bar{\nu}_n - \bar{\nu}| = O(|\nu_n - \nu|)\), these conditions are equivalent to \(\nu_n = \nu + o(1/\sqrt{\log n})\) for \(\gamma_n\) and \(\nu_n = \nu + o(1/\log n)\) for \(\alpha_n\), respectively.

Theorem 1.3 implies that also the random variable \(Q\) is remarkably universal, in the sense that it always involves two independent martingale limit variables corresponding to the branching processes, and a Gumbel distribution.

Let \(L_n(i)\) denote the weight of the \(i\)th shortest path, so that \(L_n = L_n(1)\), and let \(H_n(i)\) denote its length. Further, let \(\bar{H}_n(i)\) and \(\bar{L}_n(i)\) denote the re-centered and normalized quantities as in Theorem 1.2. The same proof for the optimal path easily extends to prove asymptotic results for the joint distribution of the weights and hopcount of these ranked paths. To keep the study to a manageable length, we shall skip a proof of this easy extension.

**Theorem 1.5 (Multiple paths).** Under the conditions of Theorem 1.2, for every \(m \geq 1\),

\[
((\bar{H}_n(i), \bar{L}_n(i)))_{i \in [m]} \xrightarrow{d} ((Z_i, Q_i))_{i \in [m]},
\]
as $n \to \infty$, where $Z_1, \ldots, Z_m$ are i.i.d. standard normals independently of $(Q_i)_{i \in [m]}$, with

$$Q_i = \frac{1}{\alpha} (-\log \mathcal{W}^{(1)} - \log \mathcal{W}^{(2)} - \Lambda_i + c),$$

where $(\Lambda_i)_{i \in [m]}$ are the ordered (minimal) points of an inhomogeneous Poisson point process with intensity $\lambda(t) = e^t$.

1.4. Examples. We treat some examples of edge weight distributions that have appeared in the literature and have been treated via distribution-specific techniques.

(i) We start with exponential edge weights [3, 8–10]. In this case, it is immediate from (1.15) and (1.18) that

$$\alpha = \nu - 1, \quad \tilde{\nu} = \tilde{\sigma} = 1/\nu,$$

hence Theorems 1.2–1.3 show that $H_n$ converges to a normal distribution, with asymptotic mean and asymptotic variance both equal to $\frac{\nu}{\nu - 1} \log n$. Furthermore, Theorem 1.2 induces the convergence of the minimal weight in [9, 10]. Observe that the random variable $M$, which appears in [9], (C.19), is equal to $-\Lambda$. In [10], the special case of the Erdős–Rényi random graph with exponential edge weights was tackled. There is a small error in the expression of the limiting random variable in [10], (4.16).

(ii) By studying weights of the form $\xi_e = 1 + E_e/k$, where $(E_e)$ are i.i.d. exponentials with mean 1, and consecutively sending $k \to \infty$, one would expect to obtain results which are close to limiting results on the graph distance between a pair of uniformly chosen vertices in $[n]$, conditioned to be connected. Indeed, the results match up nicely with those in [18] for the Norros–Reittu model and [22] for the CM. For the sake of brevity, we leave the derivation to the reader.

(iii) As a third example one can consider the CM with fixed degrees $r$, and where each edge is given an edge weight $E^s$, $s > 0$, where $E \sim \text{Exp}(1)$, a variant of the weak disorder models in statistical physics [14]. One can formally consider the case $r = n - 1$, although this does not satisfy the conditions of our theorem. Here the CM with fixed degrees $n - 1$ resembles a complete graph on $[n]$ and the results match up nicely with those in [7], namely, a central limit theorem for $H_n$ with asymptotic mean $s \log n$ and asymptotic variance $s^2 \log n$, while $n^{s} [L_n - \frac{1}{\lambda} \log n]$ converges in distribution, where $\lambda = \Gamma(1 + 1/s)^s$. We refer to [7] for the derivation of these parameters.

1.5. Discussion. In this section, we give a brief discussion of our results, possible extensions and open problems.
(a) **Universality.** Our results are universal in the sense that, as Theorems 1.2–1.3 demonstrate, the CLT for the hopcount depends only on the first two moments of the size-biased offspring distribution and on the edge weight distribution, but not on any other property of the network model. Further, the form of the limit random variable for the minimal weight has a universal form in terms of the martingale limits of branching processes and a Gumbel random variable. Our conditions on the edge weights \((\xi_e)_{e \in E_n}\) are very weak, giving a highly universal result. In many other settings, such as lattice models or random graphs with infinite-variance degrees (see the next paragraph), such strong universality is not valid. The same robustness arises in the life-time distribution of age-dependent branching processes (see the books by Harris or Jagers [21, 24]).

(b) **Infinite-variance configuration model.** In [9], we have investigated the CM with exponential edge weights, but with i.i.d. degrees with \(P(D \geq x) \sim cx^{-(\tau-1)}\) and \(\tau \in (2, 3)\), so that \(\mathbb{E}[D^2] = \infty\). In this case, the result for \(L_n\) is markedly different, in the sense that \(L_n\) converges in distribution without re-centering. Further, \(H_n\) satisfies a central limit with asymptotic mean and variance equal to a multiple of \(\log n\). It would be of interest to investigate whether \(H_n\) always satisfies a central limit theorem, and, if so, whether the order of magnitude of its variance is always equal to that of its mean. See [5] for results showing that there are more universality cases in the infinite-variance setting.

(c) **The \(X \log X\)-condition.** In Condition 1.1(c), we assume that the degrees satisfy a second moment condition with an additional logarithmic factor. This is equivalent to the CTBP satisfying an \(X \log X\)-condition (uniformly in the size \(n\) of the graph). It would be of interest to investigate what happens when this condition fails.

(d) **Flooding and diameter.** In [3], the flooding time and diameter have been investigated in the context of the CM with exponential edge weights. These are \(\max_{j \in [n]}: L_n(U_1, j) < \infty\), \(L_n(U_1, j)\), respectively \(\max_{i, j \in [n]}: L_n(i, j) < \infty\), \(L_n(i, j)\), where \(L_n(i, j)\) is the minimal weight between the vertices \(i\) and \(j\) and \(U_1\) is, as before, a randomly selected vertex. It would be of interest to investigate the flooding time for general edge weights.

(e) **Superconcentration and chaos.** Analogous to various problems in statistical physics such as random polymers or FPP on the lattice, our results suggest that the FPP optimal path problem is chaotic. This means that there exists \(\varepsilon_n \to 0\) such that refreshing a fraction \(\varepsilon_n\) of the edge weights with new random variables with the same distribution would entirely change the actual optimal path, in the sense that the new optimal path would be “almost” disjoint from the original optimal path; see, for example, [15]. Such questions have also arisen in computer science wherein one is interested in judging the “importance” and fair price of various edges in the optimal path; if an edge being deleted causes a large change in the cost of the new optimal path, then that edge is deemed very valuable. These form the basis of various “truth and auction mechanisms” in computer science (see, e.g., [6, 19, 34]).
Universality of First Passage Percolation 2579

(f) Pandemics, gossip and other models of diffusion. First passage percolation models as well as models using FPP as a building block have started to play an increasingly central role in the applied probability community in describing the flow of materials, ranging from viral epidemics [16], gossip algorithms [2] and more general finite Markov interchange processes [1]. Models with more general edge weight distributions have also arisen in understanding the flow of information and reconstruction of such information networks in sociology and computer science; see [31, 32] for examples in this vast field.

Organization of this paper. In Section 2, we describe the coupling between the first-passage percolation neighborhoods in CM_{n(d)} and a CTBP. In Section 3, we state our main technical result that describes a Poisson process limit for the occurrence of short paths between U_1 and U_2 which then proves our main theorem. In Section 4, we extend results for CTBPs, as proved in [21, 24, 38], to the case of infinite-variance offspring distributions, using truncation. In Section 5, we prove bounds on our coupling. In Section 6, we give a novel proof of the asymptotics for the number of alive individuals in a CTBP in a given generation and with a given residual lifetime. This proof is tailored to deal with CTBPs observed till some time t that have an offspring distribution that depends on n where n → ∞ and t = t_n → ∞ simultaneously. In Section 7, we prove our main technical result on the Poisson process limit.

2. Coupling. In this section, we describe a coupling between FPP on CM_{n(d)} and continuous-time branching processes. We start with an informal description.

2.1. Informal description of shortest weight trees. The model CM_{n(d)} with edge set E_n together with i.i.d. lengths (also referred to as weights) (ξ_e)_{e ∈ E_n} on the edges was introduced in Section 1.2. Here, ξ_e ∼ F_ξ, with density f_ξ. Our ultimate goal is to calculate the limit distribution of the hopcount H_n and weight L_n of the shortest path between a uniformly chosen pair of connected vertices U_1 and U_2, when Condition 1.1 is satisfied.

To obtain a proper understanding of the shortest path between two vertices, we imagine a liquid that percolates through the edges of the CM at rate one. We start percolating the liquid simultaneously from both vertices U_1 and U_2 and we interpret the edge weight ξ_e on edge e as the distance between the two vertices incident to e. For any t ≥ 0, the set of half-edges that are currently being wetted by the liquid, as well as the residual time to completely wet them, starting from U_i will be informally denoted as the shortest weight tree SWT(i), i = 1, 2. A precise definition of these SWTs will be given in the next section. When the liquid has reached two vertices that are incident to a connecting edge between the two SWTs, then a possible shortest path has been found. Since at that moment the connecting edge has not yet been filled, we cannot be sure whether the given path between U_1 and U_2 is
indeed the shortest one. Hence, we have to find all connecting edges between the two SWTs and take the minimum of all these path weights to determine \( L_n \) and \( H_n \).

In the mathematical description in the next section, we build the CM simultaneously with the liquid percolating through the edges. Since we will construct the process sequentially, it is easier to index the sequence of new edge-weights added to the system as \( (\xi_j)_{j \geq 1} \). The half-edges emanating from the wetted vertices are called the \textit{alive} half-edges \( \text{AH}(t) \) at time \( t \). During the building process, we form two SWTs consisting of “alive” half-edges and vertices attached to \( U_i \), for \( i = 1, 2 \). In order to perform the building process properly, we put the i.i.d. weights \( (\xi_j)_{j \geq 1} \) on the half-edges instead of on the edges. Technically, one has to be extremely careful in constructing the process in this fashion. Imagine a situation where the liquid reaches both \( a \) and \( b \) for some edge \( e \) formed by merging the half-edges \( e = (a, b) \). Assigning independent half-edge weights \( \xi_a \) and \( \xi_b \) is then \textit{not} the same as first passage percolation. Instead we put the weight on the half-edge that is found \textit{first} by the liquid. We initiate the construction by putting weights \( \xi_1, \ldots, \xi_{dU_1} \) on the half-edges incident to \( U_1 \) and weights \( \xi_{dU_1+1}, \ldots, \xi_{dU_1+dU_2} \) on the half-edges incident to \( U_2 \). Of course, this creates a problem when these half-edges are paired to one another, which we have to take into account properly.

We construct a sequence of epoch times \( (T_k)_{k \geq 0} \) that track when a decision has to be made. Start with \( T_0 = 0 \) and wait until the end point of the first of the \( dU_1 + dU_2 \) half-edges is reached. This time is called \( T_1 \) and successive times at which further end points of “alive” half-edges are reached are called \( T_2, T_3, \ldots \). At \( t = T_1 \), we pair the exhausted half-edge, which we call \( r_1 \), with one of the \( \ell_n - 1 \) other half-edges at random; the found half-edge is called \( Pr_1 \). The formed edge \( (r_1, Pr_1) \) receives the weight of the exhausted half-edge \( r_1 \) and we connect the siblings of \( Pr_1 \) to the newly found vertex. The sibling half-edges receive i.i.d. weights from the sequence \( (\xi_j)_{j \geq 1} \), whereas the weights of the other “alive” half-edges are updated by subtracting \( T_1 \). We repeat the whole procedure by finding the minimum of the “alive” half-edges, and after adding this minimal weight to \( T_1 \) we find the second epoch time \( T_2 \). We continue this procedure until all half-edges are attached to one of the SWTs.

In general, the formed edge \( (r_1, Pr_1) \) at time \( t = T_1 \) receives a weight with the correct distribution. However, this only occurs when \( r_1 \) pairs with one of the \( \ell_n - (dU_1 + dU_2) \) so-called “free” half-edges, that is, half-edges connected to vertices which are not yet wetted. When \( r_1 \) pairs with one of its \( dU_1 - 1 \) sibling half-edges (and hence a self-loop occurs) or when \( r_1 \) pairs with one of the \( dU_{3-i} \) half-edges incident to vertex \( U_{3-i} \) (and hence a “collision” edge occurs), then we do not know what weight we should assign to the self-loop or collision edge, because the half-edge to which \( r_1 \) is paired is an alive half-edge and already had a weight. In order to resolve this issue, in the next section, we shall make sure that a weight is only assigned to \textit{one} of the half-edges of an edge. This will be achieved by first investigating whether a half-edge is paired to a “free” half-edge or to an “alive”
half-edge. In particular, we will change the order in which half-edges are paired. We are free to pair half-edges in any order we like and this property is used to remove self-loops, edges that close a cycle and collision edges beforehand. This way, all paths receive weights with the correct distribution, and after the completion of the entire construction we take the minimum of all connecting paths to find \( L_n \) and \( H_n \).

The removal of self-loops, edges that close a cycle and collision edges is done at the epoch times \( T_0, T_1, T_2, \ldots \). Conditionally on the number of “alive” half-edges and “free” half-edges, we know the success probabilities of Bernoulli random variables that determine whether a pairing results in attaching to a “free” half-edge or to an “alive” half-edge.

We will couple \( \text{SWT}(\cdot) \) both to an \( n \)-dependent continuous-time branching process (CTBP) denoted by \( \text{BP}_n(\cdot) \) and to a CTBP \( \text{BP}(\cdot) \). This results in a coupling \( (\text{SWT}(t), \text{BP}_n(t), \text{BP}(t))_{t \geq 0} \) on the same probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \), where \( \text{SWT}(t) \) consists of the alive half-edges that are connected to \( U_1 \) and \( U_2 \) by paths of weights at most \( t \), as well as their residual lifetimes, while \( \text{BP}_n(t) \) contains the same information for the \( n \)-dependent CTBP and \( \text{BP}(t) \) for the \( n \)-independent CTBP.

Since there are a number of ingredients in this coupling, let us start by giving the reader an intuitive mental picture of the key actors in this coupling. In the first step of the coupling, which is explained in full detail in Section 2.2, we couple the forward degrees in the \( \text{SWT}(t) \) to the number of offspring in the branching processes \( \text{BP}_n(t) \) and \( \text{BP}(t) \). On one and the same probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \), we introduce sequences of random variables \( (B^{(n)}_k)_{k \geq 1}, (B_k)_{k \geq 1}, (Y^{(n)}_k)_{k \geq 1}, \) and \( (X^{(n)}_k)_{k \geq 1} \), and a sequence of stopping times \( (\tau_k)_{k \geq 1} \).

The sequences \( (B^{(n)}_k)_{k \geq 1} \) and \( (B_k)_{k \geq 1} \) are i.i.d. and will be used as the number of offspring in a branching process, where the first branching process depends on \( n \), while the second one is independent of \( n \). In the coupling, there is a strong dependence between \( B^{(n)}_k \) and \( B_k \) for any \( k \). The sequence \( (Y^{(n)}_k)_{k \geq 1} \) will correspond to the sequence of forward degrees, that is, the degree minus one, as the liquid in \( \text{SWT} \) percolates through the graph, while \( X^{(n)}_k \) will be equal to \( Y^{(n)}_k \) minus the number of pairings that result in either a self-loop, a cycle or a collision edge. Using the stopping times \( \tau_k \), the \( k \)th variable \( X^{(n)}_k \) will be successfully coupled to \( B^{(n)}_{\tau_k} \) precisely when \( X^{(n)}_k = Y^{(n)}_k \) and \( Y^{(n)}_k = B^{(n)}_{\tau_k} \).

Since this coupling is not perfect, in the second step of the coupling performed in Section 2.3, we discuss the evolution of the processes \( (\text{SWT}(t), \text{BP}_n(t), \text{BP}(t))_{t \geq 0} \), including the evolution of the children of the alive half-edges that are miscoupled. Finally, in Section 2.4, we state the main bounds on our coupling. After this high-level explanation, let us now give the details of our coupling construction.

2.2. Coupling forward degrees of the SWT. In this section, we will give a precise definition of the coupling of the forward degrees in the SWT to the number
of offspring in an \( n \)-dependent continuous-time branching process \( \text{BP}_n \) and secondly to that in a continuous-time branching process \( \text{BP} \) that is independent of \( n \). As mentioned above, our purpose is to couple the forward degrees \( (Y^{(n)}_k)_{k \geq 1} \) and the reduced forward degrees \( (X^{(n)}_k)_{k \geq 1} \) to the number of offspring \( (B^{(n)}_k)_{k \geq 1} \) and \( (B_k)_{k \geq 1} \). Especially the definition of \( X^{(n)}_k \) involves additional work that unfortunately makes the coupling less transparent. In defining \( X^{(n)}_k \), we will use that pairing half-edges in \( \text{CM}_n(d) \) can be done in any order. More specifically, to decide whether one of the \( Y^{(n)}_k \) half-edges closes a cycle or connects the two SWT’s, we use a test involving Bernoulli random variables. After completing this test for all sibling half-edges involved, we can determine the number of half-edges that resulted in closing a cycle or in a connecting edge, and hence define \( X^{(n)}_k \) equal to \( Y^{(n)}_k \) minus this number. We now start with a procedure which abstractly formalizes the way how we count the number of occurring self-loops, edges that close a cycle or collision edges.

The idea is as follows. Consider a partition of a set \([m]\) into \( A \) and \( B = A^c \). One can achieve a uniform draw from \([m]\) in two steps, first by performing a Bernoulli experiment with success probability \( p_A = |A|/m \); if the outcome of this experiment is one, then we draw an object uniformly from \( A \), otherwise we draw a uniform object from \( B \). In fact, we do not even have to actually draw the latter uniform element. We will perform such a construction repeatedly, with \( A \) denoting the set of alive half-edges at appropriate stopping times and where the set \([m]\) is recursively defined. When doing so, we can think of this as “testing” whether a half-edge creates a self-loop, cycle or collision edge or whether it connects to a “free” half-edge. This is formalized in the following procedure. We first need to set up some notation.

The procedure takes as its input one nonempty set, the “alive” set \( \text{AS} = \{a_1, a_2, \ldots, a_s\} \) having \( s \) elements, and the size of the “free” set \( \text{N} \). The elements \( a_j \in \text{AS}, j = 1, 2, \ldots, r \) for \( r \leq s \) are special. We will view them as a list (namely an ordered set) \( \text{TS} = (a_1, a_2, \ldots, a_r) \), of \( r \) elements. Abusing notation, the procedure initializes with \( \text{AS}(0) := \text{AS}, \text{TS}(0) := \text{TS} \) and \( \text{N}(0) := \text{N} \) and we will sequentially update these sets and this number as follows using a sequence of conditionally independent Bernoulli random variables \( (\beta_i)_{1 \leq i \leq r} \) and a sequence of sets \( (\text{S}_i)_{i \geq 1} \):

**PROCEDURE 2.1 (Preprocessing the matchings).**

(a) **Initialization:** Define the success probability and set \( \text{S}_1 \) as

\[
(2.1) \quad p_1 = \frac{|\text{AS}(0)| - 1}{|\text{AS}(0)| + \text{N}(0) - 1} \quad \text{and} \quad \text{S}_1 = \text{AS}(0) \setminus \{a_1\} = \{a_2, \ldots, a_s\}.
\]

Let \( \beta_1 \sim \text{Bernoulli}(p_1) \).

(i) If \( \beta_1 = 1 \), then select element \( b_1 \) uniformly at random from the set \( \text{S}_1 \) and update the sets as \( \text{AS}(1) = \text{AS}(0) \setminus \{a_1, b_1\} \) and if \( b_1 \in \text{TS}(0) \) then \( \text{TS}(1) = \text{TS}(0) \setminus \{a_1, b_1\} \), else \( \text{TS}(1) = \text{TS}(0) \setminus \{a_1\} \). We do not change \( \text{N}(1) := \text{N}(0) \).
(ii) If $\beta_1 = 0$, then we do not select any element from $S_1$ and we update the sets as $AS(1) = AS(0) \setminus \{a_1\}$ and $TS(1) = TS(0) \setminus \{a_1\}$, while $N(1) = N(0) - 1$.

(b) Recursion: For $k \geq 1$, we proceed recursively as above, taking the first element at the front of the list $TS(k)$ which, abusing notation, we still call $a_{k+1}$ [note, this element may not be $a_{k+1}$ in $TS(0)$, if that element is already drawn at a previous time step] now defining

\begin{equation}
\beta_{k+1} = \frac{|AS(k)| - 1}{|AS(k)| + N(k) - 1} \quad \text{and} \quad S_{k+1} = AS(k) \setminus \{a_k\},
\end{equation} 

generating $\beta_{k+1} \sim \text{Bernoulli}(p_{k+1})$ and proceeding as in (i) and (ii) above, with $a_1, b_1, S_1$ replaced by $a_{k+1}, b_{k+1}, S_{k+1}$, respectively.

(c) Termination: We stop when the list equals the empty set, that is, $TS(k) = \varnothing$.

Let us give a brief and informal sketch of how we use Procedure 2.1. Consider the informal description in Section 2.1 of the liquid percolating through a network started from two vertices $U_1$ and $U_2$ simultaneously. Assume that at some time $t$, the liquid from vertex $U_i$, where $i = 1, 2$, hits a new vertex $V$. These to half-edges incident to $V$, called $HE_V$, except the one used by $V$ to connect to $SWT(i)$, are now deemed active since the flow has encountered this new vertex $V$. Write $AS = AH^{(1)}(t) \cup AH^{(2)}(t) \cup HE_V$ for the collection of alive half-edges at this time, and $TS = (HE_V)$ for the set of half edges incident to $V$. Then the above procedure tests for each one of these newly added half-edges whether it pairs to a half-edge in $AH(t)$, which corresponds to the “$\beta_k = 1$” events, or instead connects to a new half-edge not in $SWT^{(1)}(t) \cup SWT^{(2)}(t)$, which corresponds to the “$\beta_k = 0$” events. In the latter case, we actually do not connect the half-edge, but only record that the half-edge is paired to a free half-edge, and thus decrease the number of free half-edges $N(k)$ by 1. Further note that in each “$\beta_k = 1$” event, the new edge created could either be (a) a self-loop or cycle when the half-edge pairs to an alive half-edge in $SWT^{(i)}(t)$, namely the same cluster that sees $V$ for the first time or (b) arguably more importantly, creates a collision edge when it is paired to a half edge in $SWT^{(3-i)}(t)$, the other cluster. These collision edges are the ones that potentially create the shortest path.

Let us now turn to the precise definition of the probability space for the coupling of the forward degrees of $SWT$ and the associated branching processes. We start on one and the same probability space $(\Omega, \mathcal{F}, P)$ with the following ingredients:

(i) Two vertices $U_1$ and $U_2$ chosen at random from $[n]$;

(ii) Label the $\ell_n$ half-edges by $[\ell_n]$ with the half-edges of vertex 1 labeled $1, 2, \ldots, d_1$, the half-edges of vertex 2 labeled $d_1 + 1, \ldots, d_1 + d_2$, etc. We will require repeated draws with replacement from $[\ell_n]$, which results in an i.i.d. sequence $(\sigma_i)_{i \geq 1}$. We will also require a second sequence of i.i.d. draws $(\tilde{\sigma}_i)_{i \geq 1}$ that is independent of the draws $(\sigma_i)_{i \geq 1}$;
(iii) To each $\sigma_i$ and $\tilde{\sigma}_i$, we associate random variables $B_i^{(n)}$ and $\tilde{B}_i^{(n)}$ that correspond to the forward degree of the vertices incident to the half-edges $\sigma_i$ and $\tilde{\sigma}_i$. To each $B_i^{(n)}$ and $\tilde{B}_i^{(n)}$, we associate random variables $B_i$ and $\tilde{B}_i$, whose distributions only depend on $B_i^{(n)}$ and $\tilde{B}_i^{(n)}$ and not on any of the other randomness involved;

(iv) An i.i.d. sequence $(\xi_i)_{i \geq 1}$ of edge-weights with distribution $F_\xi$, and a second sequence of i.i.d. weights $(\tilde{\xi}_i)_{i \geq 1}$, also with distribution $F_\xi$, that is independent of the edge-weights $(\xi_i)_{i \geq 1}$;

(v) Recall that vertex $j \in [n]$ had degree $d_j$. Recall the uniform choices $b_1, b_2, \ldots$ in Procedure 2.1 modulated by the values of the Bernoulli sequence $\beta_1, \beta_2, \ldots$. Analogously, we construct random variables $(b_j(1), \ldots, b_j(d_j))$, taking values in the set of half-edges $[\ell_n]$ modulated by a sequence of Bernoulli random variables $\beta_j(1), \ldots, \beta_j(d_j)$. The distribution of both these random variables will depend on $(U_1, U_2, (\sigma_i)_{i \geq 1}, (\xi_i)_{i \geq 1})$. The precise laws of these ingredients will be specified as we sequentially apply Procedure 2.1 below.

Before using the above ingredients to construct SWT, let us first describe how they are used to construct the offspring of the branching processes $BP_n$ and $BP$. We define, for $i \geq 1$,

\begin{equation}
B_i^{(n)} = \sum_{j=1}^{n} (d_j - 1) \mathbb{1}_{[d_1 + \cdots + d_{j-1} < \sigma_i \leq d_1 + \cdots + d_j]},
\end{equation}

that is, when $\sigma_i$ chooses one of the half-edges incident to vertex $j$, $B_i^{(n)}$ is the forward degree (i.e., degree minus one) of that vertex $j$. Obviously, the sequence $(B_i^{(n)})_{i \geq 1}$ is i.i.d. with common distribution given by

\begin{equation}
g_k^{(n)} = \mathbb{P}(B_i^{(n)} = k) = \frac{k + 1}{\ell_n} \sum_{j=1}^{n} \mathbb{1}_{[d_j = k + 1]}, \quad k \geq 0.
\end{equation}

Note that $g_k^{(n)} = \mathbb{P}(D_n^{*} = k + 1)$, where $D_n^{*}$ has the same distribution as the size-biased version of $D_n$, the degree of a randomly selected vertex; see (1.7). Assuming Condition 1.1, we have that $g_k^{(n)} \to g_k$, as $n \to \infty$, where

\begin{equation}
g_k = \frac{(k + 1) f_{k+1}}{\sum_{j=1}^{\infty} j f_j}, \quad k \geq 0,
\end{equation}

and where $f_j = F(j) - F(j - 1)$.

We next construct an i.i.d. sequence $(B_i)_{i \geq 1}$ with common distribution (2.5) by using the already constructed $(B_i^{(n)})_{i \geq 1}$ sequence as follows: For each $i \geq 1$, $B_i$ depends only on $B_i^{(n)}$ and is generated via the conditional distribution

\begin{equation}
\mathbb{P}(B_i = k \mid B_i^{(n)} = l) = \frac{p_{kl}}{\sum_{j=0}^{\infty} p_{jl}},
\end{equation}
where

\begin{equation}
    p_{kl}^{(n)} = \begin{cases} 
    \min \{g_k, g_k^{(n)}\}, & \text{for } k = l, \\
    (g_k - \min \{g_k, g_k^{(n)}\})(g_l^{(n)} - \min \{g_l, g_l^{(n)}\}), & \text{for } k \neq l.
\end{cases}
\end{equation}

It is easy to check that \((B_i)_{i \geq 1}\) is an i.i.d. sequence of random variables having probability mass function \((g_k)_{k \geq 0}\) in (2.5). In fact, the joint distribution \((p_{kl}^{(n)})_{k,l \geq 0}\) is the one that maximizes the coupling probability between the two probability mass functions \((g_k)_{k \geq 0}\) and \((g_k^{(n)})_{k \geq 0}\) (alternatively, the coupling that minimizes the total variation distance between the two distributions [40]).

Let us now proceed to the more involved construction of the shortest weight tree SWT using the above probabilistic ingredients. The main ingredient of our construction are the continuous-time processes of “alive” half-edges \((AH(t))_{t \geq 0}\) and “free” half-edges \((F(t))_{t \geq 0}\). We also introduce two new random sequences \((Y(n)_k)_{k \geq 1}\) and \((X(n)_k)_{k \geq 0}\). We will need an additional superscript \(i\) to denote whether \(Y(n)_k\) and/or \(X(n)_k\) belongs to the SWT of \(U_{i,i}=1,2\). The continuous-time processes \((AH(t))_{t \geq 0}\) and \((F(t))_{t \geq 0}\) only change at random times \(T_0 = 0 < T_1 < T_2 < \cdots\) and therefore a full description of the continuous-time evolution can be given by a specification of how the random times above are constructed and how these processes “jump” at each of these times.

At time \(t = T_0\), we start by testing whether any of the \(d_{U_1}\) half-edges incident to \(U_1, i = 1, 2\), are paired to one another. This is performed vertex by vertex, and we start with \(U_1\). Let us define \(HE_j\), for \(j \in [n]\), as the set of \(d_j\) half-edges that belong to vertex \(j\). We define \(Y_0^{(n,1)}\) as the number of half-edges incident to \(U_1\), that is,

\begin{equation}
    Y_0^{(n,1)} = d_{U_1} = |HE_{U_1}|.
\end{equation}

Now put \(AS = HE_{U_1}\), \(TS = (AS)\), where the parentheses \((\cdot)\) indicate that we consider a list instead of a set, and \(N = \ell_n - d_{U_1}\), and apply Procedure 2.1 to remove all half-edges from the total set of \(d_{U_1}\) half-edges that are part of a self-loop. We then define \(RHE_{U_1}\) as the set of unpaired half-edges after the self-loops incident to \(U_1\) are removed and

\begin{equation}
    X_0^{(n,1)} = |RHE_{U_1}|,
\end{equation}

as the number of unpaired half-edges of \(U_1\) after the self-loops have been removed. We attach i.i.d. weights to each of the half-edges in \(RHE_{U_1}\) by taking the first \(X_0^{(n,1)}\) weights from \((\xi_i)_{i \geq 1}\).

We continue with the \(d_{U_2}\) half-edges incident to \(U_2\), and test whether they are paired to one of the \(X_0^{(n,1)}\) remaining half-edges incident to \(U_1\), or any of the \(d_{U_2}\) half-edges incident to \(U_2\). We do this by applying Procedure 2.1 with \(AS = RHE_{U_1} \cup HE_{U_2}\), \(TS = (HE_{U_2})\) and

\begin{equation}
    N = \ell_n - d_{U_1} - d_{U_2} - X_0^{(n,1)},
\end{equation}
which equals the total number of half-edges that are still available to be connected to (noting that the ones that are paired to the half-edges incident to $U_1$ are no longer available). The subtraction of $X^{(n,1)}_0$ in (2.10) originates from the fact that $X^{(n,1)}_0$ free half-edges are claimed to match the reduced forward degree. A self-loop is formed when during this test a half-edge is paired to one of the $d_{U_2}$ sibling half-edges. A so-called collision edge is formed when during this test a half-edge is paired to one of the $X^{(n,1)}_0$ remaining half-edges incident to vertex $U_1$. The weight of this collision edge is the weight of the half-edge incident to $U_1$, which it has already obtained in the previous step. A collision produces a path between vertices $U_1$ and $U_2$, which possibly is the minimal weight path between $U_1$ and $U_2$. We define $\text{RHE}_{U_2}$ as the set of unpaired half-edges incident to vertex $U_2$ after the removal of the self-loops and collision edges. Furthermore, we define

\begin{equation}
Y^{(n,2)}_0 = d_{U_2},
\end{equation}

and

\begin{equation}
X^{(n,2)}_0 = |\text{RHE}_{U_2}|,
\end{equation}

that is, $X^{(n,2)}_0$ denotes the number of unpaired half-edges of vertex $U_2$ after the test for collision edges and self-loops has been performed. We attach i.i.d. weights to each of the half-edges in $\text{RHE}_{U_2}$ by taking the first $X^{(n,2)}_0$ available weights from the i.i.d. sequence $(\xi_i)_{i \geq 1}$ (note that the first $X^{(n,1)}_0$ weights have already been assigned to the half edges in $\text{RHE}_{U_1}$). By construction, the remaining $X^{(n,1)}_0 + X^{(n,2)}_0$ half-edges incident to the vertices $U_1, U_2$ are paired to fresh vertices, that is, vertices distinct from $U_1$ and $U_2$.

For the moment, we collect the possible collision edges at time $T_0$, together with their weights, which is equal to the weight of the half-edge incident to $U_1$ that forms one half of the collision edge, and continue with the description. All half-edges that are not paired to one of the other $d_{U_1} + d_{U_2} - 1$ half-edges incident to either $U_1$ or $U_2$ together form the set $\text{AH}(0)$, the set of active half-edges at time 0, that is,

\begin{equation}
\text{AH}(0) = \text{RHE}_{U_1} \cup \text{RHE}_{U_2}.
\end{equation}

For $y \in \text{AH}(0)$, we define the height $H(y) = 1$ and its index $I(y) = i$, if the half-edge $y$ is connected to $U_i$, $i = 1, 2$, and $R_0(y)$ for the weight with distribution $F_\xi$ that the half edge received earlier. This initiates the construction with

\begin{equation}
\text{SWT}(0) = (y, H(y), I(y), R_0(y))_{y \in \text{AH}(0)},
\end{equation}

and we let $\text{SWT}^{(i)}(0) = (y, H(y), R_0(y))_{y \in \text{AH}(0), I(y) = i}$, be the subset of $\text{SWT}(0)$ that is connected to vertex $U_i$, $i = 1, 2$.

After the above initialization, let us now describe how to construct the process $\text{SWT}(t) = \text{SWT}^{(1)}(t) \cup \text{SWT}^{(2)}(t)$ and $\text{SWT}^{(i)}$ with $i = 1, 2$, for general $t > 0$. Abusing notation, we call $\text{SWT}^{(i)}(t)$ the shortest path tree emanating from vertex
Using the information \((\mathsf{SWT}^{(i)}(s))_{0 \leq s \leq t}\), we can construct the genealogical tree representing how the liquid percolates from the source \(U_i\) but this process contains much more information including edge lengths encountered by the process. As for \(t = 0\), the process \(\mathsf{SWT}(t)\) has a set of “alive” half-edges \(\mathsf{AH}(t)\), which we formally define below. For \(y \in \mathsf{AH}(t)\), we record its index \(I(y) \in \{1, 2\}\) if \(y \in \mathsf{SWT}^{(i)}(s)\) and we let \(H(y)\) denote the graph distance of \(y\) to \(U_I(y)\) [viewing \((\mathsf{SWT}^{(i)}(s))_{s \in [0, t]}\) as a tree]. Further, for \(y \in \mathsf{AH}(t)\), we let \(R_t(y)\) denote the residual lifetime of \(y\) at time \(t\). Then we let

\[
\tag{2.15} \mathsf{SWT}(t) = (y, H(y), I(y), R_t(y))_{y \in \mathsf{AH}(t)},
\]

denote the set of alive half-edges together with their indices, their heights and residual lifetimes. At a later state, we will also define \(\mathsf{BP}_n(t)\) and \(\mathsf{BP}(t)\), the CTBP analogs of \(\mathsf{SWT}(t)\).

We next recursively define the evolution of \((\mathsf{SWT}(t))_{t \geq 0}\). Define \(T_1 = \min_{y \in \mathsf{AH}(0)} R_0(y)\) and denote the half-edge equal to the argument of this minimum by \(y_0^{*}\), hence \(R_0(y_0^*) = \min_{y \in \mathsf{AH}(0)} R_0(y)\). Since the distribution of the weights (lifetimes) admits a density \(f_\xi\), \(y_0^*\) is a.s. unique. Now set

\[
\tag{2.16} \mathsf{AH}(t) = \mathsf{AH}(0), \quad 0 \leq t < T_1,
\]

that is, the active set remains unchanged during the interval \([0, T_1)\). This defines the shortest weight tree in (2.15) for \(0 \leq t < T_1\), where \(I(y)\) and \(H(y)\) are defined above and \(R_t(y) = R_0(y) - t, 0 \leq t < T_1\), denotes the remaining lifetime of half-edge \(y\) at time \(t\).

At time \(t = T_1\), we continue by describing the pairing of the half-edge \(y_0^*\) with \(z_0 = P_{y_0^*}\) and at this place we will introduce the coupling between \(Y_1^{(n)}\) and \(B_1^{(n)}\) [see (2.3)]. For a half-edge \(y\), let \(V_y\) denote the vertex incident to it. By construction, \(z_0 = P_{y_0^*}\) is chosen such that \(V_{z_0}\) is not equal to \(U_i, i = 1, 2\). This is achieved by taking

\[
\tag{2.17} \tau_1 = \min\{m \geq 1 : V_{\sigma_m} \neq U_1, V_{\sigma_m} \neq U_2\},
\]

and we define

\[
\tag{2.18} Y_1^{(n)} = B_1^{(n)} \quad \text{and} \quad z_0 = \sigma_{\tau_1}.
\]

When \(\tau_1 = 1\), we see that \(Y_1^{(n)} = B_1^{(n)}\), while when \(\tau_1 > 1\), the forward degree \(Y_1^{(n)}\) of the chosen vertex \(V_{z_0}\) is not successfully coupled to the random variable \(B_1^{(n)}\).

At time \(t = T_1\), we remove \(y_0^*\) from the set \(\mathsf{AH}(t-.)\). Then, for each of the \(d_{V_{z_0}} - 1\) other half-edges incident to vertex \(V_{z_0}\) we test, using Procedure 2.1, with

\[
\tag{2.19} \mathsf{AS} = \mathsf{AH}(t-) \cup (\mathsf{HE}_{V_{z_0}} \setminus \{z_0\}), \quad \mathsf{TS} = (\mathsf{HE}_{V_{z_0}} \setminus \{z_0\}),
\]

and

\[
\tag{2.20} N = \ell_n - d_{U_1} - d_{U_2} - d_{V_{z_0}} - |\mathsf{AH}(0)|,
\]
which again has the interpretation of the number of available half-edges at the time of finding \( V_{z_0} \), whether it is part of a self-loop or paired to a half-edge from the set \( \text{AH}(t-) \). All half-edges incident to \( V_{z_0} \) that are part of a self-loop or incident to \( \text{AH}(t-) \) are removed from vertex \( V_{z_0} \); we also remove the involved half-edges from the set \( \text{AH}(t-) \). For all the remaining sibling half-edges \( x \) of \( z_0 \), we do the following: \( x \) is added to \( \text{AH}(t-) \), \( I(x) = I(y_0^*) \), \( H(x) = H(y_0^*) + 1 \), while \( R_{T_1}(x) \) is the next available i.i.d. lifetime from the sequence \( (\xi_i)_{i \geq 1} \). This constructs \( \text{AH}(T_1) \).

We now set \( \text{AH}(t) = \text{AH}(T_1), \quad T_1 \leq t < T_2, \)

where \( T_2 = T_1 + \min_{y \in \text{AH}(T_1)} R_{T_1}(y) \), and where the minimizing half-edge is called \( y_1^* \).

We continue using induction, by defining \( \text{AH}(t) \) and \( \text{SWT}(t) \) during the random interval \([T_k, T_{k+1})\) for \( k \geq 1 \), given that the processes are defined on \([0, T_k)\). By construction, we know that \( z_{k-1} = P_{y_{k-1}^*} \) is chosen such that \( V_{z_{k-1}} \) is not equal to \( U_i, i = 1, 2 \) or one of the previously chosen vertices \( V_{z_j}, 1 \leq j \leq k-2 \) (for \( k = 1 \), the latter is an empty condition). Therefore, we take

\[
\tau_k = \min\{m \geq \tau_{k-1} + 1: V_{\sigma_m} \notin \{U_1, U_2, V_{z_0}, \ldots, V_{z_{k-2}}\}\},
\]

and we define

\[
Y_k^{(n)} = B_{\tau_k}^{(n)} \quad \text{and} \quad z_{k-1} = \sigma_{\tau_k}.
\]

When \( \tau_k = \tau_{k-1} + 1 \), we see that \( Y_k^{(n)} = B_{\tau_{k-1}+1}^{(n)} \), while for \( \tau_k > \tau_{k-1} + 1 \), the forward degree \( Y_k^{(n)} \) of the chosen vertex \( V_{z_{k-1}} \) is not coupled to the random variable \( B_{\tau_{k-1}+1}^{(n)} \) and we call the vertex \( V_{z_{k-1}} \) degree-miscoupled. At time \( t = T_k \), we remove \( y_{k-1}^* \) from the set \( \text{AH}(t-) \). Then, for each of the \( d_{V_{z_{k-1}}} - 1 \) other half-edges incident to vertex \( V_{z_{k-1}} \), we use Procedure 2.1, with

\[
\text{AS} = \text{AH}(t-) \cup (\text{HE}_{V_{z_{k-1}}} \setminus \{z_{k-1}\}), \quad \text{TS} = (\text{HE}_{V_{z_{k-1}}} \setminus \{z_{k-1}\}),
\]

and

\[
N = \ell_n - d_{U_1} - d_{U_2} - \sum_{j=0}^{k-1} d_{V_{z_j}} - |\text{AH}(T_{k-1})|,
\]

to test whether it is part of a self-loop or paired to a half-edge from the set \( \text{AH}(t-) \). It is part of Procedure 2.1 that all half-edges incident to \( V_{z_{k-1}} \) that are part of a self-loop or incident to \( \text{AH}(t-) \) are removed from vertex \( V_{z_{k-1}} \); we also remove the involved half-edges from the set \( \text{AH}(t-) \). We will discuss the role of the half-edges incident to \( V_{z_{k-1}} \) that are paired to half-edges in \( \text{AH}(t-) \) in more detail below.

We sequentially order the remaining siblings half-edges of \( z_{k-1} \) in an arbitrary order. In this order, we do the following: Let \( x \) be one such half-edge of \( V_{z_{k-1}} \), then
add $x$ to $\text{AH}(t-)$, and set $I(x) = I(y^*_k)$ and $H(x) = H(y^*_k) + 1$, while $R_{T_k}(x)$ is the next in line of the i.i.d. sequence $(\xi_i)_{i \geq 1}$. This constructs $\text{AH}(T_k)$. We now set

$$AH(t) = AH(T_k), \quad T_k \leq t < T_{k+1},$$

where $T_{k+1} = T_k + \min_{y \in \text{AH}(T_k)} R_{T_k}(y)$, and where the minimizing half-edge is called $y^*_k$.

For $t \in [T_k, T_{k+1})$, we define $\text{SWT}(t)$ by (2.15), where $R_t(y) = R_{T_k}(y) - (t - T_k)$, $T_k \leq t \leq T_{k+1}$. Finally, we denote the number of the $d_{V_{z_{k-1}}} - 1$ other half-edges incident to vertex $V_{z_{k-1}}$ that do not form a self-loop and that are not paired to a half-edge from the set $\text{AH}(t-)$ by $X_{k}^{(n)}$. We say that the vertex $V_{z_{k-1}}$ is successfully degree-coupled to the corresponding individual in a branching process that has offspring $B_{T_{k-1}+1}^{(n)}$ (this will show up again in the next section) when both

$$Y_{k}^{(n)} = B_{T_{k-1}+1}^{(n)} \quad \text{and} \quad X_{k}^{(n)} = Y_{k}^{(n)},$$

and otherwise we call it degree-miscoupled.

We finally denote $S_{k}^{(n)} = |\text{AH}(T_k)|$, so that $S_{0}^{(n)} = X_{0}^{(n,1)} + X_{0}^{(n,2)}$, while $S_{k}^{(n)}$ satisfies the recursion

$$S_{k}^{(n)} = S_{k-1}^{(n)} + Y_{k}^{(n)} - 2(Y_{k}^{(n)} - X_{k}^{(n)}) - 1, \quad k \geq 1.$$ 

This describes the evolution of $(\text{SWT}(t))_{t \geq 0}$.

**Cycle edges and collision edges.** At time $T_k, k \geq 1$, we find the half-edge $y^*_k$ that is paired to $z_{k-1} = P_{y^*_k}$, and for each of the other half-edges $x$ incident to $V_{z_{k-1}}$, we check, using Procedure 2.1, whether or not a self-loop has been formed or whether or not $P_x \in \text{AH}(T_k-).$ The newly found half-edges that are paired to already alive half-edges in $\text{AH}(T_k-)$ are special. Indeed, the edge $(x, P_x)$ creates a cycle when $I(x) = I(P_x)$, while $(x, P_x)$ completes a path between $U_1$ and $U_2$ when $I(x) = 3 - I(P_x)$. Precisely the latter edges can create the shortest-weight path between $U_1, U_2$. Let us describe these collision edges in more detail.

At time $T_k$ and when we create a collision edge consisting of $x$ and $P_x$, then we record

$$(2.28) \quad (T_k, I(x), H(x), H(P_x), R_{T_k}(P_x)),$$

where

$$(2.29) \quad I(x) = I(y^*_k), \quad H(x) = H(y^*_k) + 1.$$ 

Order the times at which collision edges occur as $(T_{j}^{(\text{col})})_{j \geq 1}$, and let $(x_j, P_{x_j})$ be the corresponding collision edge (so that $P_{x_j}$ is in the other $\text{SWT}$ as $x_j$). If multiple collision edges are created at the same time, then order them arbitrarily. We will
see that the probability of such events in the time scale of interest converges to zero as $n \to \infty$. Write

\begin{equation}
C := \left( T_j^{(\text{col})}, I(x_j), H(x_j), H(P_{x_j}), R_{T_j^{(\text{col})}}(P_{x_j}) \right)_{j \geq 1},
\end{equation}

for the collection of all collision edges collected by the process.

It is possible (albeit unlikely) that multiple half-edges incident to $V_{zk-1}$ create collision edges, and if so, we collect all of them in the list in (2.30). With some abuse of notation we denote the $j$th collision edge by $(x_j, P_{x_j})$. For the collection of all collision edges collected by the process.

Let $I^\star$ denote the minimizer of $j \mapsto 2T_j^{(\text{col})} + R_{T_j^{(\text{col})}}(P_{x_j})$, then

\begin{equation}
H_n = H(x_{I^\star}) + H(P_{x_{I^\star}}) + 1.
\end{equation}

Of course, (2.31) and (2.32) need a proof, which we give now.

\textit{Proof that $L_n$ in (2.31), and $H_n$ in (2.32) yield the minimal weight and hopcount, respectively.} Observe that the weight of each path between $U_1$ and $U_2$ can be written in the form $L = 2T_i^{(\text{col})} + R_{T_i^{(\text{col})}}(P_{x_i})$, for some $i \geq 0$. Indeed, let $(i_0 = U_1, i_1, i_2, \ldots, i_k = U_2)$ form a path with weight $L$, and denote the weight on $(i_{j-1}, i_j)$ by $\xi_{e_j}$, for $1 \leq j \leq k$. For $k = 1$, we obviously find $\xi_{e_1} = 2T_0 + \xi_{e_1}$. For general $k \geq 1$, take the maximal $j \geq 0$ such that $\xi_{e_1} + \cdots + \xi_{e_j} \leq L/2$. Then, since $L = \sum_{s=1}^k \xi_{e_s}$, we have that $\sum_{s=1}^j \xi_{e_s} \leq \sum_{s=j+1}^k \xi_{e_s}$, so that

\[ L = 2 \sum_{s=1}^j \xi_{e_s} + \left[ \sum_{s=j+1}^k \xi_{e_s} - \sum_{s=1}^j \xi_{e_s} \right], \]

which is of the form $L = 2T_j^{(\text{col})} + R_{T_j^{(\text{col})}}(y)$, for some $j \geq 0$ and some half-edge $y$. Note that in the construction of the SWTs, instead of putting weight on the edges, we have given weights to half-edges instead. In the representation (2.31), full edge weight is given to the active half-edges and weight 0 to the ones to which they are paired. At time $T_j^{(\text{col})}$ when a collision edge has been found, the path-weight of the edges belonging to the same vertex $U_i$ as half-edge $y^\star$ add up to $T_j^{(\text{col})}$, the path-weight of all completed edges connected to $3 - U_i$ together with the residual
lifetime $R_j^{(\text{col})}(P_x)$ of the half-edge $P_x$ has to be added to $T_j^{(\text{col})}$ in order to yield the total weight of the path between $U_1$ and $U_2$.

The proof of (2.32) follows because the number of edges of the path between $U_1$ and $U_2$ that passes through the collision edge $(x_j, P_{x_j})$ is equal to the sum of the heights of the vertices incident to $x_j$, $P_{x_j}$, respectively, and we add 1 for the edge $(x_j, P_{x_j})$ itself. This completes the proof of the claim.

2.3. Coupling: Process level. In the above, we have described the coupling between reduced forward degrees $(X^{(n)}_i)_{i \geq 1}$ in SWT and i.i.d. random variables $((B^{(n)}_i, B_i))_{i \geq 1}$, where $(B^{(n)}_i)_{i \geq 1}$ has marginal distribution (2.4) and $(B_i)_{i \geq 1}$ has marginal distribution (2.5), and they are coupled as in (2.6)–(2.7). We have used this coupling to describe the evolution of $(\text{SWT}(t))_{t \geq 0}$, and at the end of this process, we know of each vertex that is found by the liquid, whether it is successfully degree-coupled or not. As long as no degree-miscouplings occur, this can be thought of as a coupling between SWT and two CTBPs with lifetimes having distribution $F_\xi$ and offsprings $(B^{(n)}_i)_{i \geq 1}$ and $(B_i)_{i \geq 1}$, respectively, but the evolutions will start to diverge as soon as degree-miscouplings start to appear. We now extend this coupling.

Recall that the SWT$(i)$ consists of half-edges and their attributes, connected to $U_i$, for $i = 1, 2$. We aim to couple each SWT$(i)$, $i = 1, 2$, to an independent CTBP BP$(i)$, so that SWT is coupled to BP, which consists of two independent CTBPs, that is, BP = (BP$^{(1)}$, BP$^{(2)}$), as well as to an $n$-independent limiting CTBP BP that also consists of two independent CTBPs, that is, BP = (BP$^{(1)}$, BP$^{(2)}$). If $Y^{(n)}_k$ is the forward degree of vertex $V_{z_k-1}$, then $I(y^{*}_{k-1})$ indicates to which SWT the vertex belongs. We recall that we say that the vertex $V_{z_k-1}$ in the SWT is degree-miscoupled to the corresponding individual (which we also refer to as vertex) if

$$Y^{(n)}_k \neq B^{(n)}_{r_{k-1}+1} \text{ or if } X^{(n)}_k \neq Y^{(n)}_k.$$  

Vertices that are degree-miscoupled will appear both in the SWT as well as in the CTBP BP$^n$. However, after being degree-miscoupled, the evolution of vertices in the CTBP and SWT diverge, as we explain now. For the SWT, we say that an alive half-edge is miscoupled if the shortest-weight path to the vertex incident to that half-edge uses at least one degree-miscoupled vertex. In particular, the evolution of the SWT is such that half-edges of degree-miscoupled vertices are by definition attached to miscoupled half-edges. The same is true for the CTBP BP$^n$, that is, offspring of degree-miscoupled individuals are by definition miscoupled.

The weights assigned to half-edges incident to miscoupled vertices in the SWT and individuals in the CTBP are independent. For this, we have introduced a second sequence of i.i.d. weights $(\xi_i)_{i \geq 1}$ that is independent of the edge-weights of correctly coupled half-edges $(\xi_i)_{i \geq 1}$. Each time that a half-edge is found by the SWT, we perform Procedure 2.1 and the coupling to the CTBPs BP$^n$ and BP. When, instead, a half-edge is found by (one of the) CTBPs only, we pair it to a
uniformly chosen half-edge chosen from \([\ell_n]\) without replacement. These choices are determined by the i.i.d. sequence \((\tilde{\sigma}_i)_{i \geq 1}\), and, from these, the random variables \(\tilde{B}^{(n)}_i\) and \(\tilde{B}_i\) are obtained as explained in (2.3) and (2.6)–(2.7). We use the variables \(\tilde{B}^{(n)}_i\) for the evolution of \(\text{BP}_n\), and \(\tilde{B}_i\) for the evolution of \(\text{BP}\). Thus, the evolution of miscoupled individuals in the CTBPs \(\text{BP}_n\) and \(\text{BP}\) is completely independent of the evolution of \(\text{SWT}\). When differences arise in \(\text{BP}_n\) and \(\text{BP}\), also their evolutions are mutually independent.

We close this section by defining the sets of alive individuals in the coupling of the random variables \((\text{SWT}(t), \text{BP}_n(t), \text{BP}(t))_{t \geq 0}\). Both \(\text{BP}_n(t)\) as well as \(\text{BP}(t)\) each have their sets of alive individuals that we denote by \(\mathcal{A}_n(t)\) and \(\mathcal{A}(t)\), respectively. For \(\text{BP}_n(t)\), we can think of these alive individuals as corresponding to repeated draws of half-edges. By our coupling, these sets of alive individuals in \(\text{BP}_n(t)\) and \(\text{BP}(t)\) are effectively coupled to the alive half-edges in \(\mathcal{A}(t)\). The successfully coupled half-edges in \(\text{SWT}(t)\) and \(\text{BP}_n(t)\) at time \(t\) form \(\mathcal{A}(t) \cap \mathcal{A}_n(t)\), the successfully coupled individuals in \(\text{BP}_n(t)\) and \(\text{BP}(t)\) form \(\mathcal{A}_n(t) \cap \mathcal{A}(t)\). We note that each alive half-edge, individual \(y\) in \(\mathcal{A}(t)\) and \(\mathcal{A}_n(t)\) has a residual lifetime \(R_{y}(t)\), as well as an index \(I(y)\) indicating which subtree \(y\) is an element of and a height \(H_{y}(t)\) denoting the generation of \(y\). Similarly to (2.15), we then define

\[
\text{BP}_n(t) = (y, I(y), H(y), R_{y}(t))_{y \in \mathcal{A}_n(t)},
\]

\[
\text{BP}(t) = (y, I(y), H(y), R_{y}(t))_{y \in \mathcal{A}(t)}.
\]

This completes the coupling of the FPP processes \((\text{SWT}(t), \text{BP}_n(t), \text{BP}(t))_{t \geq 0}\) and defines the probability space \((\Omega, \mathcal{F}, \mathbb{P})\) on which this coupling of \((\text{SWT}(t), \text{BP}_n(t), \text{BP}(t))\) lives. We let \((\mathcal{F}_t)_{t \geq 0}\) be the filtration generated by all the randomness used in the construction up to time \(t\), that is, \(\mathcal{F}_t\) contains all the information needed to construct \((\text{SWT}(s), \text{BP}_n(s), \text{BP}(s))_{s \in [0,t]}\). Under this coupling law, we can speak of convergence in probability, and we shall frequently do this in the sequel.

**Summary of the coupling.** For completeness and future references, we resume how differences arise in the coupling. Degree-miscouplings arise due to three effects:

1. **MISC-miscouplings** occur between the forward degree \(Y_k^{(n)}\) (which are not i.i.d. due to draws being without replacement) and the i.i.d. draws \((B_i^{(n)})_{i \geq 1}\), because \(\tau_k > \tau_{k-1} + 1\);

2. **cycle-events** occur referring to self-loops or cycles that makes \(X_k^{(n)} < Y_k^{(n)}\). In these cases, we remove the self-loop or the edge that gives rise to a cycle from the set of alive half-edges. This amounts to removing up to at most \(Y_k^{(n)} - X_k^{(n)}\) half-edges incident to vertex \(V_{z_k-1}\), as well as up to at most \(Y_k^{(n)} - X_k^{(n)}\) half-edges to which they are paired from \(\text{SWT}\).
(3) collision edges are found. In this case, precisely one of the vertices to which the collision edge is incident is degree-miscoupled. We want to emphasize here that the degree-miscoupling caused by finding the collision edge at time $T^{(\text{col})}_j$ does not effect the coupling of the shortest-weight paths. When the collision edge is removed, we are left with two paths connecting $U_i$ to one of the vertices incident to the two half-edges of which the collision edge consists. It should be checked that at any time prior to $T^{(\text{col})}_j$ each of these paths is not miscoupled, that is, does not contain any earlier degree-miscoupled vertices, since otherwise the path is not necessarily present in the random graph.

In all these three cases, the vertices involved are called degree-miscoupled, and any further offspring of degree-miscoupled vertices (in the SWT or in the CTBP) are called miscoupled. Thus, any miscoupling gives rise to a tress of miscoupled children half-edges in the SWT, respectively, offspring in the CTBP.

2.4. Main coupling results. We consider the process coupling $(\text{SWT}(t), \text{BP}_n(t), \text{BP}(t))_{t \geq 0}$ defined in the previous section, as well as the associated filtration $(\mathcal{F}_t)_{t \geq 0}$. We recall that $\text{BP}_n(t) = (\text{BP}^{(1)}_n(t), \text{BP}^{(2)}_n(t))$ are two independent CTBPs starting with offspring distribution $D_n$ in the first generation and offspring law $B^{(n)}_n = D^* - 1$ in the second and further generations, and $\text{BP}(t) = (\text{BP}^{(1)}(t), \text{BP}^{(2)}(t))$ which are two independent CTBPs starting with offspring distribution $D$ in the first generation and offspring distribution $B = D^* - 1$ in the second and further generations. For this coupling $(\text{SWT}(t), \text{BP}_n(t), \text{BP}(t))_{t \geq 0}$, we let $\text{AH}(t) \triangle \text{AI}_n(t)$ denote the set of miscoupled half-edges at time $t$. With a slight abuse of notation, we write $|\text{SWT}(t)| = |\text{AH}(t)|$ and $|\text{SWT}(t) \triangle \text{BP}_n(t)| = |\text{AH}(t) \triangle \text{AI}_n(t)|$. Finally, we denote the set of all miscoupled half-edges and individuals up to time $t$ by

$$\bigcup_{s \in [0, t]} \text{SWT}(s) \triangle \text{BP}_n(s).$$

In this section, we state two key propositions concerning the coupling. Proposition 2.2(a) shows that there exists some $s_n \to \infty$ such that, w.h.p., there are no miscouplings up to time $s_n$. In Proposition 2.2(b) and Proposition 2.3, we investigate the size of $\text{SWT}(t)$ for $t$ close to $t_n = \log n / (2\alpha_n)$.

**Proposition 2.2 (Coupling the SWT to a BP).**

(a) There exists $s_n \to \infty$ such that, for the coupling defined in Sections 2.2–2.3,

$$\mathbb{P}\left((\text{SWT}(s))_{s \in [0, s_n]} = (\text{BP}_n(s))_{s \in [0, s_n]} = (\text{BP}(s))_{s \in [0, s_n]}\right) = 1 - o(1).$$

Consequently, with $W^{(i)}_{s_n} = e^{-\alpha_n s_n} |\text{SWT}^{(i)}(s_n)|$, $i = 1, 2$,

$$\lim_{\varepsilon \downarrow 0} \lim_{n \to \infty} \mathbb{P}(W^{(1)}_{s_n} \in [\varepsilon, 1/\varepsilon], W^{(2)}_{s_n} \in [\varepsilon, 1/\varepsilon] \mid W^{(1)}_{s_n} W^{(2)}_{s_n} > 0) = 1.$$
(b) Let $t_n = \log n / (2\alpha_n)$. For the coupling of $(\text{SWT}(s))_{s \geq 0}$ and $(\text{BP}_n(s))_{s \geq 0}$ defined in Sections 2.2–2.3, there exist sequences $\varepsilon_n \to 0$ and $B_n \to \infty$ such that, conditionally on $\mathcal{F}_{s_n}$, and for every $t \leq t_n + B_n$,

$$P\left(\bigcup_{s \in [0,t]} \text{SWT}(s) \triangle \text{BP}_n(s) \geq \varepsilon_n \sqrt{n} \mid \mathcal{F}_{s_n}\right) \to 0. \quad (2.38)$$

The proof of Proposition 2.2 is deferred to Section 5. We warn the reader to beware for confusion between the (large) constant $B_n$ and the i.i.d. random variables $(B_i)_{i \geq 1}$. Fix the deterministic sequence $s_n \to \infty$ from Proposition 2.2. Now let

$$t_n = \frac{1}{2\alpha_n} \log n, \quad \tilde{t}_n = \frac{1}{2\alpha_n} \log n - \frac{1}{2\alpha_n} \log (\mathcal{W}_{s_n}^{(1)} / \mathcal{W}_{s_n}^{(2)}). \quad (2.39)$$

Note that $e^{\alpha_n t_n} = \sqrt{n}$; it will turn out that both $|\text{SWT}^{(i)}(t_n)|$, for $i = 1, 2$, are of order $\sqrt{n}$. Further, it will turn out that collision edges start to appear when these clusters grow to be of this size. Consequently, the variable $t_n$ denotes the typical time at which collision edges start appearing, and the time $\tilde{t}_n$ incorporates for stochastic fluctuations in the size of the SWTs.

For $i \in \{1, 2\}$, $k \geq 0$, and $s, t \geq 0$, we define

$$|\text{SWT}^{(i)}_k[t, t+s]| = |\{y \in \text{AH}(t) : I(y) = i, H(y) = k, R_t(y) \in [0, s]\}|, \quad (2.40)$$

as the number of alive half-edges at time $t$ that (i) are in the SWT of vertex $U_i$, (ii) have height $k$, and (iii) have remaining (or residual) lifetime at most $s$. We further write

$$|\text{SWT}^{(i)}_{\leq k}[t, t+s]| = |\{y \in \text{AH}(t) : I(y) = i, H(y) \leq k, R_t(y) \in [0, s]\}|, \quad (2.41)$$

for the number of alive half-edges that have height at most $k$. To formulate the CLT for the height of vertices, we will choose

$$k_n(t, x) = \frac{t}{\tilde{v}_n} + x \sqrt{\frac{t}{\tilde{v}_n^3}}, \quad (2.42)$$

where $\tilde{v}_n$, $\tilde{v}$ and $\tilde{\sigma}^2$ are defined in (1.19).

Define the residual life-time distribution $F_r$ to have density $f_r$ given by

$$f_r(x) = \frac{\int_0^\infty e^{-\alpha y} f_\xi(x + y) dy}{\int_0^\infty e^{-\alpha y} [1 - F_\xi(y)] dy} = \frac{\alpha \nu}{v - 1} \int_0^\infty e^{-\alpha y} f_\xi(x + y) dy. \quad (2.43)$$

Below, we let $\Phi$ denote the standard normal distribution function. Finally, for a half-edge $y \in \text{AH}(t)$, we let $X^*_y = d_{V_y} - 1$. 
Proposition 2.3 (Ages and heights in SWT). Fix $j \in \{1, 2\}$, and numbers $x, y, t \in \mathbb{R}$, $s_1, s_2 > 0$, all independent of $n$. Then, conditionally on $\mathcal{F}_{n}$ and $\mathcal{W}_{n}^{(1)} \mathcal{W}_{n}^{(2)} > 0$:

(a) \[
e^{-2\alpha t n} |\text{SWT}_{\leq \delta_n(n, x)}^{(j)}(\tilde{t}_n + t, \tilde{t}_n + t + s_1)| |\text{SWT}_{\leq \delta_n(n, y)}^{(3-j)}(\tilde{t}_n + t, \tilde{t}_n + t + s_2)| \xrightarrow{p} e^{2\alpha t} \Phi(x) \Phi(y) F_R(s_1) F_R(s_2),
\]

(b) \[
e^{-2\alpha t n} |\text{SWT}_{\leq \delta_n(n, x)}^{(j)}(\tilde{t}_n + t, \tilde{t}_n + t + s_1)| \sum_v \chi_v^* \mathbb{1}_{\{v \in \text{SWT}_{\leq \delta_n(n, y)}^{(3-j)}(\tilde{t}_n + t, \tilde{t}_n + t + s_2)\}} \xrightarrow{p} \nu e^{2\alpha t} \Phi(x) \Phi(y) F_R(s_1) F_R(s_2).
\]

The proof of Proposition 2.3 is deferred to Section 6.

3. Main ingredient: Poisson point process limit. In this section, we state our main result that implies Theorems 1.2–1.3. To state this result, we need some additional definitions.

Recall the collection of collision edges $\mathcal{C}$ from (2.30). Here, the $j$th collision edge is given by $(x_j, P_{x_j})$, where $P_{x_j}$ is an alive half-edge and $x_j$ the half-edge which pairs to $P_{x_j}$. Rescaling time by $\tilde{t}_n$ [see (2.39)], we define

\[
\bar{T}_j^{(\text{col})} = T_j^{(\text{col})} - \tilde{t}_n, \quad \bar{H}_j^{(\text{or})} = \frac{H(x_j) - t_n/\bar{\nu}_n}{\sqrt{\bar{\sigma}^2 t_n/\bar{\nu}^3}},
\]

\[
\bar{H}_j^{(\text{de})} = \frac{H(P_{x_j}) - t_n/\bar{\nu}_n}{\sqrt{\bar{\sigma}^2 t_n/\bar{\nu}^3}},
\]

and write the random elements $(\Xi_j)_{j \geq 1}$ with $\Xi_j \in \mathcal{S} := \mathbb{R} \times \{1, 2\} \times \mathbb{R} \times \mathbb{R} \times [0, \infty)$, by

\[
\Xi_j = (\bar{T}_j^{(\text{col})}, I(x_j), \bar{H}_j^{(\text{or})}, \bar{H}_j^{(\text{de})}, R_{T_j^{(\text{col})}}(P_{x_j})).
\]

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

\[
\Pi_n(A) = \sum_{j \geq 1} \delta_{\Xi_j}(A),
\]

where $\delta_x$ gives measure 1 to the point $x$. Let $\mathcal{M}(\mathcal{S})$ denote the space of all simple locally-finite point processes on $\mathcal{S}$ equipped with the vague topology (see, e.g., [29]). On this space, one can naturally define the notion of weak convergence of a sequence of random point processes $\Pi_n \in \mathcal{M}(\mathcal{S})$. This is the notion of convergence referred to in the following theorem.
THEOREM 3.1 (PPP limit of collision edges). Consider the distribution of the point process $\Pi_n \in M(S)$ defined in (3.3) conditioned on $F_{s_n}$ and $W^{(1)}_{s_n}W^{(2)}_{s_n} > 0$, for $s_n$ as in Proposition 2.2. Then, as $n \to \infty$, $\Pi_n$ converges in distribution to a Poisson Point Process (PPP) $\Pi$ with intensity measure

$$\lambda(dt \times i \times dx \times dy \times dr)$$

(3.4)

$$= \frac{2vf_r(0)}{E[D]} e^{2at} dt \otimes \{1/2, 1/2\} \otimes \Phi(dx) \otimes \Phi(dy) \otimes F_r(dr).$$

Theorem 3.1 will be proved in Section 7. In order to reduce our main theorems to PPP convergence in Theorem 3.1, we rely on the fact that $\bar{T}_{\text{col}}^{(1)}$ is tight. This tightness is the content of the next proposition.

PROPOSITION 3.2 (Tightness of appearance of first collision edge). The sequence of random variables $(\bar{T}_{\text{col}}^{(1)})_{n \geq 1}$ is tight.

We will prove Proposition 3.2 at the end of Section 5.

Completion of the proof of Theorems 1.2 and 1.3. Let us now prove Theorem 1.2 subject to Theorem 3.1 and Proposition 3.2. Recall (2.31)–(2.32). First of all, by (3.1), (3.2), (2.31) and (2.32),

$$H_n - \frac{1}{\alpha_n} \log n \sqrt{\frac{\sigma^2}{\nu^3}} \log n$$

(3.5)

is a continuous function of the point process $\Pi_n$, and, therefore, by the continuous mapping theorem, the above random vector converges in distribution to some random limit $(Z, Q)$.

Recall that $I^*$ denotes the minimizer of $i \mapsto 2T^{(\text{col})}_i + R^{(\text{col})}_i(P_{X_i})$. By (2.31), the weight $L_n$ as well as the value of $I^*$, are functions of the first and the last coordinates of $\Pi_n$. The hopcount $H_n$ is a function of the third and the fourth, instead. By the product form of the intensity in (3.4), we obtain that the limits $(Z, Q)$ are independent. Therefore, it suffices to study their marginals.

We start with the limit distribution of the hopcount. By (3.1), (2.32) and (2.39),

$$H_n - \frac{1}{\alpha_n} \log n \sqrt{\frac{\sigma^2}{\nu^3}} \log n$$

(3.6)

$$= \frac{1}{2} \sqrt{2\tilde{H}_{I^*}^{(o)}} + \frac{1}{2} \sqrt{2\tilde{H}_{I^*}^{(d)}} + o_p(1).$$

By Theorem 3.1, the random variables $(\tilde{H}_{I^*}^{(o)}, \tilde{H}_{I^*}^{(d)})$, converge to two independent standard normals, so that also the left-hand side of (3.6) converges to a standard normal.
Theorem 3.1 implies that, conditionally on Proposition 3.2 proves that (\(\pi_i\))

\[
L_n - \frac{1}{\alpha_n} \log n = L_n - 2t_n = L_n - 2\tilde{t}_n - \frac{1}{\alpha_n} \log(\mathcal{W}_{s_n}^{(1)} \mathcal{W}_{s_n}^{(2)})
\]

(3.7)

\[= - \frac{1}{\alpha_n} \log(\mathcal{W}_{s_n}^{(1)} \mathcal{W}_{s_n}^{(2)}) + \min_{i \geq 1} \left[ 2\tilde{T}_i^{(\text{col})} + R_{T_i^{(\text{col})}}(P_{x_i}) \right].\]

By Proposition 2.2, \((\mathcal{W}_{s_n}^{(1)}, \mathcal{W}_{s_n}^{(2)}) \xrightarrow{d} (\mathcal{W}^{(1)}, \mathcal{W}^{(2)})\), which are two independent copies of the random variable in (1.21). We will prove that, conditionally on \(\mathcal{F}_{s_n}\) and \(\mathcal{W}_{s_n}^{(1)} \mathcal{W}_{s_n}^{(2)} > 0\), also \(\min_{i \geq 1} \left[ 2\tilde{T}_i^{(\text{col})} + R_{T_i^{(\text{col})}}(P_{x_i}) \right] \xrightarrow{d} \min_{i \geq 1} \left[ 2\pi_i + R_i \right]\).

By Proposition 2.2, \((\mathcal{W}_{s_n}^{(1)}, \mathcal{W}_{s_n}^{(2)}) \xrightarrow{d} (\mathcal{W}^{(1)}, \mathcal{W}^{(2)})\), which are two independent copies of the random variable in (1.21). We will prove that, conditionally on \(\mathcal{F}_{s_n}\) and \(\mathcal{W}_{s_n}^{(1)} \mathcal{W}_{s_n}^{(2)} > 0\), also \(\min_{i \geq 1} \left[ 2\tilde{T}_i^{(\text{col})} + R_{T_i^{(\text{col})}}(P_{x_i}) \right] \xrightarrow{d} \min_{i \geq 1} \left[ 2\pi_i + R_i \right]\).

The limit distribution of the weight \(L_n\) is slightly more involved. By (2.31) and (2.39),

\[
L_n - \frac{1}{\alpha_n} \log n = L_n - 2t_n = L_n - 2\tilde{t}_n - \frac{1}{\alpha_n} \log(\mathcal{W}_{s_n}^{(1)} \mathcal{W}_{s_n}^{(2)})
\]

(3.7)

\[= - \frac{1}{\alpha_n} \log(\mathcal{W}_{s_n}^{(1)} \mathcal{W}_{s_n}^{(2)}) + \min_{i \geq 1} \left[ 2\tilde{T}_i^{(\text{col})} + R_{T_i^{(\text{col})}}(P_{x_i}) \right].\]

In order to prove that, conditionally on \(\mathcal{F}_{s_n}\) and \(\mathcal{W}_{s_n}^{(1)} \mathcal{W}_{s_n}^{(2)} > 0\), \(\min_{i \geq 1} \left[ 2\tilde{T}_i^{(\text{col})} + R_{T_i^{(\text{col})}}(P_{x_i}) \right] \xrightarrow{d} \min_{i \geq 1} \left[ 2\pi_i + R_i \right]\), let \(\mathcal{L}(K)\) denote the points in the PPP for which the first coordinate is in \([-K, K]\) and the fifth coordinate is in \([0, K]\). Then Theorem 3.1 implies that, conditionally on \(\mathcal{F}_{s_n}\) and \(\mathcal{W}_{s_n}^{(1)} \mathcal{W}_{s_n}^{(2)} > 0\),

\[
\min_{i \in \mathcal{L}(K)} \left[ 2\tilde{T}_i^{(\text{col})} + R_{T_i^{(\text{col})}}(P_{x_i}) \right] \xrightarrow{d} \min_{i \in \mathcal{L}(K)} \left[ 2\pi_i + R_i \right].
\]

Below, we show by direct calculation that \(M = \min_{i \geq 1} \left[ 2\pi_i + R_i \right] \leq K\) with high probability, so that also

\[
\lim_{K \to \infty} \limsup_{n \to \infty} \mathbb{P}(\exists i \text{ such that } \tilde{T}_i^{(\text{col})} \leq K, R_{T_i^{(\text{col})}}(P_{x_i}) \leq K) = 1.
\]

Proposition 3.2 proves that

\[
\lim_{K \to \infty} \liminf_{n \to \infty} \mathbb{P}(\tilde{T}_1^{(\text{col})} \geq -K) = 1.
\]

We conclude that

\[
\lim_{K \to \infty} \liminf_{n \to \infty} \mathbb{P}\left( \min_{i \in \mathcal{L}(K)} \left[ 2\tilde{T}_i^{(\text{col})} + R_{T_i^{(\text{col})}}(P_{x_i}) \right] \right) = \min_{i \geq 1} \left[ 2\tilde{T}_i^{(\text{col})} + R_{T_i^{(\text{col})}}(P_{x_i}) \right] = 1.
\]

The conditional convergence of \(\min_{i \geq 1} \left[ 2\tilde{T}_i^{(\text{col})} + R_{T_i^{(\text{col})}}(P_{x_i}) \right] \xrightarrow{d} \min_{i \geq 1} \left[ 2\pi_i + R_i \right]\) now follows from (3.12) and (3.9).
We next identify the distribution of \( M = \min_{i \geq 1} [2\pi i + R_i] \). First, \((2\pi i)_i \geq 1\) forms a Poisson process with intensity \( \nu f_R(0) e^{\alpha t} dt \). According to [37], Example 3.3 on page 137, the point process \((2\pi i + R_i)_i \geq 1\) is a nonhomogeneous Poisson process with mean-measure the convolution of \( \mu(-\infty, x] = \int_{-\infty}^x \nu f_R(0) e^{\alpha t} dt \) and \( F_R \). Hence, \( P(M \geq x) \) equals the Poisson probability of 0, where the parameter of the Poisson distribution is \( (\mu * F_R)(x) \), so that

\[
P(M \geq x) = \exp\left\{ -\frac{\nu f_R(0)}{\mathbb{E}[D]} e^{\alpha x} \int_0^\infty F_R(z)e^{-\alpha z} d\,dz \right\}.
\]

Let \( \Lambda \) have a Gumbel distribution, that is, \( P(\Lambda \leq x) = e^{-e^{-x}}, x \in \mathbb{R} \), then, from (3.13),

\[
M = \min_{i \geq 1} (2\pi i + R_i) \overset{d}{=} -\alpha^{-1} \Lambda - \alpha^{-1} \log(\nu f_R(0) B / \mathbb{E}[D]),
\]

with \( B = \int_0^\infty F_R(z)e^{-\alpha z} d\,dz \). In the following lemma, we simplify these constants:

**Lemma 3.3 (The constant).** The constants \( B = \int_0^\infty F_R(z)e^{-\alpha z} d\,dz \) and \( f_R(0) \) are given by

\[
B = \bar{v} / (\nu - 1), \quad f_R(0) = \alpha / (\nu - 1).
\]

Consequently, the constant \( c \) in the limit variable (1.23) equals

\[
c = -\log(\nu f_R(0) B / \mathbb{E}[D]) = \log(\mathbb{E}[D](\nu - 1)^2 / (\alpha \nu \bar{v})).
\]

**Proof.** According to (2.43) and (1.15),

\[
f_R(0) = \frac{\alpha \nu}{\nu - 1} \int_0^\infty e^{-\alpha y} f_\xi(y) dy = \alpha / (\nu - 1).
\]

For \( B \), we use partial integration and substitution of (2.43). This yields

\[
B = \int_0^\infty F_R(z)e^{-\alpha z} d\,dz = \frac{1}{\alpha} \int_0^\infty f_R(z)e^{-\alpha z} d\,dz
\]

\[
= \frac{\nu}{\nu - 1} \int_0^\infty e^{-\alpha z} \int_0^\infty e^{-\alpha y} f_\xi(y + z) dy d\,dz
\]

\[
= \frac{\nu}{\nu - 1} \int_0^\infty s f_\xi(s)e^{-\alpha s} ds = \frac{1}{(\nu - 1)} \int_0^\infty s \tilde{F}_\xi(ds) = \tilde{v} / (\nu - 1).
\]

This completes the proof of Theorems 1.2 and 1.3 subject to Theorem 3.1 and Proposition 3.2. 
\( \square \)
4. Height CLT and residual lifetime for CTBP. In this section, we set the stage for the proof of Proposition 2.3 for CTBPs. We make use of second moment methods similar to the ones in [21, 24, 25, 38], but with a suitable truncation argument to circumvent the problem of infinite-variance offspring distributions.

As in the first part of Section 1.3, we consider a (standard) CTBP process [21], Chapter 6, with lifetime distribution \( F_\xi \) admitting a density \( f_\xi \), and random offspring \( X = X_v \), satisfying (1.13) and the \( X \log X \) condition in (1.14). We define

\[
\eta = \nu \int_0^\infty e^{-2\alpha s} dF_\xi(s) \quad \text{and} \quad m_j = K \eta^{-j}, \quad j \geq 1,
\]

for some \( K > 1 \). Note that \( \eta \in (0, 1) \), since \( \alpha \) is such that \( \nu \int_0^\infty e^{-\alpha s} dF_\xi(s) = 1 \). The truncated CTBP \( \text{BP}^{(\vec{m})} \) has for each individual in generation \( j \) offspring \((X \wedge m_j)\) instead of \( X \).

We denote the number of alive individuals in the CTBP at time \( t \) by \(|\text{BP}(t)|\). By \( |\text{BP}_k(t)|, |\text{BP}_k[t, t+s]| \), we denote the number of alive individuals in generation \( k \) at time \( t \), number of alive individuals in generation \( k \) at time \( t \) with residual lifetime at most \( s \), respectively. We warn the reader that \(|\text{BP}_k(t)|\) refers to the number of individuals in generation \( k \), and not to the \( n \)-dependence. When dealing with \( n \)-dependent CTBPs, we will use the notation \(|\text{BP}_{n,k}(t)|\) instead.

Here, the generation of the first individual equals 0, and the generation of its offspring is equal to 1, etc. For the truncated process \( \text{BP}^{(\vec{m})}(t) \), we define, analogously to the definitions above, \(|\text{BP}^{(\vec{m})}(t)|, |\text{BP}^{(\vec{m})}_k(t)|, \) and \(|\text{BP}^{(\vec{m})}_k[t, t+s]| \). Furthermore,

\[
|\text{BP}^{(\vec{m})}_{\leq k}[t, t+s]| = \sum_{j=0}^k |\text{BP}^{(\vec{m})}_j[t, t+s]|,
\]

\[
|\text{BP}^{(\vec{m})}[t, t+s]| = \sum_{j=0}^\infty |\text{BP}^{(\vec{m})}_j[t, t+s]|.
\]

A key ingredient to the proof of Proposition 2.3 is Proposition 4.1 below.

**Proposition 4.1 (First and second moment CLT).** Consider the branching process \( \text{BP}(t) \) introduced above, with i.i.d. lifetimes \( F_\xi \) admitting a density and random offspring \( X \) satisfying \( \nu = \mathbb{E}[X] > 1 \), and \( \mathbb{E}[X \log(X)_+] < \infty \). Choose \( m_j = K \eta^{-j} \) as in (4.1). Then, with \( A = (\nu - 1)/(\alpha \nu \bar{\nu}) \):

(a)

\[
\lim_{t \to \infty} e^{-\alpha t} \mathbb{E}[|\text{BP}(t)|] = A,
\]

(b) there exists a constant \( C > 0 \), such that uniformly in \( t \in [0, \infty) \),

\[
e^{-2\alpha t} \mathbb{E}[|\text{BP}^{(\vec{m})}(t)|^2] \leq CK.
\]
(c) 
\[
\lim_{K \to \infty} \lim_{t \to \infty} e^{-\alpha t} \mathbb{E}\left[|\text{BP}_{k(t,x)} - \text{BP}_{k(t,x)}(t)|\right] = A \Phi(x) F_R(s),
\]
where $F_R$ is defined through (2.43) and $k(t,x) = t/\bar{v} + x \sqrt{t\bar{\sigma}^2/(\bar{v})^2}$.

(d) Replace in the above statements $\text{BP}$ by $\text{BP}_n$, with offspring $X_n$ depending on $n$ in such a way that $X_n \xrightarrow{d} X$, $\nu_n = \mathbb{E}[X_n] \to \mathbb{E}[X]$ and $\lim_n \mathbb{E}[X_n \log(X_n/K_n)] = 0$, for any $K_n \to \infty$. Furthermore, now define $m_j$ by $m_j = K_n \eta_n^{-j}$, with $\eta_n = \nu_n \int_0^\infty e^{-2\alpha_n s} dF_\xi(s)$, and replace $k(t,x)$ by $k_n(t,x)$ defined in (2.42). Then Part (a) and Part (c) hold with $\alpha = \alpha_n$ and $t = t_n$ and with the limits replaced by $\lim_n$, for any sequence $t_n \to \infty$. Similarly, under these conditions and substitutions, Part (b) holds for all $n \geq 1$, with $K$ replaced by $K_n$, uniformly in $t$.

**Proof.** We start by proving Proposition 4.1(a). The first claim of Part (a) is proved in [21], Theorem 17.1.

We bound the first moment of the difference between the truncated and the original branching process. Let $\nu^{(j)} = \mathbb{E}[(X \wedge m_j)]$. We compute that for $t > 0$,
\[
\begin{align*}
\mathbb{E}\left[ \sum_{k=0}^\infty |\text{BP}_k(t) - \text{BP}_k^{(m)}(t)| \right] &= e^{-\alpha t} \sum_{k=0}^\infty \left[ \nu^k - \prod_{j=1}^k \nu^{(j)} \right] \left[ F_\xi^*k(t) - F_\xi^{*(k+1)}(t) \right],
\end{align*}
\]
with $F_\xi^*k$ denoting the $k$-fold convolution of $F_\xi$, and where, by convention, $F_\xi^*0(t) = 1$ for every $t \geq 0$. In order to bound the differences $\nu^k - \prod_{j=1}^k \nu^{(j)}$, we rely on the following lemma.

**Lemma 4.2 (Effect of truncation on expectation CTBP).** Under the conditions of Proposition 4.1, uniformly in $k \geq 1$,
\[
\begin{align*}
1 - \prod_{j=1}^k \frac{\nu^{(j)}}{\nu} &\leq (\log(1/\eta))^{-1} \mathbb{E}[X \log (X/K)_+] = o_K(1),
\end{align*}
\]
where $o_K(1)$ denotes a quantity that converges to zero as $K \to \infty$.

**Proof.** Since $\nu^{(j)} \leq \nu$, for all $j \geq 1$, it is easily shown by induction that
\[
1 - \prod_{j=1}^k \frac{\nu^{(j)}}{\nu} \leq \sum_{j=1}^k \left(1 - \frac{\nu^{(j)}}{\nu}\right) \leq \sum_{j=1}^\infty \left(1 - \frac{\nu^{(j)}}{\nu}\right).
\]
Now, using that \( \nu > 1 \),
\[
\sum_{j=1}^{\infty} \left( 1 - \frac{\nu^{(j)}}{\nu} \right) \leq \sum_{j=1}^{\infty} \mathbb{E}[X \mathbb{1}_{X > m_j}] = \mathbb{E} \left[ X \sum_{j=1}^{\infty} \mathbb{1}_{m_j < X} \right],
\]
and we note that the number of \( j \) for which \( m_j = K \eta^{-j} < x \) is at most \( \log (x/K)/\log (1/\eta) \) \( \vee \) 0. Therefore, the inequality in (4.7) holds. Since \( \mathbb{E}[X \log (X/K)] = o_K(1) \), the equality in (4.7) follows. \( \square \)

By Lemma 4.2 and (4.6),
\[
e^{-\alpha t} \mathbb{E} \left[ \sum_{k=0}^{\infty} \left[ |BP_k(t)| - |BP^{(\tilde{m})}_k(t)| \right] \right] = o_K(1) e^{-\alpha t} \mathbb{E} \left[ \sum_{k=0}^{\infty} |BP_k(t)| \right],
\]
which completes the proof of Proposition 4.1(a).

We continue with the proof of the second moment estimate in Proposition 4.1(b). We follow the proof in [38], keeping attention to the truncation. We introduce the generating functions
\[
h(s) = \mathbb{E}[s^{X}], \quad h_j(s) = \mathbb{E}[s^{(X \wedge m_j)}],
\]
where \( m_j \) is given by (4.1). Parallel to calculations in the proof of [38], Lemma 4,
\[
\mathbb{E} \left[ \left| BP^{(\tilde{m})}_k(t) \right|^2 \right] = h''_1(1) \left( \mathbb{E} \left[ \left| BP^{(\tilde{m}_1)}_k(t) \right|^2 \right] \right)^2 \ast F_{\bar{\xi}} + h'_1(1) \mathbb{E} \left[ \left| BP^{(\tilde{m}_1)}_k(t) \right|^2 \right] \ast F_{\bar{\xi}},
\]
where \( \tilde{m}_1 = (m_2, m_3, \ldots) \) is \( \tilde{m} \) with the first element removed, and where for simplicity of reading the argument \( t \) has been left out. Transforming to
\[
\left| \overline{BP}^{(\tilde{m})}_k(t) \right| = e^{-\alpha t} \left| BP^{(\tilde{m})}_k(t) \right|,
\]
we obtain, after multiplying both sides of (4.12) by \( e^{-2\alpha t} \),
\[
\mathbb{E} \left[ \left| \overline{BP}^{(\tilde{m})}_k(t) \right|^2 \right] = \frac{\eta h''_1(1)}{\nu} \left( \mathbb{E} \left[ \left| BP^{(\tilde{m}_1)}_k(t) \right|^2 \right] \right)^2 \ast Q + \frac{\eta h'_1(1)}{\nu} \mathbb{E} \left[ \left| BP^{(\tilde{m}_1)}_k(t) \right|^2 \right] \ast Q,
\]
where
\[
\bar{F}_{\bar{\xi}}(t) = \nu \int_0^t e^{-\alpha y} dF_{\bar{\xi}}(y),
\]
\[
Q(t) = \eta^{-1} \int_0^t e^{-\alpha y} d\bar{F}_{\bar{\xi}}(y) = \eta^{-1} \nu \int_0^t e^{-2\alpha y} d\bar{F}_{\bar{\xi}}(y),
\]
and where we recall that \( \eta = \int_0^\infty e^{-\alpha y} d\bar{F}_{\bar{\xi}}(y) < 1 \) and \( \nu = h'(1) \). Iteration of (4.14) yields
\[
\mathbb{E} \left[ \left| \overline{BP}^{(\tilde{m})}_k(t) \right|^2 \right] = \sum_{j=1}^{\infty} f_1 \cdots f_{j-1} e_j \left( \mathbb{E} \left[ \left| \overline{BP}^{(\tilde{m}_j)}_k(t) \right|^2 \right] \right) \ast Q^* j,
where
\begin{equation}
(4.17) \quad e_j = \frac{\eta h''_j(1)}{\nu}, \quad f_j = \frac{\eta h'_j(1)}{\nu},
\end{equation}
and where \( \tilde{m}_j = (m_{j+1}, m_{j+2}, \ldots) \). Obviously, for \( t \to \infty \),
\begin{equation}
(4.18) \quad \mathbb{E}[|\mathbb{B}^\tilde{m}_j(t)|] \leq \mathbb{E}[|\mathbb{B}^\tilde{m}_j(t)|] \leq \mathbb{E}[|\mathbb{B}^m(t)|] \to A,
\end{equation}
by Part (a). Hence, provided that the sum \( \sum_{j \geq 1} f_1 \cdots f_{j-1} e_j \) converges, which we will establish in Lemma 4.3 below, we have that, uniformly in \( t \),
\begin{equation}
(4.19) \quad \mathbb{E}[|\mathbb{B}^\tilde{m}_j(t)|^2] \leq C \sum_{j=1}^\infty f_1 \cdots f_{j-1} e_j,
\end{equation}
for some constant \( C \geq A^2 \).

**Lemma 4.3** (Effect of truncation on variance of CTBP). For \( m_j = K \eta^{-j}, \) and with \( \nu = \mathbb{E}[X] > 1 \),
\begin{equation}
(4.20) \quad \sum_{j=1}^\infty f_1 \cdots f_{j-1} e_j \leq \frac{2\nu K}{1-\eta}.
\end{equation}

**Proof.** We bound \( f_j \leq \eta \), and
\begin{equation}
(4.21) \quad e_j \leq \eta \mathbb{E}[(X \wedge m_j)^2] = \eta (m_j^2 \mathbb{P}(X > m_j) + \mathbb{E}[X^2 \mathbb{1}_{\{X \leq m_j\}}]),
\end{equation}
so that
\begin{equation}
(4.22) \quad \sum_{j=1}^\infty f_1 \cdots f_{j-1} e_j \leq \sum_{j=1}^\infty m_j^2 \mathbb{P}(X > m_j) \eta^j + \sum_{j=1}^\infty \mathbb{E}[X^2 \mathbb{1}_{\{X \leq m_j\}}] \eta^j.
\end{equation}
We bound both terms separately. The first contribution equals
\begin{equation}
(4.23) \quad K^2 \sum_{j=1}^\infty \mathbb{P}(X > K \eta^{-j}) \eta^{-j} = K^2 \mathbb{E} \left[ \sum_{j=1}^\infty \eta^{-j} \mathbb{1}_{\{K \eta^{-j} < X\}} \right] \\
= K^2 \mathbb{E} \left[ \frac{\eta^{-a(X)} - 1}{1-\eta} \right],
\end{equation}
where \( a(x) = \max\{j : K \eta^{-j} < x\} = \lfloor \log (x/K) / \log (1/\eta) \rfloor \). Therefore, \( \eta^{-a(X)} \leq X/K \), so that
\begin{equation}
(4.24) \quad \sum_{j=1}^\infty m_j^2 \mathbb{P}(X > m_j) \eta^j \leq \frac{K^2}{1-\eta} \mathbb{E}[X/K] = \frac{K \nu}{1-\eta}.
\end{equation}
The second contribution is bounded in a similar way as
\[
\sum_{j=1}^{\infty} \mathbb{E}[X^2 \mathbb{1}_{X \leq j}] \eta_j = \mathbb{E} \left[ \sum_{j=1}^{\infty} X^2 \eta_j \mathbb{1}_{X \leq K \eta^j} \right] = \mathbb{E} \left[ X^2 \sum_{j=b(X)}^{\infty} \eta_j \right] = \mathbb{E} \left[ \frac{X^2 \eta^{b(X)}}{1-\eta} \right],
\]
where \( b(x) = \min\{j : K \eta^j \geq x\} \), so that \( \eta^{b(X)} \leq K/X \). Therefore,
\[
\sum_{j=1}^{\infty} \mathbb{E}[X^2 \mathbb{1}_{X \leq m_j}] \eta_j \leq K \mathbb{E}[X] = K \nu \frac{1-\eta}{1-\eta}.
\]

Proposition 4.1(b) follows by combining (4.19) and (4.20).

For Proposition 4.1(c), we start by showing that
\[
e^{-\alpha t} \sum_{j=0}^{k(t,x)} \mathbb{E}[|BP_j[t,t+s]|] \to A \Phi(x) F_R(s).
\]
Observe that \(|BP[t, t+s]| = \sum_{j=0}^{\infty} |BP_j[t, t+s]|\) is the total number of alive individuals at time \( t \), with residual lifetime at most \( s \), so that by applying [21], Lemma 2, Appendix Chapter VI, on the renewal equation
\[
\mathbb{E}[|BP[t, t+s]|] = F_\xi(t+s) - F_\xi(t) + \nu \int_0^{\infty} \mathbb{E}[|BP(t-y, t+s-y)|] dF_\xi(y),
\]
we readily obtain (compare the derivation of [21], Theorem 24.1),
\[
\lim_{t \to \infty} e^{-\alpha t} \mathbb{E}[|BP[t, t+s]|] = \lim_{t \to \infty} e^{-\alpha t} \sum_{j=0}^{\infty} \mathbb{E}[|BP_j[t, t+s]|] = A F_R(s),
\]
with \( A = (\nu - 1)/(\alpha \nu \tilde{v}) \) as given in the proposition. For fixed \( s > 0 \), define
\[
|BP_{>m}[t, t+s]| = \sum_{j=m+1}^{\infty} |BP_j[t, t+s]| = \sum_{j=m+1}^{\infty} e^{-\alpha t} |BP_j[t, t+s]|.
\]
Then (4.27) follows if we show that
\[
\mathbb{E}[|BP_{>m}[t, t+s]|] \to AF_R(s) - AF_R(s) \Phi(x) = AF_R(s) \Phi(-x).
\]
Conditioning on the lifetime (with cumulative distribution function equal to \( F_\xi \)) of the first individual,
\[
\mathbb{E}[|BP_j[t, t+s]|] = \nu \int_0^t \mathbb{E}[|BP_{j-1}[t-y, t+s-y]|] dF_\xi(y).
\]
Changing to $\mathcal{B}_j$ and $F_\xi$ and iteration of (4.31) yields

$$
(4.32) \quad E[|\mathcal{B}_{k(t,x)}[t,t+s]|] = \int_0^t E[|\mathcal{B}_j[t-y,t-y+s]|] dF_\xi^{*(k(t,x)+1)}(y),
$$

where $F_\xi^{*j}$ is the $j$-fold convolution of $F_\xi$, and hence the distribution function of the independent sum of $j$ copies of a random variable each having cumulative distribution function $F_\xi$. This is the point where we will use the CLT. Take an arbitrary $\varepsilon > 0$ and take $t_0$ so large so that for $t > t_0$,

$$
(4.33) \quad |E[|\mathcal{B}_{k(t,x)}[t,t+s]|] - AF_R(s)| \leq \varepsilon.
$$

Then

$$
(4.34) \quad |E[|\mathcal{B}_{k(t,x)}[t,t+s]|] - AF_R(s)\Phi(-x)| \leq \varepsilon F_\xi^{*(k(t,x)+1)}(t) + AF_R(s) dF_\xi^{*(k(t,x)+1)}(y).
$$

The last term vanishes since $E[|\mathcal{B}_j[t,t+s]|]$ is uniformly bounded and $F_\xi^{*(k(t,x))}(t-t_0) = o(1)$, as $t \to \infty$. Furthermore, with $m = k(t,x) \to \infty$,

$$
(4.35) \quad k(t,x) \sim \frac{t}{\bar{v}} + x \sqrt{t \frac{\bar{\sigma}^2}{\bar{v}^3}} \iff t \sim m \bar{v} - x \bar{\sigma} \sqrt{m}.
$$

As a result, by the CLT and the fact that $\bar{v}$ and $\bar{\sigma}^2$ are the mean and the variance of the distribution function $F_\xi$,

$$
(4.36) \quad \lim_{t \to \infty} F_\xi^{*(k(t,x))}(t) = \Phi(-x).
$$

Together with (4.34), this proves the claim in (4.30), and hence (4.27). Finally, we use the second statement of Part (a) to show that

$$
(4.37) \quad e^{-\alpha t} \sum_{j=0}^{k(t,x)} E[|\mathcal{B}_j[t,t+s]|] - e^{-\alpha t} \sum_{j=0}^{k(t,x)} E[|\mathcal{B}_j^{(m)}[t,t+s]|] \leq e^{-\alpha t}(E[|\mathcal{B}(t)|] - E[|\mathcal{B}^{(m)}(t)|]) \to 0,
$$

as first $t \to \infty$ and then $K \to \infty$. This shows Proposition 4.1(c).

We continue with the proof of Proposition 4.1(a) for the $n$-dependent CTBP. We denote the number of alive individuals at time $t$ in the $n$-dependent CTBP by $|\mathcal{B}_n(t)|$. We then have to show that for any sequence $t_n \to \infty$, as $n \to \infty$,

$$
(4.38) \quad e^{-\alpha n t_n} E[|\mathcal{B}_n(t_n)|] \to A,
$$
where $A = (\nu - 1)/\langle \nu \tilde{\nu} \rangle$. Denote by $\varphi(s) = \int_0^\infty e^{-sy} f_\xi(y) \, dy$, the Laplace transform of $f_\xi$, the density of the lifetime distribution $F_\xi$. Then

$$\int_0^\infty e^{-st} \mathbb{E}[|BP_n(t)|] \, dt = \frac{1 - \varphi(s)}{s(1 - \nu_n \varphi(s))}. \tag{4.39}$$

This equation follows directly from [21], equation (16.1), with $m$ replaced by $\nu_n$ and is valid when the real part of $s$ satisfies $\text{Re}(s) > \alpha_n$, where $\alpha_n > 0$ is defined as the unique solution to $\nu_n \varphi(\alpha_n) = 1$ [compare (1.15)]. From the inversion formula for Laplace transforms, we obtain

$$\mathbb{E}[|BP_n(t)|] = \frac{1}{2\pi i} \int_{\Gamma} e^{st} \frac{1 - \varphi(s)}{s(1 - \nu_n \varphi(s))} \, ds, \tag{4.40}$$

where $\Gamma$ is the path $(c_0 - i \infty, c_0 + i \infty)$, with $c_0 > \alpha_n$. Since $\alpha_n \to \alpha$ and $\nu_n \to \nu > 1$ and $\varphi(s)$ is the Laplace transform of a probability density, the function $s(1 - \nu_n \varphi(s))$ has a simple zero $s = \alpha_n$, but no other zeros in a small strip $|\text{Re}(s) - \alpha_n| < \epsilon$, for some $\epsilon > 0$. It is then easy to conclude from Cauchy’s theorem, calculating the residue at $s = \alpha_n$, that

$$\mathbb{E}[|BP_n(t_n)|] = e^{\alpha_n t_n} \left( 1 + O(e^{-\epsilon t_n}) \right), \tag{4.41}$$

where

$$A_n = \frac{\nu_n - 1}{\alpha_n \nu_n^2 \int_0^\infty y e^{-\alpha_n y} f_\xi(y) \, dy} = \frac{\nu_n - 1}{\alpha_n \nu_n \tilde{\nu}_n}. \tag{4.42}$$

Since $A_n \to A$, the claim (4.38) follows.

For the second statement in Proposition 4.1(a) for the $n$-dependent CTBP, we replace the inequality in (4.7) by the equivalent $n$-dependent statement, uniformly in $k \geq 1$,

$$1 - \prod_{j=1}^k \frac{\nu^{(j)}_n}{\nu_n} \leq (\log(1/\eta_n))^{-1} \mathbb{E}[X_n \log(X_n/K_n)_+] \tag{4.43}$$

Since $\lim_{n \to \infty} \mathbb{E}[X_n \log(X_n/K_n)_+] = 0$, as $n \to \infty$, the statement follows as in (4.10).

For the $n$-dependent case of Proposition 4.1(b), we need to show that for all $n \geq 1$ and uniformly in $t$,

$$e^{-2\alpha_n t} \mathbb{E}[|BP_n^{(m)}(t)|^2] \leq C K_n, \tag{4.44}$$

for some constant $C$, and where $K_n$ is the cut-off variable used in $m_j = K_n \eta_n^{-j}$. Copying the derivation which leads to (4.16), we obtain

$$\mathbb{E}[|BP_n^{(m)}|^2] = \sum_{j=1}^\infty f_1^{(n)} \cdots f_{j-1}^{(n)} e_j^{(n)} \mathbb{E}[|BP_n^{(m_j)}|^2]^{\ast Q_n^{j}}, \tag{4.45}$$
where

\[ e^{(n)}_j = \frac{\eta_n \mathbb{E}[(X_n \wedge m_j)^2]}{v_n}, \quad f^{(n)}_j = \frac{\eta_n \mathbb{E}[(X_n \wedge m_j)]}{v_n}, \]

and

\[ \bar{F}_{n, \xi}(t) = v_n \int_0^t e^{-\alpha_n y} dF_\xi(y), \quad Q_n(t) = \eta_n^{-1} v_n \int_0^t e^{-2\alpha_n y} dF_\xi(y). \]

From the proof of Lemma 4.3, we readily obtain that

\[ \sum_{j=1}^{\infty} f^{(n)}_1 \cdots f^{(n)}_{j-1} e^{(n)}_j \leq \sum_{j=1}^{\infty} m_j^2 \mathbb{P}(X_n > m_j) \eta_{n,j} + \sum_{j=1}^{\infty} \mathbb{E}[X_n^2 \mathbb{1}_{[X_n \leq m_j]}] \eta_n \leq \frac{2Knv_n}{1-\eta_n}. \]

Since \( v_n \to v \) and \( \eta_n \to \eta \), as \( n \to \infty \), we find, by combining (4.45) and (4.48), that given \( \varepsilon > 0 \), there is an \( n_0 \) so that for \( n > n_0 \), and uniformly in \( t \),

\[ e^{-2\alpha_n t} \mathbb{E}[|BP^n(t)|^2] \leq \frac{2Kn(v + \varepsilon)(A^2 + \varepsilon)}{(1-\eta - \varepsilon)} \leq CK_n. \]

By enlarging the constant \( C \), we see that (4.44) holds for all \( n \geq 1 \) and uniformly in \( t \).

Finally, we will give the proof of Proposition 4.1(c) for the \( n \)-dependent CTBP. We denote by \(|BP_{n,j}[t, t+s)|\) the number of individuals in generation \( j \) having residual lifetime at most \( s \) at time \( t \) of the CTBP with offspring given by \( X_n \). Then we obtain similarly as in (4.32),

\[ \mathbb{E}[|BP_{n,j}[t, t+s)|] = \int_0^t \mathbb{E}[|BP_{n}[t-y, t+s-y])|] d\bar{F}^{*}_{n, \xi}(y). \]

The expectation \( \mathbb{E}[|BP_{n}[t, t+s])| \) satisfies the renewal equation

\[ \mathbb{E}[|BP_{n}[t, t+s])|] = F_\xi(t+s) - F_\xi(t) + v_n \int_0^t \mathbb{E}[|BP_{n}[t-y, t+s-y])|] dF_\xi(y). \]

For \( s > 0 \) fixed, we denote by

\[ \tilde{K}_n(v, s) = \int_0^\infty e^{-vt} \mathbb{E}[|BP_{n}[t, t+s])|] dt, \]

\[ \tilde{f}(v, s) = \int_0^\infty e^{-vt}[F_\xi(t+s) - F_\xi(t)] dt, \]
the Laplace transforms of $\mathbb{E}[|BP_n[t, t+s]|]$ and $[F_{\xi}(t+s) - F_{\xi}(t)]$, respectively. Then (4.51) yields $\tilde{K}_n(v, s) = \tilde{f}(v, s)/(1 - \nu_n \varphi(v))$. From the inversion theorem for Laplace transforms, we obtain [compare (4.40)],

$$ \mathbb{E}[|BP_n[t, t+s]|] = \frac{1}{2\pi i} \int_{\Gamma} e^{vt} \tilde{f}(v, s) \frac{1}{(1 - \nu_n \varphi(v))} dv, $$

where $\Gamma$ is the same path as in (4.40), so that from the theory of residues, for some $\varepsilon > 0$,

$$ \mathbb{E}[|BP_n[t_n, t_n+s]|] = \frac{e^{\alpha_n t_n} \tilde{f}(\alpha_n, s)}{-\nu_n \varphi'(\alpha_n)} (1 + O(e^{-\varepsilon t_n})) $$

(4.54)

$$ = e^{\alpha_n t_n} A_n F_{n,R}(s) (1 + O(e^{-\varepsilon t_n})), $$

with $A_n$ defined in (4.42) and with

$$ F_{n,R}(s) = \frac{\alpha_n \nu_n}{\nu_n - 1} \int_0^\infty e^{-\alpha_n y} [F_{\xi}(y+s) - F_{\xi}(y)] dy. $$

Since $A_n \to A$ and $F_{n,R}(s) \to F_R(s)$ for $n \to \infty$ [see (2.43) for the definition of $f_R$], we obtain that, for any sequence $t_n \to \infty$,

$$ \lim_{n \to \infty} \mathbb{E}[|BP_n[t_n, t_n+s]|] = \lim_{n \to \infty} e^{-\alpha_n t_n} \mathbb{E}[|BP_n[t_n, t_n+s]|] = A F_R(s). $$

The $n$-dependent definition $k_n(t_n, x)$ yields that $m = k_n(t_n, x) \to \infty$ implies $t_n \sim m \tilde{\nu}_n - x \tilde{\sigma} \sqrt{m}$, so that since $\tilde{\sigma}_n \to \tilde{\sigma}$,

$$ \tilde{F}_{n,\xi}(k_n(t_n, x)) \to \Phi(-x). $$

(4.56)

Since $\nu_n \to \nu$, $\alpha_n \to \alpha$, we obtain, similarly as in (4.30) and for any sequence $t_n \to \infty$, that

$$ \mathbb{E}[|BP_{n,k_n(t_n,x)}[t_n, t_n+s]|] $$

(4.57)

$$ = \int_0^{t_n} \mathbb{E}[|BP_n[t_n-y, t_n-y+s]|] d\tilde{F}_{n,\xi}(k_n(t_n, x) + 1) = AF_R(s) \Phi(-x). $$

The remaining details of the proof follow from Part (a) and an argument as in (4.37). □

5. Bounds on the coupling: Proof of propositions 2.2 and 3.2.

5.1. Some simple lemmas concerning miscouplings. In (2.26), we have coupled the forward degrees in the SWT $(Y_k^{(n)})_{k \geq 1}$, as well as the (possibly) reduced forward degrees $(X_k^{(n)})_{k \geq 1}$, to an i.i.d. sequence $(B_k^{(n)})_{k \geq 1}$ with distribution equal to that of $D_n^* - 1$ given in (2.5).
We next investigate some simple consequences of this coupling. For this, it will be useful to note that when $D_n$, having distribution function $F_n$ in (1.2), satisfies Condition 1.1(c), the maximal degree $\Delta_n = \max_{i \in [n]} d_i$ satisfies
\begin{equation}
\Delta_n = o(\sqrt{n/\log n}).
\end{equation}
Indeed, suppose that $\Delta_n \geq \varepsilon \sqrt{n/\log n}$. Then, pick $K_n = n^{1/4}$ to obtain that
\begin{equation}
\mathbb{E}[D_n^2 \log (D_n/K_n)] = \frac{1}{n} \sum_{k=1}^{n} d_k^2 \log(d_k/n^{1/4}) \geq \frac{\Delta_n^2}{n} \log(\Delta_n/n^{1/4})
\end{equation}
\begin{equation}
\geq n^{-1}(\varepsilon \sqrt{n/\log n})^2 \log(n^{1/4}/(\log n)^{1/2}) \geq \varepsilon^2/8.
\end{equation}
This is in contradiction to Condition 1.1(c), so we conclude that (5.1) holds.

On $(\Omega, \mathcal{F}, P)$ we define the sigma-algebra $\mathcal{G}_k$ by
\begin{equation}
\mathcal{G}_k = \sigma(d_{U_1}, d_{U_2}, \tau_j, X_j^{(n)}, Y_j^{(n)}, (B_{t_i}^{(n)})_{i \leq j})_{j \leq k}.
\end{equation}
In the following lemma, we investigate the conditional probability of $Y_k^{(n)} \neq B_{t_{k-1}+1}^{(n)}$ given $\mathcal{G}_{k-1}$. In its statement, we recall the definition of $S_k^{(n)}$ in (2.27).

**Lemma 5.1 (Miscoupling of forward degree).** Assume that Condition 1.1(c) holds. For all $k \leq m_n$, and assuming that $m_n \leq \sqrt{n \log n}$,
\begin{equation}
P(Y_k^{(n)} \neq B_{t_{k-1}+1}^{(n)} | \mathcal{G}_{k-1}) \leq \frac{1}{\ell_n(1 - o(1))} \left( S_0^{(n)} + \sum_{s=1}^{k-1} (Y_s^{(n)} + 1) \right) = o_P(1).
\end{equation}

**Proof.** We have that $Y_k^{(n)} \neq B_{t_{k-1}+1}^{(n)}$ precisely when we pair the half-edge $y_{k-1}^*$ to a half-edge of a previously chosen vertex. Now let $V_{z_0}, \ldots, V_{z_{k-2}}$ be the previously chosen vertices and let $Y_s^{(n)} = B_{t_s}^{(n)}$, for $s \leq k-1$, be the forward degree of vertex $V_{z_s-1}$, $s \leq k-1$. Then the total number of half-edges incident to chosen vertices is at most
\begin{equation}
S_0^{(n)} + \sum_{s=1}^{k-1} (Y_s^{(n)} + 1).
\end{equation}
By (5.1), $\Delta_n = o(\sqrt{n/\log n})$, so that
\begin{equation}
S_0^{(n)} + \sum_{s=1}^{k-1} (Y_s^{(n)} + 1) \leq (k+1)\Delta_n \leq (m_n + 1)\Delta_n = o(n).
\end{equation}
From (5.5), it is clear that we draw each time from at least $\ell_n - o(n) = \ell_n(1 - o(1))$ half-edges. This shows (5.4). □
Lemma 5.2 (Probability of drawing at least one alive half-edge). Assume that Condition 1.1(c) holds. For all $k \leq m_n$, and assuming that $m_n \leq \sqrt{n \log n}$,

\begin{equation}
\mathbb{P}(X_k^{(n)} < Y_k^{(n)} \mid G_{k-1}) \leq \frac{\mathbb{E}[Y_k^{(n)} \mid G_{k-1}]}{\ell_n(1 - o(1))} \left(S_0^{(n)} + \sum_{s=1}^{k-1} Y_s^{(n)}\right).
\end{equation}

**Proof.** Recall the definition of $S_k^{(n)}$ in (2.27). We have that $X_k^{(n)} < Y_k^{(n)}$ when we pair at least one of the $Y_k^{(n)}$ half-edges to a half-edge incident to $\{U_1, U_2, V_{z_0}, \ldots, V_{zk-2}\}$. Since there are precisely $Y_k^{(n)}$ half-edges that need to be paired, and the number of half-edges incident to $\{U_1, U_2, V_{z_0}, \ldots, V_{zk-2}\}$, given $G_{k-1}$, equals $S_{k-1}^{(n)}$, we find

\begin{equation}
\mathbb{P}(X_k^{(n)} < Y_k^{(n)} \mid G_{k-1}, Y_k^{(n)}) \leq \frac{Y_k^{(n)} \cdot S_{k-1}^{(n)}}{\ell_n - \sum_{s=1}^{k-1} (Y_s^{(n)} - 1) - S_0^{(n)} - 1}.
\end{equation}

Clearly, $S_{k-1}^{(n)} \leq S_0^{(n)} + \sum_{s=1}^{k-1} Y_s^{(n)}$. Consequently, we obtain (5.6) from the tower-property for conditional expectations. \qed

5.2. Proof of Proposition 2.2(a). The i.i.d. sequences $(B_i^{(n)})_{i \geq 1}$ and $(B_i)_{i \geq 1}$ have probability mass functions $(g_k^{(n)})_{k \geq 0}$ and $(g_k)_{k \geq 0}$ given in (2.4) and (2.5), respectively. Since $(g_k^{(n)})_{k \geq 0}$ and $(g_k)_{k \geq 0}$ are discrete distributions and since by Condition 1.1, the distribution $(g_k^{(n)})_{k \geq 1}$ converges as $n \to \infty$ in distribution to $(g_k)_{k \geq 1}$, it follows that

\begin{equation}
d_{TV}(B_1^{(n)}, B_1) = \frac{1}{2} \sum_{k=0}^{\infty} |g_k - g_k^{(n)}| \to 0,
\end{equation}

where $d_{TV}$ denotes the total variation distance; see, for instance, [40], Theorem 6.1.

Take $s_n \to \infty$ such that

\begin{equation}
e^{2\alpha s_n} d_{TV}(B_1^{(n)}, B_1) \to 0.
\end{equation}

According to (1.20), and with $i \in \{1, 2\}$, we obtain

\begin{equation}
e^{-\alpha s_n} |\mathbb{B}^{(i)}(s_n)| \xrightarrow{a.s.} \tilde{\mathbb{W}}^{(i)},
\end{equation}

where $\tilde{\mathbb{W}}^{(i)}$ are two independent copies of $\mathbb{W}$. Since $\mathbb{P}(\tilde{\mathbb{W}}^{(i)} < \infty) = 1$ and $e^{\alpha s_n} \to \infty$, we conclude that $|\mathbb{B}(s_n)| \leq k_n$, w.h.p., if we take $k_n = \lfloor e^{2\alpha s_n} \rfloor$. If this $k_n$ does not satisfy $k_n = o(\sqrt{n})$, then we lower $s_n$ so that the corresponding value of $k_n = \lfloor e^{2\alpha s_n} \rfloor$ does satisfy $k_n = o(\sqrt{n})$.

Recall the definition of $G_k$ in (5.3). By Boole’s inequality,

\begin{equation}
\mathbb{P}(X_k^{(n)} \neq B_k^{(n)} \mid G_{k-1}) \leq \mathbb{P}(X_k^{(n)} < Y_k^{(n)} \mid G_{k-1}) + \mathbb{P}(Y_k^{(n)} \neq B_k^{(n)} \mid G_{k-1}).
\end{equation}
Using Lemmas 5.1–5.2, and by taking the expectation, a lower bound for the probability of all forward-degree-couplings being successful during the first $k_n = o(\sqrt{n})$ pairings is

$$\Pr\left(\bigcap_{k=1}^{k_n} X_k(n) = B_k(n)\right) = 1 - \Pr\left(\bigcup_{k=1}^{k_n} \{X_k(n) \neq B_k(n)\}\right)$$

$$\geq 1 - \frac{1}{\ell_n(1 - o(1))} \sum_{k=1}^{k_n} \mathbb{E}\left[S_0^{(n)} + \sum_{s=1}^{k-1} (Y_s^{(n)} + 1)\right]$$

$$- \frac{1}{\ell_n(1 - o(1))} \sum_{k=1}^{k_n} \mathbb{E}\left[\mathbb{E}\left[Y_k^{(n)} | G_{k-1}\right] \left(S_0^{(n)} + \sum_{s=1}^{k-1} Y_s^{(n)}\right)\right]$$

$$\geq 1 - ck^2_n/n \to 1,$$

where we rely on the inequality

$$\mathbb{E}\left[Y_k^{(n)} | G_{k-1}\right] \leq \sum_{j \in [n]} \frac{d_j(d_j - 1)}{\ell_n - 2k_n \Delta_n} = v_n(1 + o(1)),$$

whenever $k_n \Delta_n = o(n)$, which follows from (5.1).

The lower bound (5.12) implies that, w.h.p., the number of half-edges $(|AH(s)|)_{s \in [0, n]}$ is perfectly coupled to the number of (alive) individuals of the $n$-dependent CTBP $(|A_n(s)|)_{s \in [0, n]}$. In turn, (5.9) shows that $(BP_n(s))_{s \leq sn}$ is w.h.p. perfectly coupled to $(BP(s))_{s \leq sn}$. This proves Proposition 2.2(a).

We close this section by investigating moments of the size-biased random variables $(Y_k^{(n)})_{k \geq 1}$, which play a crucial role in the remainder of the paper.

**Lemma 5.3 (Moments of size-biased samplings).** Assume that Condition 1.1(a)–(c) holds. For all $k \leq mn$, and assuming that $mn \leq \sqrt{n \log n}$, and for any $K_n \to \infty$ such that $K^2_n = o(n/mn)$,

$$\mathbb{E}\left[Y_k^{(n)} \mathbb{1}_{\{Y_k^{(n)} \leq Kn\}} | G_{k-1}\right] = (1 + o(1))v_n,$$

$$\mathbb{E}\left[Y_k^{(n)} \mathbb{1}_{\{Y_k^{(n)} > Kn\}} | G_{k-1}\right] = o_{P}(1).$$

**Proof.** We use the upper bound

$$\mathbb{E}\left[Y_k^{(n)} \mathbb{1}_{\{Y_k^{(n)} \geq a\}} | G_{k-1}\right] \leq \frac{1}{\ell_n(1 - o(1))} \sum_{l \in [n]} d_l(d_l - 1)\mathbb{1}_{\{d_l \geq a\}},$$

where we again use that, since $mn \leq \sqrt{n \log n}$,

$$\ell_n - S_0^{(n)} - \sum_{j=1}^{k-1} Y_j^{(n)} \geq \ell_n - 2m_n \Delta_n = \ell_n(1 - o(1)).$$
This provides the necessary upper bound in (5.14) by taking \( a = 0 \) and from the identity \( \nu_n = \sum_{l \in [n]} d_l (d_l - 1) / \ell_n \). For (5.15), this also proves the necessary bound, since by Condition 1.1(c),

\[
\frac{1}{\ell_n} \sum_{l \in [n]} d_l (d_l - 1) \mathbb{1}_{[d_l \geq K_n]} = o(1).
\]

For the lower bound in (5.14), we bound, instead,

\[
\mathbb{P}[Y_k^{(n)} \mathbb{1}_{[Y_k^{(n)} \leq K_n]} \mid G_{k-1}]
\geq \frac{1}{\ell_n (1 - o(1))} \left[ \sum_{l \in [n]} d_l (d_l - 1) \mathbb{1}_{[d_l \leq K_n]} \right.
- \left. \sum_{l \in [n]} d_l (d_l - 1) \mathbb{1}_{[d_l \geq K_n]} \mathbb{1}_{[l \text{ is chosen}]} \right],
\]

where the event “\( l \) is chosen” means that vertex \( l \) belongs to the set of already chosen vertices \( U_1, U_2, V_0, \ldots, V_{z_m n - 2} \). The first term equals \( \nu_n (1 + o(1)) \). The second term is a.s. bounded by \( m_n K_n^2 / \ell_n = o(1) \), since \( K_n^2 = o(n/m_n) \). □

### 5.3. Completing the coupling: Proof of Proposition 2.2(b)

In this section, we use Proposition 4.1 to prove Proposition 2.2(b). In order to bound the difference between BP\((t)\) and SWT\((t)\), we will introduce several events. Let \( B_n, C_n, \varepsilon_n, m_n, m_n \) denote sequences of real numbers for which \( B_n, C_n \to \infty \) and \( \varepsilon_n \to 0 \) arbitrarily slowly, and \( m_n \gg \sqrt{n}, m_n \ll \sqrt{n} \). Later in this proof, we will formulate precisely how to choose these sequences.

Define the event \( \mathcal{A}_n \) by

\[
\mathcal{A}_n = \left\{ \left| \bigcup_{s \in [0, t_n + B_n]} \text{SWT}(s) \triangle \text{BP}_n(s) \right| < \varepsilon_n \sqrt{n} \right\},
\]

where we recall that \( \text{SWT}(s) \triangle \text{BP}_n(s) \) is the set of alive half-edges at time \( s \) that are miscoupled, and where we recall further that an alive half-edge is miscoupled if the shortest-weight path from the root to the vertex incident to that half-edge uses at least one degree-miscoupled vertex. Similarly, an alive individual is miscoupled if at least one of its ancestors is degree miscoupled. Note that when \( \mathcal{A}_n \) holds, then \( \left| \bigcup_{s \in [0, t]} \text{SWT}(s) \triangle \text{BP}_n(s) \right| < \varepsilon_n \sqrt{n} \) for any \( t \leq t_n + B_n \) by monotonicity in \( t \) [see Definition (2.35)].

In terms of the above notation, Proposition 2.2(b) can be reformulated as

\[
\mathbb{P}(\mathcal{A}_n^c \mid \mathcal{F}_s) = o_\varepsilon(1).
\]

Hence only a tiny fraction of the alive half-edges or individuals is miscoupled and the alive half-edges that are not miscoupled are connected to the root via a path containing only successfully coupled vertices.
In order to prove (5.21), we introduce the following events:

\[ B_n = \{ Y^{(BP)}(t_n + B_n) \leq \bar{m}_n \} \cap \{ Y^{(SWT)}(t_n + B_n) \leq \bar{m}_n \} \]

(5.22)

\[ C_n = \left\{ \| SWT(t) \| = \| BP_n(t) \|, \forall t \leq t_n - B_n \right\}, \]

(5.23)

\[ D_n = \{ \exists i \text{ such that } T_i \leq t_n + B_n, X^{(n)}_i \neq B^{(n)}_{t_i - 1 + 1}, dV_{z_i} \geq C_n \}, \]

(5.24)

where

\[ Y^{(BP)}(t) = \left| \{ v: v \in BP_n(s) \text{ for some } s \leq t \} \right|, \]

(5.25)

denotes the total number of individuals ever born into the BP before time \( t \) and

\[ Y^{(SWT)}(t) = \left| \{ v: v \in AH(s) \text{ for some } s \leq t \} \right|, \]

(5.26)

denotes the number of half-edges in the SWT that have ever been alive before time \( t \). Informally, on \( B_n \), the total number of half-edges in SWT and individuals in the CTBP are not too large. On \( C_n \), there is no early degree-miscoupled vertex, while on \( D_n \), there is no degree miscoupling involving a vertex that has high degree, until a later stage.

Obviously,

\[ \mathbb{P}(A_n^C | F_{s_n}) \]

(5.27)

\[ \leq \mathbb{P}(B_n^C | F_{s_n}) + \mathbb{P}(C_n^C \cap B_n | F_{s_n}) + \mathbb{P}(D_n^C \cap B_n \cap C_n | F_{s_n}) \]

\[ + \mathbb{P}(A_n^C \cap B_n \cap C_n \cap D_n | F_{s_n}). \]

To bound conditional probabilities of the form \( \mathbb{P}(E_n^C | F_{s_n}) \) as appearing in (5.21), we note that it suffices to prove that \( \mathbb{P}(E_n^C) = o(1) \), since then, by the Markov inequality and for every \( \epsilon > 0 \),

\[ \mathbb{P}(\mathbb{P}(E_n^C | F_{s_n}) \geq \epsilon) \leq \mathbb{E}[\mathbb{P}(E_n^C | F_{s_n})] / \epsilon = \mathbb{P}(E_n^C) / \epsilon = o(1). \]

(5.28)

Thus, we are left to prove that

\[ \mathbb{P}(B_n^C) = o(1), \quad \mathbb{P}(C_n^C \cap B_n) = o(1), \]

(5.29)

\[ \mathbb{P}(D_n^C \cap B_n \cap C_n) = o(1), \quad \mathbb{P}(A_n^C \cap B_n \cap C_n \cap D_n) = o(1). \]

We will do so in the above order.

**Lemma 5.4 (Expected number of particles born).** For all \( t \geq 0 \),

\[ \mathbb{E}[Y^{(BP)}(t)] = 2 \left( 1 - \frac{\mathbb{E}[D_n]F_\xi(t)}{v_n - 1} \right) + \frac{v_n}{v_n - 1} \mathbb{E}[|BP_n(t)|]. \]

(5.30)

Moreover, when \( e^{\alpha_n(t_n + B_n)} = o(\bar{m}_n) \) and \( e^{\alpha_n(t_n - B_n)} = o(\bar{m}_n) \), then

\[ \mathbb{P}(B_n^C) = o(1). \]

(5.31)
PROOF. Note that we grow two sets of alive half-edges and two BPs, which explains the factor 2 in (5.30). As is well known, the expected number of descendants in generation \( k \) of a BP equals \( \nu_k n \), where \( \nu_n \) denotes the mean offspring. Here, we deal with a delayed BP \( n \) where in the first generation the mean number of offspring equals \( E[D_n] \); the factor \( F^{(k)}_\xi (t) - F^{(k+1)}_\xi (t) \) represents the probability that an individual of generation \( k \) is alive at time \( t \). Together this yields

\[
\mathbb{E}[|BP_n(t)|] = \sum_{k=1}^{\infty} 2\mathbb{E}[D_n]v_n^{k-1} [F^{(k)}_\xi (t) - F^{(k+1)}_\xi (t)],
\]

(5.32)

\[
\mathbb{E}[Y^{(BP)}(t)] = 2 + \sum_{k=1}^{\infty} 2\mathbb{E}[D_n]v_n^{k-1} F^{(k)}_\xi (t).
\]

It is not difficult to deduce (5.30) from the two identities above.

To bound \( P(B_n^c) \), we note that we have to bound events of the form \( P(Y^{(BP)}(t) \geq m) \) and \( P(Y^{(SWT)}(t) \geq m) \) for various choices of \( m \) and \( t \). We use the Markov inequality and (5.30) to bound

\[
P(Y^{(BP)}(t) \geq m) \leq \frac{\mathbb{E}[Y^{(BP)}(t)]}{m} \leq \frac{\nu_n m}{m(\nu_n - 1)} \mathbb{E}[|BP_n(t)|] + \frac{2}{m}.
\]

(5.33)

According to (4.41), after conditioning on the offspring of the first individual,

\[
\mathbb{E}[|BP_n(t_n)|] = 2\mathbb{E}[D_n] A_n e^{\alpha_n t_n} (1 + o(1)),
\]

so that, since \( \mathbb{E}[D_n] \to \mathbb{E}[D] \),

\[
P(Y^{(BP)}(t_n) \geq m_n) = \Theta(e^{\alpha_n t_n} / m_n).
\]

(5.35)

The conditions on \( t \) and \( m \) in Lemma 5.4 have been chosen precisely so that \( e^{\alpha_n (t_n - B_n)} / m_n \to 0 \), and \( e^{\alpha_n (t_n + B_n)} / m_n \to 0 \).

We continue with \( P(Y^{(SWT)}(t) \geq m) \). We use the same steps as above, and start by computing

\[
\mathbb{E}[Y^{(SWT)}(t)] = 2 + 2 \sum_{k=0}^{\infty} F^{(k)}_\xi (t) \mathbb{E}[P^*_k],
\]

(5.36)

where \( P^*_0 = \ell_n / n \) and

\[
P^*_k = \sum_{|\pi| = k, \pi \subseteq CM_n(d)} (d_{\pi} - 1)/n, \quad k \geq 1,
\]

is the sum of the number of half-edges at the ends of paths of lengths \( k \) in \( CM_n(d) \), from a uniformly selected starting point. Following [27], Proof of Lemma 5.1, we find that

\[
\mathbb{E}[P^*_k] = \frac{1}{n} \sum_{v_0, \ldots, v_k} d_{v_0} \prod_{i=1}^{k} \frac{d_{v_i} (d_{v_i} - 1)}{\ell_n - 2i + 1} \leq \mathbb{E}[D_n]v_n^k,
\]

(5.37)
where the sum is taken over distinct vertices in \([n]\). Note that our definition of \(\nu_n\) in (1.16) deviates from the one given in [27], (2.3), which explains the difference between the right-hand side of (5.37) and the result in [27]. We obtain, for \(\nu_n \geq 1\), which holds for \(n\) sufficiently large,

\[
(5.38) \quad \mathbb{E}[Y^{(\text{SWT})}(t)] \leq 2 + 2\sum_{k=0}^{\infty} F^{(k)} \mathbb{E}[D_n] \nu_n^k \leq 2\mathbb{E}[D_n] + \nu_n \mathbb{E}[Y^{(\text{BP})}(t)],
\]

and we can repeat our arguments for \(\mathbb{E}[Y^{(\text{BP})}(t)]\).

**Lemma 5.5 (No early degree-miscoupling).** When \(m_n = o(\sqrt{n})\), then

\[
(5.39) \quad \mathbb{P}(C_n \cap B_n) = o(1).
\]

**Proof.** On \(B_n\), the inequality \(Y^{(\text{SWT})}(t_n - B_n) \leq m_n\) holds. By (5.12), the probability that there exists a degree-miscoupling before the draw of the \(m_n\)th half-edge is \(o(1)\) when \(m_n = o(\sqrt{n})\).

**Lemma 5.6 (No late miscouplings of high degree).** If \(m_n \leq \sqrt{n} \log n\), and \(C_n\) satisfies

\[
(5.40) \quad \frac{m_n^2}{\ell_n} \sum_{i \in [n]} d_i^2 1_{\{d_i \geq C_n\}} = o(n),
\]

then

\[
(5.41) \quad \mathbb{P}(D_n^c \cap B_n \cap C_n) = o(1).
\]

**Proof.** On \(B_n\), we have that \(Y^{(\text{SWT})}(t_n + B_n) \leq m_n\), and hence for a degree-miscoupling, when \(D_n^c\) holds, one of the vertices \(i \in [n]\) with \(d_i \geq C_n\) has to be chosen twice during the first \(m_n\) pairings. By Boole’s inequality and (5.40), an upper bound for this probability is

\[
(5.42) \quad \frac{m_n^2}{\ell_n} \sum_{i \in [n]} \frac{d_i^2}{(\ell_n - o(n))^2} 1_{\{d_i \geq C_n\}} = \frac{m_n^2}{(\ell_n - o(n))^2} \cdot \frac{\ell_n o(n)}{m_n^2} = o(1). \tag*{□}
\]

**Proposition 5.7 (Degree-miscoupled half-edges have small offspring).** If \(m_n \leq \sqrt{n} \log n\) and \(e^{2\alpha_n B_n} C_n m_n^2 / \ell_n = o(\varepsilon_n \sqrt{n})\), then

\[
(5.43) \quad \mathbb{P}(A_n^c \cap B_n \cap C_n \cap D_n) = o(1).
\]

**Proof.** We split the proof into the contribution of the degree-miscoupled vertices in \(|\text{BP}_n(t) \setminus \text{SWT}(t)|\) and those in \(|\text{SWT}(t) \setminus \text{BP}_n(t)|\), \(t \leq t_n + B_n\).
A bound on $|\text{BP}_n(t) \setminus \text{SWT}(t)|$. Recall that at time $T_k$, the vertex $V_{z_k-1}$ is degree-miscoupled when at least one of the equalities in (2.26) fails.

When $Y_k^{(n)} \neq B_k^{(n)}$, we can give an upper bound on the contribution to $\text{BP}_n(t)\setminus\text{SWT}(t)$ of the trees of miscoupled individuals by drawing from the i.i.d. sequence $(B_i^{(n)})_{i \geq 1}$. As a result, the total contribution to $|\text{BP}_n(t) \setminus \text{SWT}(t)|, t \leq t_n + B_n$ can be bounded above by

$$Y_k^{(BP)}(t_n + B_n - T_k),$$

where, for different $k \geq 1$, $(Y_k^{(BP)}(t))_{t \geq 0}$ are independent CTBPs. On $C_n$, we have that $T_k \geq t_n - B_n$ so that $t_n + B_n - T_k \leq 2B_n$, while on $D_n$, each degree-miscoupling starts with a vertex with degree at most $C_n$. Therefore, using (4.41),

$$\mathbb{E}[Y_k^{(BP)}(t_n + B_n - T_k) 1_{C_n \cap D_n}] \leq C_n A_n e^{2\alpha_n B_n} (1 + o(1)).$$

On $B_n$, the expected number of miscouplings is at most $O(\frac{m^2}{\ell_n})$, hence

$$\mathbb{E}\left[\sum_{\{k : T_k \leq t_n + B_n\}} Y_k^{(BP)}(t_n + B_n - T_k) 1_{B_n \cap C_n \cap D_n}\right] \leq O(\frac{m^2}{\ell_n}) C_n e^{2\alpha_n B_n}.$$

By assumption, the right-hand side is $o(\varepsilon \sqrt{n})$. Therefore, by the Markov inequality,

$$\mathbb{P}\left(\left|\text{BP}_n(s)_{s \in [0, t_n + B_n]} \setminus \text{SWT}(s)_{s \in [0, t_n + B_n]}\right| \geq \varepsilon \sqrt{n}\right)$$

$$\cap B_n \cap C_n \cap D_n \cap \text{MISC} \right) = o(1),$$

where the intersection with MISC indicates that we only deal with degree-miscouplings of the form $Y_k^{(n)} \neq B_k^{(n)} + 1$.

When $Y_k^{(n)} \neq X_k^{(n)}$, a self-loop or cycle-creating event occurs and the two half-edges that form the last edge in the cycle are removed from $\text{SWT}(t)$, but they are kept in $\text{BP}_n(t)$. In case of a removal of a collision edge, only one individual is kept in $\text{BP}_n(t)$. Whether one or two individuals are kept in the $\text{BP}_n(t)$ is of no consequence for the argument below.

Again, on the event $B_n$, the expected number of degree miscouplings is bounded by $O(\frac{m^2}{\ell_n})$. Furthermore, on the event $B_n \cap C_n$, the expected offspring of the half-edges involved in cycle-creating events is at most

$$\nu_n \mathbb{E}[Y_k^{(BP)}(2B_n)],$$

where $(Y_k^{(BP)}(t))_{t \geq 0}$ is the total number of individuals that have ever been alive in a CTBP where all individuals have i.i.d. offspring with law $(g_k^{(n)})_{k \geq 0}$. Indeed, we have no information about the remaining lifetime of the half-edge involved in an event that is caused by $Y_k^{(n)} \neq X_k^{(n)}$. As a result, rather than waiting for the residual life-time to be completed, we instantaneously take as offspring an i.i.d. draw from $(g_k^{(n)})_{k \geq 0}$, and start the various $\text{BP}_n(t)$ from there (on the average there are $\nu_n$ of these $\text{BP}_n$). The total number of individuals ever alive only increases by this
change. On the event $\mathcal{D}_n$, we have that $\mathbb{E}[Y^{(BP)}(2B_n)] \leq C_n A_n e^{2\alpha_n B_n} (1 + o(1))$. By assumption, $\frac{m_n^2}{B_n} C_n A_n e^{2\alpha_n B_n} / \ell_n = O(\varepsilon_n \sqrt{n})$. Therefore, the total contribution to $|\mathcal{B}_n(t) \setminus \mathcal{S}(t)|, t \leq t_n + B_n$, due to degree miscoupling events of the kind $Y_k^{(n)} \neq X_k^{(n)}$ is $O(\varepsilon_n \sqrt{n})$, as required.

A bound on $|\mathcal{S}(t) \setminus \mathcal{B}_n(t)|$. By construction the number of miscoupled half-edges in $\mathcal{S}(s), s \in [0,t]$ at any time $t$ is bounded from above by

$$\sum_{j=1}^{\text{MIS}(t)} Y_j^{(\mathcal{S})}(t - T_j),$$

where $\text{MIS}(t)$ denotes the number of degree-miscoupled vertices and $Y_j^{(\mathcal{S})}(t - T_j)$ is the number of half-edges reached by the liquid during $[T_j, t)$, and which are in the tree with root $V_{z_j-1}$. On the event $\mathcal{C}_n$, we have that $T_1 \geq t_n - B_n$. Therefore, on the event $\mathcal{C}_n$,

$$|\mathcal{S}(t)_{t \in [0,t_n+B_n]} \setminus \mathcal{B}_n(t)_{t \in [0,t_n+B_n]}| \leq \sum_{j=1}^{\text{MIS}(t_n+B_n)} Y_j^{(\mathcal{S})}(2B_n).$$

By the Markov inequality,

$$\mathbb{P}(\{|\mathcal{S}(t)_{t \in [0,t_n+B_n]} \setminus \mathcal{B}_n(t)_{t \in [0,t_n+B_n]}| \geq \varepsilon_n \sqrt{n}\} \cap \mathcal{B}_n \cap \mathcal{C}_n \cap \mathcal{D}_n) \leq (\varepsilon_n \sqrt{n})^{-1} \mathbb{E}\left[ \mathbb{1}_{\mathcal{B}_n \cap \mathcal{D}_n} \sum_{j=1}^{\text{MIS}(t_n+B_n)} Y_j^{(\mathcal{S})}(2B_n) \right].$$

We rewrite

$$\mathbb{E}\left[ \mathbb{1}_{\mathcal{B}_n \cap \mathcal{D}_n} \sum_{j=1}^{\text{MIS}(t_n+B_n)} Y_j^{(\mathcal{S})}(2B_n) \right] \leq (1 + o(1)) \sum_{j \in [n]} \mathbb{P}(j \text{ is degree-miscoupled, } \mathcal{B}_n \cap \mathcal{D}_n) \mathbb{E}[Y^{(\mathcal{S})}(2B_n)]$$

$$\leq (1 + o(1)) \sum_{j \in [n]} \left( \frac{d_j m_n}{\ell_n} \right)^2 \mathbb{1}_{\{d_j < C_n\}} \mathbb{E}[Y^{(\mathcal{S})}(2B_n)],$$

where we use that, upon degree-miscoupling of vertex $j$, we redraw a vertex from the size-biased distribution, for which the number of half-edges found before time $2B_n$ is equal to $\mathbb{E}[Y^{(\mathcal{S})}(2B_n)](1 + o(1))$ since $m_n \leq \sqrt{n \log n}$ and $\mathcal{B}_n$ occurs. Since $\mathbb{E}[Y^{(\mathcal{S})}(t)] \leq 2\mathbb{E}[D_n] + v_n \mathbb{E}[Y^{(BP)}(t)]$, we obtain that

$$\mathbb{E}[Y^{(\mathcal{S})}(2B_n)] \leq v_n A_n e^{2\alpha_n B_n} (1 + o(1)).$$
Therefore, we arrive at

\[
\mathbb{E}
\left[
\mathbbm{1}_{\mathcal{B}_n \cap \mathcal{D}_n} \sum_{j=1}^{\text{MIS}(\mathcal{T}_n + \mathcal{B}_n)} Y_j^{(\text{SWT})}(2\mathcal{B}_n)
\right]
\leq \nu_n A_n e^{2\alpha_n B_n} (1 + o(1)) \sum_{j \in [n]} \left(\frac{d_j \bar{m}_n}{\ell_n}\right)^2 \mathbb{1}_{[d_j < C_n]}.
\]

(5.54)

Bounding \(\sum_{j \in [n]} d_j^2 \mathbb{1}_{[d_j < C_n]} \leq C_n \ell_n\), the right-hand side of (5.54) is bounded by \(\nu_n A_n e^{2\alpha_n B_n} (1 + o(1)) C_n \bar{m}_n^2 / \ell_n = o(\nu_n \sqrt{n})\). Combining this with (5.51) proves that \(|\text{SWT}(t)_{t \in [\mathcal{T}_n + \mathcal{B}_n]} \setminus \text{BP}_n(t)_{t \in [\mathcal{T}_n + \mathcal{B}_n]}| = o(\nu_n \sqrt{n})\) on \(\mathcal{B}_n \cap \mathcal{C}_n \cap \mathcal{D}_n\). \(\square\)

**PROOF OF PROPOSITION 2.2(b).** Take

(5.55)

\[m_n = \sqrt{n}/(\log \log n)^{\alpha/2}, \quad \bar{m}_n = \sqrt{n}(\log n)^{1/4},\]

and

(5.56)

\[B_n = \log \log n, \quad C_n = n^{1/4}, \quad \varepsilon_n = 1/\log n.\]

By Condition 1.1(c) applied with \(K_n = \sqrt{C_n}/e\), and using that \(\log \sqrt{C_n} \leq \log(eD_n / \sqrt{C_n})_+\) when \(D_n \geq C_n\),

\[
\frac{1}{n} \sum_{i \in [n]} d_i^2 \mathbb{1}_{[d_i \geq C_n]} = \mathbb{E}\left[D_i^2 \mathbb{1}_{[D_i \geq C_n]}\right] \leq \mathbb{E}\left[D_n^2 \log(e \cdot D_n / \sqrt{C_n})_+ / \log \sqrt{C_n}\right]
\]

(5.57)

\[= o((\log n)^{-1}),\]

which verifies (5.40). All other conditions in Lemmas 5.4–5.6 and Proposition 5.7 are straightforward. Therefore, (5.21) follows, which completes the proof of Proposition 2.2(b). \(\square\)

We close this section with the proof of Proposition 3.2.

**PROOF OF PROPOSITION 3.2.** The fact that

(5.58)

\[\lim_{K \to \infty} \limsup_{n \to \infty} \mathbb{P}(\tilde{T}_1^{(\text{col})} \geq K) = 0,\]

follows from the PPP-convergence in Theorem 3.1. It thus suffices to prove that \(\mathbb{P}(\tilde{T}_1^{(\text{col})} \leq -K)\) satisfies the same asymptotics.

Recall the definition of \(\tilde{T}_1^{(\text{col})}\) in (3.1) and that of \(t_n\) and \(\tilde{t}_n\) in (2.39). Since \((W_{s_n}^{(1)}, W_{s_n}^{(2)}) \xrightarrow{d} (W^{(1)}, W^{(2)})\), it suffices to prove that

(5.59)

\[\lim_{K \to \infty} \limsup_{n \to \infty} \mathbb{P}(T_1^{(\text{col})} \leq t_n - K) = 0.\]
Recall the definition of $Y^{(\text{SWT})}(t)$ in (5.26) and let $Y_1^{(\text{SWT})}(t)$ and $Y_2^{(\text{SWT})}(t)$ be the corresponding objects for $U_1$ and $U_2$. We have at most $Y_i^{(\text{SWT})}(t_n - K/2)$ half-edges incident to $\text{SWT}^{(i)}(t)$ at any time $t \leq t_n - K/2$, for $i = 1, 2$. When $T_1^{(\text{col})} \leq t_n - K$, one of the half-edges incident to $\text{SWT}^{(2)}(t)$ for some $t \leq t_n - K/2$ should be paired with one of the at most $Y_1^{(\text{SWT})}(t_n - K/2)$ half-edges incident to $\text{SWT}^{(1)}(t)$, which has probability at most

\begin{equation}
\frac{Y_1^{(\text{SWT})}(t_n - K/2)}{\ell_n(1 - o(1))}.
\end{equation}

Since there are at most $Y_2^{(\text{SWT})}(t_n - K/2)$ half-edges incident to $\text{SWT}^{(2)}(t)$ at any time $t \leq t_n - K/2$, we aim to show that, for $i = 1, 2$,

\begin{equation}
\lim_{K \to \infty} \limsup_{n \to \infty} \mathbb{P}(Y_i^{(\text{SWT})}(t_n - K/2) \geq \sqrt{n}e^{-\alpha_n K/4}) = 0.
\end{equation}

Indeed, when $Y_i^{(\text{SWT})}(t_n - K/2) \leq \sqrt{n}e^{-\alpha_n K/4}$, for $i = 1, 2$, the probability that $T^{(\text{col})}_1 \leq t_n - K$ is at most $Y_1^{(\text{SWT})}(t_n - K/2)Y_2^{(\text{SWT})}(t_n - K/2) \leq \sqrt{n}e^{-\alpha_n K/4} \sqrt{n}e^{-\alpha_n K/4} \leq \ell_n(1 - o(1)) = O(e^{-\alpha_n K/2}),$

which proves (5.59). To prove (5.61), we use the Markov inequality

\begin{equation}
\mathbb{P}(Y_i^{(\text{SWT})}(t_n - K/2) \geq \sqrt{n}e^{-\alpha_n K/4}) \leq n^{-1/2}e^{\alpha_n K/4}\mathbb{E}[Y_i^{(\text{SWT})}(t_n - K/2)],
\end{equation}

and use (5.38) followed by (5.30) and (5.34) to deduce that

\begin{equation}
\mathbb{E}[Y_i^{(\text{SWT})}(t_n - K/2)] \leq A_n e^{\alpha_n (t_n - K/2)(1 + o(1))} = O(1)\sqrt{n}e^{-\alpha_n K/2}.
\end{equation}

Substituting (5.64) into (5.63) proves (5.61), and thus Proposition 3.2. □

6. Height CLT and stable age: Proof of Proposition 2.3. We first prove Proposition 2.3(a). Throughout this proof, we abbreviate $k_n = k_n(t_n, x)$ as in (2.42). The proof contains several key steps.

Reduction to a single BP. We start by showing that, in order for Proposition 2.3(a) to hold, it suffices to prove that for $j \in \{1, 2\}$, $x, t \in \mathbb{R}$ and $s > 0$, such that $t + s < B_n$, conditionally on $\mathcal{F}_{s_n}$ and on $\mathcal{W}_{s_n}^{(j)}$ and $\mathcal{W}_{s_n}^{(2)} > 0$,

\begin{equation}
e^{-\alpha_n t_n} |\mathbb{P}\left[\mathcal{W}_{s_n}^{(j)}\left[\tilde{t}_n + t, \tilde{t}_n + t + s\right]\right]| \overset{p}{\longrightarrow} e^{\alpha t} \Phi(x) F_R(s) \sqrt{\mathcal{W}_{s_n}^{(j)} / \mathcal{W}_{s_n}^{(3-j)}},
\end{equation}

where we use (2.37) in Proposition 2.2(a) to see that $\sqrt{\mathcal{W}^{(j)} / \mathcal{W}^{(3-j)}} \in [\varepsilon, 1/\varepsilon]$ w.h.p. Indeed, by Proposition 2.2(b) and the fact that $e^{-\alpha_n t_n} = O(n^{-1/2})$, (6.1)
implies that for $t + s < B_n$,

$$
e^{-\alpha_n t_n} |\text{SWT}^{(j)}_{\leq k_n}[\tilde{t}_n + t, \tilde{t}_n + t + s]| = e^{-\alpha_n t_n} |\text{BP}^{(j)}_{n, \leq k_n}[\tilde{t}_n + t, \tilde{t}_n + t + s]| + e^{-\alpha_n t_n} o_p(\varepsilon_n \sqrt{n})
$$

\[ (6.2) \]

which proves Proposition 2.3(a) by the independence of the two CTBPs involved.

**Using the branching property.** To prove (6.1), we note that $(\text{BP}^{(j)}_n(s))_{s \geq s_n}$ is the collection of alive individuals in the different generations of a CTBP, starting from the alive particles in $\text{BP}^{(j)}_n(s_n)$. Then, conditionally on $\mathcal{F}_{s_n}$,

$$
|\text{BP}^{(j)}_{n, \leq k_n}[\tilde{t}_n + t, \tilde{t}_n + t + s]|
$$

\[ (6.3) \]

$$
= \sum_{i \in \text{BP}^{(j)}_n(s_n)} \sum_{k=1}^{k_n-G_i^{(j)}} |\text{BP}^{(i,j)}_{n,k}[\tilde{t}_n + t - s_n - R_i, \tilde{t}_n + t + s - s_n - R_i]|,
$$

where $G_i^{(j)}$ is the generation of individual $i \in \text{BP}^{(j)}_n(s_n)$, while $R_i = R_i(s_n)$ is its remaining lifetime at time $s_n$, and $(\text{BP}^{(i,j)}(t))_{t \geq 0}$ are i.i.d. CTBPs for different $i$, for which the offspring for each individual has distribution $(g_k^{(n)})_{k \geq 0}$, and where the branching process starts with one individual that dies immediately.

**Truncating the branching process.** We continue by proving that we can truncate the branching process at the expense of an error term that converges to zero in probability. We let $\text{BP}^{(i,j,\vec{m})}_n$ denote the branching process $\text{BP}^{(i,j)}_n$ obtained by truncating particles in generation $l$ (measured from the root $i$) by $m_l = K_n \eta_n^{-l}$. We take $K_n \to \infty$ such that $K_n e^{-\alpha_n s_n} = o(1)$. We first show that, as $t_n \to \infty$, we can replace $e^{-\alpha_n t_n} |\text{BP}^{(i,j)}_{n, \leq k_n}[\tilde{t}_n, \tilde{t}_n + s]|$ by $e^{-\alpha_n t_n} |\text{BP}^{(i,j,\vec{m})}_{n, \leq k_n}[\tilde{t}_n, \tilde{t}_n + s]|$, at the expense of an $o_p(1)$-term. Indeed, with

$$
\sum_{k=1}^{\infty} |\text{BP}^{(i,j,k)}_{n,\vec{m}}(t)| = \sum_{k=1}^{\infty} |\text{BP}^{(i,j,\vec{m})}_{n,k}(t)|, \quad |\text{BP}^{(i,j,\vec{m})}_{n,k}(t)| = \sum_{k=1}^{\infty} |\text{BP}^{(i,j,\vec{m})}_{n,k}(t)|,
$$

by the $n$-dependent version of Proposition 4.1(a) formulated in Proposition 4.1(d), which we apply to each of the individuals born at time $s_n + R_i$, and for each sequence $u_n \to \infty$, we have that

$$
e^{-\alpha_n u_n} \mathbb{E}[|\text{BP}^{(i,j,\vec{m})}_n(u_n)| - |\text{BP}^{(i,j,\vec{m})}_n(u_n)|] \leq e^{-\alpha_n u_n} \mathbb{E}[|\text{BP}^{(i,j)}_n(u_n)| - |\text{BP}^{(i,j)}_n(u_n)|] = o(1).
$$
Therefore, using that the law of \((\mathcal{B}_n^{(i,j)}(t))_{t \geq 0}\) only depends on \(\mathcal{F}_{s_n}\) through \(R_i, \bar{t}_n, e^{-\alpha n} \sum_{i \in \mathcal{B}_n^{(j)}(s_n)} \sum_{k=1}^{k_n-G^{(j)}_i} \mathbb{E}[|\mathcal{B}_n^{(i,j)}(\bar{t}_n + t - s_n - R_i)|]\)

\[-|\mathcal{B}_n^{(i,j,\bar{m})}(\bar{t}_n + t - s_n - R_i)|\mathcal{F}_{s_n}]\]

(6.5) \[
\leq e^{-\alpha n t_n} \sum_{i \in \mathcal{B}_n^{(j)}(s_n)} \mathbb{E}[|\mathcal{B}_n^{(i,j)}(\bar{t}_n + t - s_n - R_i)|] - |\mathcal{B}_n^{(i,j,\bar{m})}(\bar{t}_n + t - s_n - R_i)| |R_i, \bar{t}_n| \]

\[= o(1) \sum_{i \in \mathcal{B}_n^{(j)}(s_n)} e^{\alpha n} \bar{t}_n - t_n + t - s_n - R_i) = o_p(1)e^{-\alpha s_n} \sum_{i \in \mathcal{B}_n^{(j)}(s_n)} e^{-\alpha R_i},\]

since the random variable \(|\bar{t}_n - t_n|\) is tight, and assuming that \(s_n \to \infty\) so slowly that \(s_n|\alpha_n - \alpha| = o(1)\). Since \(R_i \geq 0\),

(6.6) \[
e^{-\alpha s_n} \sum_{i \in \mathcal{B}_n^{(j)}(s_n)} e^{-\alpha R_i} \leq e^{-\alpha s_n} |\mathcal{B}_n^{(j)}(s_n)| \xrightarrow{d} \tilde{W}^{(j)} = o_p(1),\]

by the “perfect” coupling between \(\mathcal{B}_n\) and \(\mathcal{B}\) at time \(s_n\) stated in Proposition 2.2(a), and using that [see (1.20)],

(6.7) \[
e^{-\alpha s_n} |\mathcal{B}_n^{(j)}(s_n)| \xrightarrow{d} \tilde{W}^{(j)}.
\]

We conclude that

\(e^{-\alpha n t_n} |\mathcal{B}_n^{(j)}(\bar{t}_n + t, \bar{t}_n + t + s)|\)

(6.8) \[
e^{-\alpha n t_n} \sum_{i \in \mathcal{B}_n^{(j)}(s_n)} \sum_{k=1}^{k_n-G^{(j)}_i} |\mathcal{B}_n^{(i,j,\bar{m})}(\bar{t}_n + t - s_n - R_i, \bar{t}_n + t + s - s_n - R_i)|
\]

\[+ o_p(1).
\]

**A conditional second moment method: Expectation.** We next use a conditional second moment estimate on the sum on the right-hand side of (6.8), conditionally on \(\mathcal{F}_{s_n}\). By the \(n\)-dependent version of Proposition 4.1(c) formulated in Proposition 4.1(d), for \(\bar{t}_n \to \infty\), and for each \(i \in \mathcal{B}_n^{(j)}(s_n),\)

(6.9) \[
e^{-\alpha n t_n} \mathbb{E}[|\mathcal{B}_n^{(i,j,\bar{m})}(\bar{t}_n, \bar{t}_n + s)|] \xrightarrow{d} A \Phi(x) F_R(s).
\]

Observe that also \(m = k_n - k_n \to \infty,\) with \(k_n = o(\sqrt{\log n}),\) implies \(t_n \sim m \bar{v}_n - x \bar{\sigma} \sqrt{m},\) so that we can conclude from (4.56) that (6.9) also holds with \(k_n\) replaced...
by $k_n - \bar{k}_n$, as long as $\bar{k}_n = o(\sqrt{\log n})$. As a result, when $i_n + t - s_n - R_i \xrightarrow{p} \infty$ and $\bar{k}_n = o(\sqrt{\log n})$ and for each $i$,

$$
e^{-\alpha_n(i_n + t - s_n - R_i)} \mathbb{E} \left[ \left| \mathbb{B}P_{n, i}^{(i,j,m)} \left[ i_n + t - s_n - R_i, \bar{i}_n + t + s - s_n - R_i \right] \right| \mid \mathcal{F}_{s_n} \right] = e^{-\alpha_n(i_n + t - s_n - R_i)}$$

$$\times \mathbb{E} \left[ \left| \mathbb{B}P_{n, i}^{(i,j,m)} \left[ i_n + t - s_n - R_i, \bar{i}_n + t + s - s_n - R_i \right] \right| \mid R_i, \bar{i}_n \right]$$

$$= A \Phi(x) F_k(s) \left[ 1 + o_P(1) \right].$$

Further, we use the general theory of vertex characteristics in [24], Theorem 6.10.1, to conclude that

$$\sum_{i \in \mathbb{B}P_{n}^{(i,j)}} e^{-\alpha_n(s_n + R_i)} \xrightarrow{p} \hat{W}^{(j)}/A.$$

Indeed, consider $Z(t) = \sum_{i \in \mathbb{B}P(t)} e^{-\alpha R_i} = \sum_{i} \mathbb{1}_{[\bar{\tau}_i, \xi_i]} e^{-\alpha(\xi_i + t - i, t)}$, where the second sum is taken over all individuals and where $\bar{\tau}_i$, $\xi_i$ are the birth and lifetime of individual $i$, respectively. Then $R_i = \bar{\tau}_i + \xi_i - t$, for $i \in \mathbb{B}P(t)$, and with the terminology of [24], Section 6.9, $Z(t) = Z^\chi(t) = \sum_{i} \chi_i(t - \bar{\tau}_i)$, where the random characteristic $\chi_i$ of individual $i$ is defined by

$$\chi_i(t) = \mathbb{1}_{[0, \xi_i]}(t)e^{-\alpha(\xi_i - t)}.$$ 

According to the aforementioned [24], Theorem 6.10.1,

$$e^{-\alpha t} Z^\chi(t) \xrightarrow{a.s.} c_{\chi} \hat{W}/k(\infty),$$

where $c_{\chi} = \int_{0}^{\infty} e^{-\alpha u} \mathbb{E} [\chi(u)] du / \int_{0}^{\infty} e^{-\alpha u} \mu(du)$, $k(\infty) = \int_{0}^{\infty} e^{-\alpha u} [1 - F_\xi(u)] du / \int_{0}^{\infty} e^{-\alpha u} \mu(du)$ and where $\mu(t) = v F_\xi(t)$. This yields $c_{\chi} = 1/v$, and $k(\infty) = (v - 1)/(\alpha v^2)$, so that $c_{\chi}/k(\infty) = 1/A$. The w.h.p. “perfect” coupling between $(\mathbb{B}P_n(s))_{s \leq s_n}$ with $(\mathbb{B}P(s))_{s \leq s_n}$ stated in Proposition 2.2(a) and using that we may take $s_n \to \infty$ so slowly that $s_n(\alpha - \alpha_n) = o(1)$ implies that (6.13) implies (6.11).

This yields that, conditionally on $\mathcal{F}_{s_n}$ and $W_{s_n}^{(1)}W_{s_n}^{(2)} > 0$, and when $G_i^{(j)} = \alpha_P(\sqrt{\log n})$ (which happens w.h.p. when $s_n$ is sufficiently small),

$$e^{-\alpha_n i_n} \sum_{i \in \mathbb{B}P_{n}^{(i,j)}} \mathbb{E} \left[ k_n - G_i^{(j)} \left[ \sum_{k=1}^{k_n-G_i^{(j)}} \left| \mathbb{B}P_{n,k}^{(i,j,m)} \left[ i_n + t - s_n - R_i, \bar{i}_n + t + s - s_n - R_i \right] \right| \right] \right]$$

$$= A e^{\alpha t} \Phi(x) F_k(s) \left[ 1 + o_P(1) \right] \sum_{i \in \mathbb{B}P_{n}^{(i,j)}} e^{\alpha_n(i_n - t - s_n - R_i)}$$

$$\xrightarrow{p} e^{\alpha t} \Phi(x) F_k(s) \sqrt{W^{(j)}/W^{(3-j)}},$$
by (6.11), again taking \( s_n \to \infty \) sufficiently slowly, and since \( e^{\alpha_n (t_n - \bar{t}_n)} = \sqrt{W^{(j)}_{s_n} W^{(3-j)}_{s_n}} \xrightarrow{p} \sqrt{W^{(j)} W^{(3-j)}} \) [see (2.39)]. Notice that in (6.14), we condition on \( W^{(j)}_{s_n} > 0 \), so that the limit \( \tilde{W}^{(j)} \) has to be replaced by \( \tilde{W}^{(j)} | \tilde{W}^{(j)} > 0 \) which is equal in distribution to \( W^{(j)} \).

A conditional second moment method: Variance. We next bound, conditionally on \( \mathcal{F}_{s_n} \), the variance of the sum on the right-hand side of (6.8). By conditional independence of \((BP^{(i,j)}_n)_{i \geq 1}\),

\[
e^{-2\alpha_n t_n} \text{Var} \left( \sum_{i \in BP^{(j)}_n(s_n)} \sum_{k=1}^{k_n-G_j^{(j)}} |BP^{(i,j,\bar{m})}_{n,k} (\bar{t}_n + t - s_n - R_i) \right) \bigg| \mathcal{F}_{s_n} \right)
\]

(6.15)

\[
e^{-2\alpha_n t_n} \sum_{i \in BP^{(j)}_n(s_n)} \text{Var} \left( \sum_{k=1}^{k_n-G_j^{(j)}} |BP^{(i,j,\bar{m})}_{n,k} (\bar{t}_n + t - s_n - R_i) \right) \bigg| \mathcal{F}_{s_n} \right).
\]

We bound

\[
\text{Var} \left( \sum_{k=1}^{k_n-G_j^{(j)}} |BP^{(i,j,\bar{m})}_{n,k} (\bar{t}_n + t - s_n - R_i) \right) \bigg| \mathcal{F}_{s_n} \right)
\]

\[
\leq \mathbb{E} \left[ \left( \sum_{k=1}^{k_n-G_j^{(j)}} |BP^{(i,j,\bar{m})}_{n,k} (\bar{t}_n + t - s_n - R_i) \right)^2 \bigg| \mathcal{F}_{s_n} \right]
\]

(6.16)

\[
\leq \mathbb{E} \left[ |BP^{(i,j,\bar{m})}_{n} (\bar{t}_n + t - s_n - R_i) \right|^2 \bigg| \mathcal{F}_{s_n} \right]
\]

\[
= \mathbb{E} \left[ |BP^{(i,j,\bar{m})}_{n} (\bar{t}_n + t - s_n - R_i) \right|^2 \bigg| R_i, \bar{t}_n \right].
\]

By the \( n \)-dependent version of Proposition 4.1(b) formulated in Proposition 4.1(d), for each \( n \geq 1 \),

\[
\sup_{t \geq 0} \{ e^{-2\alpha_n t} \mathbb{E} \left[ |BP^{(i,j,\bar{m})}_{n} (t) \right|^2 \bigg] \} \leq C K_n.
\]

(6.17)

As a result,

\[
e^{-2\alpha_n t_n} \text{Var} \left( \sum_{i \in BP^{(j)}_n(s_n)} \sum_{k=1}^{k_n-G_j^{(j)}} |BP^{(i,j,\bar{m})}_{n,k} (\bar{t}_n + t - s_n - R_i) \right)
\]
\begin{equation}
\tilde{t}_n + t + s - s_n - R_i \mid \mathcal{F}_{s_n} \leq e^{-2\alpha_n t_n} \sum_{i \in \text{BP}^{(j)}_{s_n}} \mathbb{E}[|\mathbb{B}^{(i,j)}_{n}(\tilde{t}_n + t - s_n - R_i)|^2 \mid R_i, \tilde{t}_n] \\
\leq C K_n e^{-2\alpha_n t_n} \sum_{i \in \text{BP}^{(j)}_{s_n}} e^{2\alpha_n (\tilde{t}_n + t - s_n - R_i)} \\
= C K_n e^{2\alpha_n (\bar{t}_n - t_n)} e^{-2\alpha_n s_n + 2\alpha_n t} \sum_{i \in \text{BP}^{(j)}_{s_n}} e^{-2\alpha_n R_i} \\
= O_p(1) K_n e^{-2\alpha_n s_n} \sum_{i \in \text{BP}^{(j)}_{s_n}} e^{-2\alpha_n R_i},
\end{equation}

since \(e^{2\alpha_n (\bar{t}_n - t_n) + 2\alpha_n t} = O_p(1)\). We can bound this further as in (6.6) and (6.7) by

\begin{equation} 
K_n e^{-2\alpha_n s_n} \sum_{i \in \text{BP}^{(j)}_{s_n}} e^{-2\alpha_n R_i} \leq K_n e^{-\alpha_n s_n} \mathbb{E}[|\mathbb{B}^{(j)}_{s_n}|] \\
= O_p(1) K_n e^{-\alpha_n s_n},
\end{equation}

which is \(o_p(1)\) precisely when \(K_n e^{-\alpha_n s_n} = o(1)\). Since we are free to choose \(K_n\), we can choose it such that \(K_n e^{-\alpha_n s_n} = o(1)\) indeed holds. By (6.18) and (6.19), the sum on the right-hand side of (6.8) is, conditionally on \(\mathcal{F}_{s_n}\), concentrated around its asymptotic conditional mean given in (6.14). As a result, (6.1) follows. This completes the proof of Proposition 2.3(a).

**Proof of Proposition 2.3(b).** In order to prove Proposition 2.3(b), we compare the statement of Proposition 2.3(b) with that of Proposition 2.3(a). Let \(m = |\text{SWT}^{(j)}_{s_n}(\bar{t}_n + t, \bar{t}_n + t + s_2)|\), so that \(m \to \infty\) on the event that \(W_{s_n}^{(1)} W_{s_n}^{(2)} > 0\), and consider the sum \(\sum_{i=1}^{m} X_i^*\), where \(X_i^* = d_{V_i} - 1\) are forward degrees of free vertices after time \(\bar{t}_n + t\), that is, from the vertices \([n]\), we remove the set \(S_m\) of all vertices of which at least one half-edge appeared in \((\text{SWT}(s))_{s \leq \bar{t}_n + t}\). We will prove that, conditionally on \(\mathcal{F}_{\bar{t}_n + t}\) with \(t < B_n\),

\begin{equation}
\frac{1}{m \nu_n} \sum_{i=1}^{m} X_i^* \overset{p}{\to} 1,
\end{equation}

and then the proof of Proposition 2.3(b) follows from the proof of Proposition 2.3(a), the fact that \(\nu_n \to \nu\), and because \(\mathcal{F}_{s_n} \subset \mathcal{F}_{\bar{t}_n + t}\).

Without loss of generality, we may assume that \(B_n\) in (5.22) holds, so that \(t < B_n\) implies that \(|S_m| \leq \bar{m}_n\), where \(\bar{m}_n = \sqrt{n} (\log n)^{1/4}\) as in (5.55). As a result the sequence \((d_i)_{i \in [n]} \cap S_m\) satisfies Condition 1.1 whenever \((d_i)_{i \in [n]}\) does. Hence, Lemma 5.3 holds with \(Y_i^{(n)}\) replaced by \(X_i^*\), so that in particular, from the Markov
inequality, conditionally on $F_{tn+t}$ and for every sequence $K_n \to \infty$ satisfying $K_n^2 = o(n/m_n)$,

\begin{equation}
\frac{1}{m} \sum_{i=1}^{m} X_i^{*1} \mathbb{1}_{X_i^{*} \leq K_n} \overset{p}{\to} 0.
\end{equation}

We use a conditional second moment method on $\sum_{i=1}^{m} X_i^{*1} \mathbb{1}_{X_i^{*} \leq K_n}$, conditionally on $F_{tn+t}$. By (5.14) in Lemma 5.3,

\begin{equation}
\mathbb{E}\left[ \sum_{i=1}^{m} X_i^{*1} \mathbb{1}_{X_i^{*} \leq K_n} \bigg| F_{tn+t} \right] = m\nu_n (1 + o_p(1)).
\end{equation}

This gives the asymptotics of the first conditional moment of $\sum_{i=1}^{m} X_i^{*1} \mathbb{1}_{X_i^{*} \leq K_n}$.

For the second moment, we start by bounding the covariances. We note that, for $1 \leq i < j \leq m$,

\begin{equation}
\text{Cov}(X_i^{*1} \mathbb{1}_{X_i^{*} \leq K_n}, X_j^{*1} \mathbb{1}_{X_j^{*} \leq K_n} \big| F_{tn+t})
\end{equation}

\begin{equation}
= \mathbb{E}[X_i^{*1} \mathbb{1}_{X_i^{*} \leq K_n} (\mathbb{E}[X_j^{*1} \mathbb{1}_{X_j^{*} \leq K_n} \big| F_{tn+t}, X_i^{*}, \ldots, X_i^{*}]
\end{equation}

\begin{equation}
- \mathbb{E}[X_j^{*1} \mathbb{1}_{X_j^{*} \leq K_n} \big| F_{tn+t})] \big| F_{tn+t}).
\end{equation}

By (5.14) in Lemma 5.3, as well as the fact that $i \leq m_n$,

\begin{equation}
\mathbb{E}\left[ X_i^{*1} \mathbb{1}_{X_i^{*} \leq K_n} \big| F_{tn+t}, X_1^{*}, \ldots, X_i^{*} \right] - \mathbb{E}\left[ X_j^{*1} \mathbb{1}_{X_j^{*} \leq K_n} \big| F_{tn+t} \right] = o_p(1),
\end{equation}

so that also

\begin{equation}
\text{Cov}(X_i^{*1} \mathbb{1}_{X_i^{*} \leq K_n}, X_j^{*1} \mathbb{1}_{X_j^{*} \leq K_n} \big| F_{tn+t}) = o_p(1).
\end{equation}

Further, a trivial bound on the second moment together with (5.14) in Lemma 5.3 yields that

\begin{equation}
\text{Var}(X_i^{*1} \mathbb{1}_{X_i^{*} \leq K_n} \big| F_{tn+t}) \leq K_n \mathbb{E}[X_i^{*1} \big| F_{tn+t}] = Kn\nu_n (1 + o_p(1)).
\end{equation}

As a result, whenever $K_n\nu_n = o(m^2)$, and $K_n^2 = o(n/m_n)$,

\begin{equation}
\text{Var}\left( \sum_{i=1}^{m} X_i^{*1} \mathbb{1}_{X_i^{*} \leq K_n} \big| F_{tn+t} \right) = o_p(m^2),
\end{equation}

which together with (6.22) proves that, conditionally on $F_{tn+t}$,

\begin{equation}
\frac{1}{m\nu_n} \sum_{i=1}^{m} X_i^{*1} \mathbb{1}_{X_i^{*} \leq K_n} \overset{p}{\to} 1.
\end{equation}

Together with (6.21), this proves (6.20), as required. □
7. The PPP limit for collision edges: Proof of Theorem 3.1. Recall that \((\mathcal{F}_t)_{t \geq 0}\) is the filtration generated by all the randomness used in the construction up to time \(t\), that is, \(\mathcal{F}_t = \sigma((\text{SWT}(s), \text{BP}_n(s), \text{BP}(s))_{s \in [0, t]}).\)

We will investigate the number of collision edges \((x_i, P_{x_i})\) with \(I(x_i) = j \in \{1, 2\}\), \(H(x_i) \leq k_n(t_n, x)\), \(H(P_{x_i}) \leq k_n(t_n, y)\) and \(R_{t_i^{(\text{col})}}(P_{x_i}) \in [0, s)\) created in the time interval \([\tilde{t}_n + t, \tilde{t}_n + t + \epsilon)\), where \(\epsilon > 0\) is small. We let \(I = [a, b) \times \{j\} \times (-\infty, x] \times (-\infty, y] \times [0, s)\) be a subset of \(S\), and we prove that

\[
\mathbb{P}(\Pi_n(I) = 0 | \mathcal{F}_{t_n}) \xrightarrow{p} \exp\left\{-\int_a^b 2\nu f_R(0) \frac{e^{2\alpha t}}{\mathbb{E}[D]} \frac{\Phi(x) \Phi(y) F_R(s) dt}{F_R(\epsilon)}\right\}.
\]

By [29], Theorem 4.7, this proves the claim.

We split

\[
\mathcal{I} = \bigcup_{\ell=1}^{N} \mathcal{I}_\ell^{(e)},
\]

where \(\mathcal{I}_\ell^{(e)} = [t_{\ell-1}^{(e)}, t_\ell^{(e)}] \times \{j\} \times (-\infty, x] \times (-\infty, y] \times [0, s]\), with \(t_\ell^{(e)} = a + \ell \epsilon\) and \(\epsilon = (b - a)/N\), with \(N \in \mathbb{N}\). We will let \(\epsilon \downarrow 0\) later on. For a fixed \(\epsilon > 0\), we say that a collision edge \((x_i, P_{x_i})\) is a first round collision edge when there exists \(l \in [N]\) and a half-edge \(y \in \text{AH}(t_{l-1}^{(e)})\) such that \(y\) is found by the liquid in the time interval \(\mathcal{I}_\ell^{(e)}\), \(y\) is paired to the half-edge \(P_y\) whose sibling half-edge \(x_i\) is paired to \(P_{x_i} \in \text{AH}(t_{l-1}^{(e)})\) with \(I(y) = j \neq I(P_{x_i}) = 3 - j\). We call all other collision edges second round collision edges. The second round collision edges are such that a half-edge \(y\) is found by the liquid in the interval \(\mathcal{I}_\ell^{(e)}\) (the first round), \(y\) is paired to the half-edge \(P_y\), one of the sibling half-edges \(x_i\) of \(y\) is then also found by the liquid in the time interval \(\mathcal{I}_\ell^{(e)}\) (the second round) and is paired to a half-edge \(P_{x_i}\), whose sibling half-edge \(z\) is paired to \(P_z \in \text{AH}(t_{l-1}^{(e)})\) with \(I(x_i) = j \neq I(P_z) = 3 - j\). When \(\epsilon > 0\) is quite small, the latter seems less likely, which is why we start with the first round collision edges.

Denote the point processes of first and second round collision edges by \(\Pi_n^{(FR)}\) and \(\Pi_n^{(SR)}\), so that \(\Pi_n = \Pi_n^{(FR)} + \Pi_n^{(SR)}\). The next two lemmas investigate the point processes \(\Pi_n^{(FR)}\) and \(\Pi_n^{(SR)}\):

**Lemma 7.1 (PPP limit for the first round collision edges).** For every \(s \geq 0\), \(x, y \in \mathbb{R}\), \(j \in \{1, 2\}\), \(\epsilon > 0\) and \(\ell \in [N]\), as \(n \to \infty\),

\[
\mathbb{P}(\Pi_n^{(FR)}(\mathcal{I}_\ell^{(e)}) = 0 | \mathcal{F}_{t_{\ell-1}^{(e)}}) \xrightarrow{p} \exp\left\{-\frac{2\nu}{\mathbb{E}[D]} e^{2\alpha t_{\ell-1}^{(e)}} \frac{\Phi(x) \Phi(y) F_R(s) F_R(\epsilon)}{}\right\}.
\]
The number of half-edges $z \in \mathcal{AH}(\bar{t}_n + t^{(\epsilon)}_{k-1})$ that are found by the liquid having $I(z) = j, H(z) \leq k_n(t_n, x)$ and residual lifetime $R_{\bar{t}_n + t^{(\epsilon)}_{k-1}}(z) \in [0, \varepsilon)$ is equal to
\[
|\text{SWT}^{(j)}_{\leq k_n(t_n, x)}[\bar{t}_n + t^{(\epsilon)}_{k-1}, \bar{t}_n + t^{(\epsilon)}_{k-1} + \varepsilon]|.
\]
Fix such a half-edge $z$, and note that it is paired to $P_z$ that has $X^*_z = dv_{P_z} - 1$ sibling half-edges. For each of these half-edges, we test whether it is paired to a half-edge in $\mathcal{AH}(\bar{t}_n + t^{(\epsilon)}_{k-1})$ or not. Therefore, the total number of tests performed in the time interval $[t^{(\epsilon)}_{k-1}, t^{(\epsilon)}_{k})$ is equal to
\[
\sum_z X^*_z \mathbb{I}_{[\text{SWT}^{(j)}_{\leq k_n(t_n, x)}[\bar{t}_n + t^{(\epsilon)}_{k-1}, \bar{t}_n + t^{(\epsilon)}_{k-1} + \varepsilon]]}.
\]
By construction, we test whether these half-edges are paired to half-edges that are incident to the $\text{SWT}(j)$ or not. Each of these edges is paired to a half-edge $w \in \mathcal{AH}(\bar{t}_n + t^{(\epsilon)}_{k-1})$ with $I(w) = 3 - j$ (and thus creating a collision edge) and $H(w) \leq k_n(t_n, y)$ and $R_{\bar{t}_n + t^{(\epsilon)}_{k-1}}(w) \in [0, s)$ with probability equal to
\[
\left(1 - \frac{1}{\ell_n - o(n)}\right)|\text{SWT}^{(3-j)}_{\leq k_n(t_n, y)}[\bar{t}_n + t^{(\epsilon)}_{k-1}, \bar{t}_n + t^{(\epsilon)}_{k-1} + s]|.
\]
Therefore, for $\varepsilon > 0$, conditionally on $\mathcal{F}_{\bar{t}_n + t^{(\epsilon)}_{k-1}}$, the probability that none of the half-edges found in the time interval in between $[\bar{t}_n + t^{(\epsilon)}_{k-1}, \bar{t}_n + t^{(\epsilon)}_{k-1} + s]$ creates a collision edge is asymptotically equal to
\[
\prod_{v \in \text{SWT}^{(j)}_{\leq k_n(t_n, x)}[\bar{t}_n + t^{(\epsilon)}_{k-1}, \bar{t}_n + t^{(\epsilon)}_{k-1} + s]} \left(1 - \frac{1}{\ell_n - o(n)}\right)
\times |\text{SWT}^{(3-j)}_{\leq k_n(t_n, y)}[\bar{t}_n + t^{(\epsilon)}_{k-1}, \bar{t}_n + t^{(\epsilon)}_{k-1} + s]|^{X^*_v}
\xrightarrow{d} \exp\left\{-\frac{v}{\mathbb{E}[D]}e^{2\alpha t^{(\epsilon)}_{k-1}} \Phi(x) \Phi(y) F_R(s) F_R(\varepsilon) \right\},
\]
where we use (2.45), that $e^{2\alpha t_n} = n^{-1}$, and that $\ell_n = n\mathbb{E}[D_n]$ with $\mathbb{E}[D_n] \to \mathbb{E}[D]$. The factor 2 in (7.3) is caused by the two possibilities $j \in \{1, 2\}$. □

**Lemma 7.2 (A bound on the second round collision edges).** For $x, y \in \mathbb{R}$, $j \in \{1, 2\}, \varepsilon > 0$ and $\ell \in [N]$, as $n \to \infty$,
\[
P(\Pi_n^{(SR)}(I^{(\epsilon)}_{t^{(\epsilon)}_{k-1}}) \geq 1 | \mathcal{F}_{t^{(\epsilon)}_{k-1}}) = O_p(1) F_R(\varepsilon) F_\xi(\varepsilon).
\]

**Proof.** By analogous arguments as above, the expected number of second round collision edges is of order
\[
O_p(1)e^{2\alpha t} \Phi(x) \Phi(y) F_R(s) F_R(\varepsilon) F_\xi(\varepsilon),
\]
since one of the half-edges \( z \) that is found by the liquid in the time interval \( [\tilde{t}_n + t_{\ell-1}^{(\varepsilon)}, \tilde{t}_n + t_{\ell}^{(\varepsilon)}) \) needs to satisfy that one of the \( d_{VP} - 1 \) half-edges has weight at most \( \varepsilon \), and which, upon being found, needs to create a collision edge. \( \square \)

Now we are ready to complete the proof of Theorem 3.1. We use that

\[
\mathbb{P}(\Pi_n(I) = 0 \mid \mathcal{F}_{s_n}) = \mathbb{E} \left[ \prod_{\ell=1}^{N} \mathbb{P}(\Pi_n(I_{\ell}^{(\varepsilon)}) = 0 \mid \mathcal{F}_{t_{\ell-1}^{(\varepsilon)}}) \mid \mathcal{F}_{s_n} \right].
\]

We start with the upper bound, for which we use that

\[
\mathbb{P}(\Pi_n(I_{\ell}^{(\varepsilon)}) = 0 \mid \mathcal{F}_{t_{\ell-1}^{(\varepsilon)}}) \leq \mathbb{P}(\Pi_n^{(PR)}(I_{\ell}^{(\varepsilon)}) = 0 \mid \mathcal{F}_{t_{\ell-1}^{(\varepsilon)}})
\]

\[
\xrightarrow{\mathbb{P}} \exp \left\{ -\frac{2\nu}{\mathbb{E}[D]} e^{2\alpha t_{\ell-1}^{(\varepsilon)}} \Phi(x) \Phi(y) F_R(s) F_R(\varepsilon) \right\},
\]

by Lemma 7.1. We conclude that, for \( n \) large,

\[
\mathbb{P}(\Pi_n(I) = 0 \mid \mathcal{F}_{s_n}) \leq \mathbb{E} \left[ \prod_{\ell=1}^{N} \exp \left\{ -\frac{2\nu}{\mathbb{E}[D]} e^{2\alpha t_{\ell-1}^{(\varepsilon)}} \Phi(x) \Phi(y) F_R(s) F_R(\varepsilon) \right\} \mid \mathcal{F}_{s_n} \right]
\]

\[
= \exp \left\{ -\sum_{\ell=1}^{N} \frac{2\nu}{\mathbb{E}[D]} e^{2\alpha t_{\ell-1}^{(\varepsilon)}} \Phi(x) \Phi(y) F_R(s) F_R(\varepsilon) \right\}
\]

\[
\xrightarrow{\mathbb{P}} \exp \left\{ -\frac{2\nu}{\mathbb{E}[D]} f_R(0) \int_{a}^{b} e^{2\alpha t} \Phi(x) \Phi(y) F_R(s) \, dt \right\},
\]

since \( \lim_{\varepsilon \downarrow 0} F_R(\varepsilon)/\varepsilon = f_R(0) \), and the Riemann approximation

\[
\varepsilon \sum_{\ell=1}^{N} e^{2\alpha t_{\ell-1}^{(\varepsilon)}} \rightarrow \int_{a}^{b} e^{2\alpha t} \, dt.
\]

This proves the upper bound.

For the lower bound, we instead bound

\[
\mathbb{P}(\Pi_n(I) = 0 \mid \mathcal{F}_{s_n}) \geq \mathbb{E} \left[ \prod_{\ell=1}^{N} \mathbb{P}(\Pi_n^{(PR)}(I_{\ell}^{(\varepsilon)}) = 0 \mid \mathcal{F}_{t_{\ell-1}^{(\varepsilon)}} \bigg) \mathcal{F}_{s_n} \right]
\]

\[
- \mathbb{E} \left[ \left( \sum_{\ell=1}^{N} \mathbb{P}(\Pi_n^{(SR)}(I_{\ell}^{(\varepsilon)}) \geq 1 \mid \mathcal{F}_{t_{\ell-1}^{(\varepsilon)}} \bigg) \right) \wedge 1 \right] \mathcal{F}_{s_n} \right].
\]

The first term has already been dealt with, the second term is, by Lemma 7.2, bounded by

\[
\mathbb{E} \left[ \left( O_{\mathbb{P}}(1) \sum_{\ell=1}^{N} F_R(\varepsilon) F_R(\varepsilon) \right) \wedge 1 \right] \mathcal{F}_{s_n} = o_{\mathbb{P}}(1),
\]
as $\varepsilon \downarrow 0$, by dominated convergence, since $F_R(\varepsilon) = \varepsilon f_R(0)(1 + o(1))$ and $F_\xi(\varepsilon) = o(1)$.

Acknowledgments. S. Bhamidi would like to thank EURANDOM and TU/e for generous hospitality and travel support which lead to the genesis of this work. S. Bhamidi and R. van der Hofstad thank Cornell University for its hospitality during the Cornell Probability Summer School in July 2012, where this work was completed. G. Hooghiemstra would like to thank the Fundación Cultural Knecht-Drenth, for his stay in Callosa d’en Sarrià (Spain), where he worked on the revision. We thank an anonymous referee for detailed comments that helped us to significantly improve the presentation.

REFERENCES


S. Bhamidi
Department of Statistics
and Operations Research
University of North Carolina
304 Hanes Hall
Chapel Hill, North Carolina 27599
USA
E-mail: bhamidi@email.unc.edu

R. van der Hofstad
Department of Mathematics
and Computer Science
Eindhoven University of Technology
P.O. Box 513
5600 MB Eindhoven
The Netherlands
E-mail: rhofstad@win.tue.nl

G. Hooghiemstra
DIAM
Delft University of Technology
Mekelweg 4
2628 CD Delft
The Netherlands
E-mail: g.hooghiemstra@tudelft.nl