EURANDOM PREPRINT SERIES

2014-014

May, 2014

Universality for first passage percolation on sparse uniform and rank-1 random graphs

S. Bhamidi, R. van der Hofstad, G. Hooghiemstra
ISSN 1389-2355

UNIVERSALITY FOR FIRST PASSAGE PERCOLATION ON SPARSE UNIFORM AND RANK-1 RANDOM GRAPHS

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

Abstract. In [3], we considered first passage percolation on the configuration model equipped with general independent and identically distributed edge weights, where the common distribution function admits a density. Assuming that the degree distribution satisfies a uniform $X^2 \log X$-condition, we analyzed the asymptotic distribution for the minimal weight path between a pair of typical vertices, as well as the asymptotic distribution of the number of edges on this path. Given the interest in understanding such questions for various other random graph models, the aim of this paper is to show how these results extend to uniform random graphs with a given degree sequence and rank-one inhomogeneous random graphs.

1. Introduction and results

1.1. **Motivation.** First passage percolation (FPP) is an important topic in modern probability theory. Let us start with a decription of the basic model. Let $\mathcal{G}$ be a random graph on n vertices. Assign independent and identically distributed (i.i.d.) random *weights or lengths* to the edges of 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 function of the edge weights admits a density, 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. In applied settings, 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). See e.g., [1, 4–6] for results in the specific situation of *exponential* edge weights. The paper [3] provided the above mentioned results in the setting of general edge weights for the class of configuration models. The aim of this paper is to show how such results carry over to other models including the uniform random graph with a prescribeddegree sequence and the rank-1 inhomogeneous random graph models (see [9]).

1.2. **Notation and organization.** Throughout this paper, we make use of the following standard notation. We write $[n]$ for the set $\{1, 2, \ldots, n\}$. We let $\xrightarrow{a.s.}$ denote convergence almost surely, $\xrightarrow{d}$ denotes convergence in distribution, and $\xrightarrow{\mathbb{P}}$ convergence in probability. For a sequence of random variables $(X_n)_{n\geq 1}$, we write $X_n = O_{\mathbb{P}}(b_n)$, when $|X_n|/b_n$ is a tight sequence of random variables, and $X_n = o_{\mathbb{P}}(b_n)$ when $|X_n|/b_n \xrightarrow{\mathbb{P}} 0$, as $n \to \infty$. We write $D \sim F$ to denote

[1]Department of Statistics and Operations Research, 304 Hanes Hall, University of North Carolina, Chapel Hill, NC 27599

[2]Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands.

[3]DIAM, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands

E-mail addresses: bhamidi@email.unc.edu, rhofstad@win.tue.nl,g.hooghiemstra@tudelft.nl.

Date: May 31, 2014.

2010 *Mathematics Subject Classification.* Primary: 60C05, 05C80, 90B15.

Key words and phrases. central limit theorem, continuous-time branching processes, first passage percolation, hopcount, Malthusian rate of growth, random graphs.

that the random variable D has distribution function F. For non-negative 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 $\lim_{n\to\infty} f(n)/g(n) = 0$. Finally, we say that a sequence of events $(\mathcal{E}_n)_{n\geq 1}$ occurs *with high probability* (whp) when $\mathbb{P}(\mathcal{E}_n) \to 1$.

The remainder of the paper is organized as follows. In Section 2, we describe the main random graph models of interest and the assumptions imposed on these models. We also introduce the branching process preliminaries required to state our results in that section. In Section 3, we state our main results. In Section 4, we provide the proofs of these results by first reviewing known results from [3] and then showing how these extend to the related random graphs considered in this paper.

2. Models and branching process preliminaries

2.1. **Random graph models of interest.**

We start by describing the random graph models of interest and the assumptions on the model parameters.

2.1.1. **Uniform random graphs with a prescribed degree sequence.**

Let $\boldsymbol{d} = (d_1, d_2, \ldots, d_n)$ be a given *degree sequence*, i.e., a sequence of positive integers with total degree

$$\ell_n = \sum_{i\in[n]} d_i, \tag{2.1}$$

assumed to be even. We call a graph *simple* when it contains no self-loops nor multiple edges. Given the above degree sequence $\boldsymbol{d} = (d_1, d_2, \ldots, d_n)$, let $\mathrm{UG}_n(\boldsymbol{d})$ denote a random graph chosen uniformly at random amongst all simple graphs on vertex set $[n]$ having degree sequence $\boldsymbol{d}$.

Regularity of vertex degrees. Let us now describe our regularity assumptions on the degree sequence $\boldsymbol{d} = (d_1, d_2, \ldots, d_n)$, as $n \to \infty$. We denote the degree of a uniformly chosen vertex V in $[n]$ by $D_n = d_V$, so that

$$\mathbb{P}(D_n \leq x) = F_n(x) = \frac{1}{n}\sum_{j\in[n]} \mathbb{1}_{\{d_j \leq x\}}, \tag{2.2}$$

where $\mathbb{1}_A$ denotes the indicator of the set A. Write $\log(x)_+ = \log(x)$ for $x \geq 1$ and $\log(x)_+ = 0$ for $x \leq 1$.

Condition 2.1 (Regularity conditions for vertex degrees)**.**

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

$$\lim_{n\to\infty} F_n(x) = F(x), \tag{2.3}$$

for any continuity point x of F; i.e., $D_n \xrightarrow{d} D$.

(b) **Convergence of second moment.**

$$\lim_{n\to\infty} \mathbb{E}[D_n^2] = \lim_{n\to\infty} \frac{1}{n}\sum_{j\in[n]} d_j^2 = \mathbb{E}[D^2], \tag{2.4}$$

where $D_n \sim F_n$ and $D \sim F$, and where we further assume that

$$\nu = \mathbb{E}[D(D-1)]/\mathbb{E}[D] > 1. \tag{2.5}$$

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

$$\lim_{n\to\infty} \mathbb{E}[D_n^2 \log{(D_n/K_n)_+}] = \lim_{n\to\infty} \frac{1}{n}\sum_{j\in[n]} d_j^2 \log{(d_j/K_n)_+} = 0. \tag{2.6}$$

By Condition 2.1(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 $D_n^\star$, the size-biased version of D_n,

$$F_n^\star(x) = \mathbb{P}(D_n^\star \leq x) = \mathbb{E}[D_n \mathbb{1}_{\{D_n \leq x\}}]/\mathbb{E}[D_n], \quad x \in \mathbb{R}. \tag{2.7}$$

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 2.3 below, $D_n^\star$ is closely related to a *branching-process approximation* of neighborhoods of a uniformly chosen vertex, and thus Condition 2.1(c) implies that this branching process satisfies a uniform $X \log X$-condition. By uniform integrability, Condition 2.1(c) follows from the assumption that $\lim_{n\to\infty} \mathbb{E}[D_n^2 \log{(D_n)_+}] = \mathbb{E}[D^2 \log{(D)_+}]$.

Note that that Conditions 2.1(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 (2.4) and (2.6), are *random* and we assume that the convergence in (2.3), (2.4) and (2.6) to the respective (deterministic) right-hand sides holds *in probability.* Thus, in this case, we require that, with $\mathbb{E}_n[D_n^i] = \frac{1}{n}\sum_{j\in[n]} d_j^i$ (which is now a random variable) and for every $\varepsilon > 0$ and $i \in \{1, 2\}$,

$$\lim_{n\to\infty} \mathbb{P}(|F_n(x) - F(x)| \geq \varepsilon) = 0, \quad \forall x \in \mathbb{R}, \qquad \lim_{n\to\infty} \mathbb{P}(|\mathbb{E}_n[D_n^i] - \mathbb{E}[D^i]| \geq \varepsilon) = 0. \tag{2.8}$$

A similar condition replaces (2.6). Finally we note that (2.5) in Condition 2.1(c) is equivalent to the existence of a giant component in $\mathrm{UG}_n(\boldsymbol{d})$, see e.g. [22–24].

2.1.2. **Rank-1 inhomogeneous random graphs.** Fix a sequence of positive vertex weights $\boldsymbol{w} = (w_i)_{i\in[n]}$. Assume that there exists a random variable W with distribution function F_W on $\mathbb{R}^+$ such that

$$F_{n,w}(x) = \frac{1}{n} \sum_{i\in[n]} \mathbb{1}_{\{w_i \leq x\}} \to F_W(x), \tag{2.9}$$

for each continuity point of F_W. Let W_n denote the weight of a uniformly chosen vertex in $[n]$, i.e., $W_n = w_V$, where $V \in [n]$ is chosen uniformly at random.

Given these vertex weights, we construct a random graph by forming an edge between vertices i and j with probability

$$p_{ij} = 1 - \mathrm{e}^{-w_i w_j/\ell_n}, \tag{2.10}$$

where, with some abuse of notation, we write the sum of the vertex weights as

$$\ell_n = \sum_{i\in[n]} w_i, \tag{2.11}$$

and where we choose different edges independently. Let $\nu = \mathbb{E}[W^2]/\mathbb{E}[W]$. We always assume $\nu > 1$ as this is necessary and sufficient for the existence of a giant component (see, e.g., [9]). Note that letting $w_i = \lambda$, we immediately get the Erdős-Rényi random graph with edge connection probability $1 - \mathrm{e}^{-\lambda/n} \sim \lambda/n$. Thus, this model is a natural generalization of the classical random graph model. Related models are the *generalized random graph* introduced by Britton, Deijfen and Martin-Löf in [11], for which

$$p_{ij} = \frac{w_i w_j}{\ell_n + w_i w_j}, \tag{2.12}$$

and the *random graph with given prescribed expected degrees* or Chung-Lu model, where instead

$$p_{ij} = \min(w_i w_j/\ell_n, 1), \tag{2.13}$$

and which has been studied intensively by Chung and Lu (see [12–16]). By Janson [21], when $W_n \xrightarrow{d} W$ and $\mathbb{E}[W_n^2] \to \mathbb{E}[W^2]$, the three random graph models defined above are *asymptotically* *equivalent*, meaning that all events have equal asymptotic probabilities. By [9], with $N_k(n)$ denoting the number of vertices with degree k,

$$N_k(n)/n \xrightarrow{\mathbb{P}} \mathbb{E}\left[\mathrm{e}^{-W}\frac{W^k}{k!}\right]. \tag{2.14}$$

This proves that $D_n \xrightarrow{d} D$, where D_n is the degree of a vertex, chosen uniformly at random from $[n]$ and D has the mixed Poisson distribution given in (2.14), i.e., for this model

$$F(x) = \sum_{k\leq x}\mathbb{E}\left[\mathrm{e}^{-W}\frac{W^k}{k!}\right]. \tag{2.15}$$

We impose the following regularity condition on the weight sequence $\boldsymbol{w} = (w_i)_{i\in[n]}$, which is similar in spirit to the degree regularity in Condition 2.1:

Condition 2.2 (Regularity conditions for vertex weights)**.** *Assume that the weight W_n of a uniformly chosen vertex satisfies that $W_n \xrightarrow{d} W$, $\mathbb{E}[W_n^2] \to \mathbb{E}[W^2]$, with $0 < \mathbb{E}[W^2] < \infty$, and $\nu := \mathbb{E}[W^2]/\mathbb{E}[W] > 1$. Further assume that, for every $K_n \to \infty$,*

$$\mathbb{E}[W_n^2 \log{(W_n/K_n)_+}] = o(1).$$

2.2. **Functionals of interest: edge weights and shortest paths.** Once the (random) graph $\mathcal{G}$ has been constructed, we attach edge weight ξ_e to every edge e, where $(\xi_e)_{e\in\mathcal{E}_n}$ are i.i.d. continuous random variables with density $f_\xi\colon [0,\infty) \to [0,\infty)$ and corresponding distribution function F_ξ. Now pick two vertices U_1 and U_2 at random from $[n]$ and let Γ_{12} denote the set of all paths in $\mathcal{G}$ between these two vertices. The weight of a path $\pi \in \Gamma_{12}$ is equal to $\sum_{e\in\pi}\xi_e$. Let

$$L_n = \min_{\pi\in\Gamma_{12}}\sum_{e\in\pi}\xi_e, \tag{2.16}$$

denote the weight of the minimal weight path between the vertices U_1 and U_2, and let H_n denote number of edges or the *hopcount* of this path. If the two vertices are in different components of the graph, then we set $L_n, H_n = \infty$. The aim of this paper is to understand the asymptotics for these two functionals in the above random graph models as $n \to \infty$.

2.3. **Continuous-time branching processes.** In this section, we define the limiting continuous-time branching process (CTBP) that describes the neighborhood structure of FPP on the above defined sparse random graph models. Fix an integer valued non-negative random variable D. We will assume that

$$\mathbb{E}[D^2\log(D)_+] < \infty. \tag{2.17}$$

For example, when D is the weak limit of the degree sequence $\boldsymbol{d}$ as in Condition 2.1(a)-(c), this condition is met. Define the size-biased distribution $F^\star$ of the random variable $D \sim F$ by

$$F^\star(x) = \mathbb{E}[D\mathbb{1}_{\{D\leq x\}}]/\mathbb{E}[D], \quad x\in\mathbb{R}. \tag{2.18}$$

Again if D is the weak limit of D_{V_n} as in Condition 2.1, then the function $F^\star$ is the weak limit, as $n\to\infty$, of $F_n^\star$ in (2.7). Now let $(\mathsf{BP}^\star(t))_{t\geq 0}$ 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) Recall the edge weight random variable ξ with distribution F_ξ. In the associated branching process each individual v in the branching process lives for a random amount of time which has distribution F_ξ, i.e., the edge weight distribution, and then dies. At the time of death the individual gives birth to $D^\star - 1$ children, where $D^\star \sim F^\star$. Lifetimes and numbers of offspring across individuals are independent.

Note that in the above construction, if we let $X_v = D^\star - 1$ be the number of children of an individual, then the expected number of children satisfies

$$\mathbb{E}[X_v] = \mathbb{E}[D^\star - 1] = \nu > 1. \tag{2.19}$$

Further, using (2.17), for $D^\star \sim F^\star$,

$$\mathbb{E}[D^\star \log(D^\star)_+] < \infty. \tag{2.20}$$

The CTBP defined above is a *standard* Bellman-Harris process, with lifetime distribution F_ξ and offspring distribution $D^\star - 1$ [2,17,18]. The Malthusian parameter α of the branching process $\mathsf{BP}^\star$ is the unique solution of the equation

$$\hat{\mu}(\alpha) = \nu \int_0^\infty \mathrm{e}^{-\alpha t} dF_\xi(t) = 1. \tag{2.21}$$

Since $\nu > 1$, we obtain that $\alpha \in (0, \infty)$.

Standard theory (see e.g., [2,17,18]) implies that under our assumptions on the model, namely (2.19) and (2.20), there exists a random variable $\widetilde{\mathcal{W}}^\star$ such that

$$\mathrm{e}^{-\alpha t} |\mathsf{BP}^\star(t)| \xrightarrow{a.s.} \widetilde{\mathcal{W}}^\star. \tag{2.22}$$

Here the limiting random variable $\widetilde{\mathcal{W}}^\star$ satisfies $\widetilde{\mathcal{W}}^\star > 0$ a.s. on the event of non-extinction of the branching process and is zero otherwise.

By (2.21) we can define the *stable-age* distribution function $\bar{F}_\xi$ as

$$\bar{F}_\xi(x) = \nu \int_0^x \mathrm{e}^{-\alpha y} dF_\xi(y). \tag{2.23}$$

Let $\bar{\nu}$ be the mean and $\bar{\sigma}^2$ the variance of $\bar{F}_\xi$, i.e.,

$$\bar{\nu} = \nu \int_0^\infty x \mathrm{e}^{-\alpha x}\, dF_\xi(x), \qquad \bar{\sigma}^2 = \nu \int_0^\infty (x - \bar{\nu})^2 \mathrm{e}^{-\alpha x}\, dF_\xi(x). \tag{2.24}$$

Then $\bar{\nu}, \bar{\sigma}^2 \in (0, \infty)$, since $\alpha > 0$, and F_ξ is non-degenerate.

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. 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_ξ giving rise to a $D^\star - 1$ number of children where $D^\star \sim F^\star$, the size-biased distribution function $F^\star$ as in (2.18). Writing $|\mathsf{BP}(t)|$ for the number of alive individuals at time t, it is easy to see here as well that

$$\mathrm{e}^{-\alpha t} |\mathsf{BP}(t)| \xrightarrow{a.s.} \widetilde{\mathcal{W}}, \tag{2.25}$$

and that, conditionally on the event $\{D = k\}$,

$$\widetilde{\mathcal{W}} \stackrel{d}{=} \widetilde{\mathcal{W}}^{\star,(1)} + \cdots + \widetilde{\mathcal{W}}^{\star,(k)},$$

where $D \sim F$, and $\widetilde{\mathcal{W}}^{\star,(i)}$ are i.i.d. with the distribution of the limiting random variable in (2.22). Let $\mathcal{W}$ denote a random variable distributed as $\widetilde{\mathcal{W}}$ conditioned to be positive, i.e., for every $x \geq 0$,

$$\mathbb{P}(\mathcal{W} \leq x) = \mathbb{P}(\widetilde{\mathcal{W}} \leq x \mid \widetilde{\mathcal{W}} > 0). \tag{2.26}$$

To state our main results we will need the following constants and limit random variables defined in terms of the above branching process constructs. Define

$$\gamma = \frac{1}{\alpha \bar{\nu}}, \qquad \beta = \frac{\bar{\sigma}^2}{\bar{\nu}^3 \alpha}. \tag{2.27}$$

Define the random variable Q as

$$Q = \frac{1}{\alpha} \left(-\log \mathcal{W}^{(1)} - \log \mathcal{W}^{(2)} - \Lambda + c \right), \tag{2.28}$$

where $\mathbb{P}(\Lambda \leq x) = \mathrm{e}^{-\mathrm{e}^{-x}}$ (so that Λ is a standard Gumbel random variable), $\mathcal{W}^{(1)}, \mathcal{W}^{(2)}$ are two independent copies of the variable $\mathcal{W}$ in (2.26), also independent from Λ, and c is the constant

$$c = \log(\mathbb{E}[D](\nu - 1)^2 / (\nu \alpha \bar{\nu})). \tag{2.29}$$

3. Results

We are now in a position to state our results.

3.1. **Uniform random graphs.** We start with the uniform random graph $\mathrm{UG}_n(\boldsymbol{d})$, where the degree sequence $\boldsymbol{d}$ satisfies Condition 2.1. The associated branching process is as defined in Section 2.3 with D denoting the random variable with the limiting degree distribution and the associated size-biased distribution $F^\star$ as in (2.18). Recall the definition of the Malthusian rate of growth parameter defined in (2.21) and let α_n be the solution to (2.21) with ν replaced with

$$\nu_n = \mathbb{E}[D_n(D_n-1)]/\mathbb{E}[D_n]. \tag{3.1}$$

Clearly, $\nu_n \to \nu$ when Condition 2.1 holds, and further $|\alpha_n-\alpha| = O(|\nu_n-\nu|)$. We also define $\bar{F}_{n,\xi}$ to be the distribution function $\bar{F}_\xi$ in (2.23) with ν and α replaced by ν_n and α_n, and we let $\bar{\nu}_n$ and $\bar{\sigma}_n^2$ be the corresponding mean and variance. Finally, let $\gamma_n = 1/(\alpha_n\bar{\nu}_n)$.

Recall that H_n and L_n denote the number of edges and length respectively on the optimal path between two randomly selected vertices and consider the rescaling:

$$(\bar{H}_n, \bar{L}_n) := \Big(\frac{H_n-\gamma_n\log n}{\sqrt{\beta\log n}},\ L_n - \frac{1}{\alpha_n}\log n\Big). \tag{3.2}$$

Theorem 3.1 (FPP on uniform random graphs)**.** *Consider the uniform random graph* $\mathrm{UG}_n(\boldsymbol{d})$ *with degrees* $\boldsymbol{d}$ *satisfying Condition* 2.1, *and with i.i.d. edge weights distributed according to the continuous distribution* F_ξ*. Then the rescaled hopcount* $\bar{H}_n$ *and weight* $\bar{L}_n$ *of the optimal path as in* (3.2) *between two uniformly chosen vertices, conditioned on being connected, satisfy*

$$(\bar{H}_n, \bar{L}_n) \xrightarrow{d} (Z, Q), \tag{3.3}$$

as $n\to\infty$*, where* Z *and* Q *are independent and* Z *has a standard normal distribution, while* Q *has distribution as in* (2.28) *with respect to the associated branching process.*

Remark 3.2 (Asymptotic mean)**.** *We can replace in* (3.2) α_n *and* γ_n *by their limits* α *and* γ *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* $\nu_n = \nu + o(1/\sqrt{\log n})$ *and* $\nu_n = \nu + o(1/\log n)$*, respectively.*

3.2. **Rank-one inhomogeneous random graphs.** We now consider the case of the inhomogeneous random graph where the weight sequence $\boldsymbol{w}$ satisfies Condition 2.2. Let D_n denote the degree of a randomly selected vertex in the graph. By (2.14), $D_n \xrightarrow{d} D$ where

$$\mathbb{P}(D=k) = \mathbb{E}\left[\mathrm{e}^{-W}\frac{W^k}{k!}\right], \qquad k\geq 0,$$

and where W is the weak limit of the corresponding weight sequence. Let $D^\star$ denote the associated size-biased distribution as in (2.18) and construct the associated branching process as in Section 2.3 where the root has offspring D and subsequent generations have offspring $D^\star-1$ and where all individuals have lifetimes distributed as ξ. Recall that $\nu = \mathbb{E}[W^2]/\mathbb{E}[W] = \mathbb{E}[D^\star-1]$. Let α denote the corresponding Malthusian rate of growth as in (2.21) and recall the associated constants $\bar{\nu}$ and $\bar{\sigma}^2$ as in (2.24) and γ,β as in (2.27). Now write

$$\nu_n = \frac{\sum_{i\in[n]} w_i^2}{\sum_{i\in[n]} w_i}. \tag{3.4}$$

Let α_n denote the corresponding solution to (2.21) when using ν_n instead of ν and similarly $\bar{\nu}_n$ and $\bar{\sigma}_n^2$ and let $\gamma_n = 1/(\alpha_n\bar{\nu}_n)$. Define the rescaled hopcount and length between two randomly selected vertices as in (3.2). Then we have the following result:

Theorem 3.3 (FPP on rank-1 inhomogeneous random graphs)**.** *Consider the rank-1 inhomogeneous random graphs constructed using edge probabilities in* (2.10), (2.12) *or* (2.13) *and consider first passage percolation on the constructed graph. Assume that Condition* 2.2 *holds for the weight* *sequence $\boldsymbol{w}$. Then the optimal path between two randomly selected vertices conditioned on being connected satisfies*

$$(\bar{H}_n, \bar{L}_n) \xrightarrow{d} (Z, Q),$$

as $n \to \infty$, where Z and Q are independent and Z has a standard normal distribution, while Q has distribution as in (2.28) with respect to the associated branching process.

Theorem 3.3 can be understood as follows. By [11], the generalized random graph *conditioned on its degree sequence* is a uniform random graph with the same degree sequence. Therefore, Theorem 3.3 follows from Theorem 3.1 when the conditions on the random degrees given in Condition 2.1, hold in probability. This is proved in Section 4.

4. Proofs

We now start with the proofs. In Section 4.1 we describe the result from [3] for FPP on the configuration model with general continuous edge weights and give an intuitive idea of the proof in Section 4.2. Then in Section 4.3 we show how these results can be extended to the uniform random graph model and complete the proof of Theorem 3.1. Finally, we present the proof of Theorem 3.3, the inhomogeneous random graph case, in Section 4.4.

4.1. **The configuration model.** The configuration model (CM) is a random graph with vertex set $[n] := \{1, 2, \ldots, n\}$ and with prescribed degrees. Let $\boldsymbol{d} = (d_1, d_2, \ldots, d_n)$ be a given *degree sequence*, i.e., a sequence of n positive integers with total degree

$$\ell_n = \sum_{i \in [n]} d_i, \tag{4.1}$$

assumed to be even. The CM on n vertices with degree sequence $\boldsymbol{d}$ 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. [8, 20]). We denote the resulting graph on $[n]$ by $\mathrm{CM}_n(\boldsymbol{d})$, with corresponding edge set $\mathcal{E}_n$. Finally, we attach edge weight ξ_e to every edge e, where $(\xi_e)_{e \in \mathcal{E}_n}$ are i.i.d. continuous random variables with density $f_\xi \colon [0, \infty) \to [0, \infty)$ and corresponding distribution function F_ξ.

Let us now describe the result from [3]. Recall the notation for $\mathrm{UG}_n(\boldsymbol{d})$ in Section 3.1, all the constants and definitions are the same for $\mathrm{CM}_n(\boldsymbol{d})$. Recall that H_n and L_n denote the number of edges and length respectively of the optimal path between two randomly selected vertices and recall the rescaling $(\bar{H}_n, \bar{L}_n)$ in (3.2). The main result in [3] is the following theorem:

Theorem 4.1 (Joint convergence of hopcount and weight for configuration model)**.** *Theorem 3.1 applies verbatim to the configuration model $\mathrm{CM}_n(\boldsymbol{d})$ with degrees $\boldsymbol{d}$ satisfying Condition 2.1.*

4.2. **Idea of the proof.** Let us now briefly describe the intuitive idea behind the proof of Theorem 4.1. The key conceptual idea is that one can construct the CM simultaneously with the first passage percolation problem. More precisely, pick two vertices at random and think of liquid percolating through the network starting *simultaneously* from these two vertices at rate one, using the assigned edge-weight as the distance (length) between the vertices incident to that edge.

Let us start with what happens when we start percolating liquid through the network from a single vertex (ignoring issues such as self-loops etc. that might arise in the construction). Pick a vertex V_n uniformly at random from the set $[n]$, having d_{V_n} half-edges. Each of these half-edges is assigned a random edge-weight with distribution F_ξ. When the liquid reaches the end point of one of the d_{V_n} half-edges, we sample from the remaining half-edges to discover the neighbor corresponding to the exhausted half-edge. Let v_1 denote the first new vertex seen by the liquid after vertex V_n. Note that, conditional on vertex V_n,

$$\mathbb{P}(v_1 = v | V_n) = \frac{d_v \mathbb{1}_{\{v \neq V_n\}}}{\sum_{w \neq V_n} d_w}.$$

Thus v_1 approximately has the size-biased degree distribution $D^\star$, as $n \to \infty$. Further, one of the half-edges of this vertex has been used to attach to V_n and thus asymptotically the number of free half-edges has distribution $D^\star - 1$. It is not hard to convince one-self (although a rigorous proof turns out to be remarkably challenging, see [3, Sections 2-7]) that as the liquid percolates through the network, it can be approximated by the branching process $\mathsf{BP}_n^\star$, where the initial root has offspring distribution D_n, while each subsequent generation has offspring distribution $D_n^\star - 1$ where $D_n^\star$ is the size-biased version of D_n. In particular standard CTBP-theory [2,18,19] (although quantifying the dependence on n again turns out to be technically challenging, see e.g. [3, Proposition 2.2]) suggests that this branching process grows at rate $|\mathsf{BP}_n^\star(t)| \sim \tilde{\mathcal{W}}_n \exp(\alpha_n t)$, where $\tilde{\mathcal{W}}_n \xrightarrow{d} \tilde{\mathcal{W}}$ with $\tilde{\mathcal{W}}$ as in (2.25) as $n \to \infty$ and $\alpha_n \to \alpha$, where α_n is the Malthusian rate of growth for the n-dependent process.

Now let us consider the liquid starting from two uniformly chosen vertices simultaneously. Again locally the wetted edges look like two (independent) continuous-time branching processes as above. However, now we have an added construct called *collision edges*, which in fact generate the optimal path between these two vertices. For this, consider the wetted structures started from the two vertices simultaneously and let us call these $\{\mathsf{SWT}^{(i)}(t)\}_{t\geq 0}$, for $i = 1, 2$, and note that by the previous discussion, we can couple the above these $\{\mathsf{SWT}^{(i)}(t)\}_{t\geq 0}$ to two independent continuous-time branching processes $\mathsf{BP}_n^{(i)}$. However note that occasionally the liquid from one of the sources (to fix ideas say 1) samples a half-edge which has already been wetted by the liquid from the other vertex 2. When this happens, we say a *collision* has happened and note the time T when this collision occurs. In this case, the liquid from 2 has already traversed some distance on this half-edge and thus has some residual length R left before the liquid from 2 comes to the end of this half-edge. Note that this implies that there is a path between vertex 1 and vertex 2 of length $2T + R$. Thus, if one is able to keep track of the collision edges, in principal one can reconstruct the optimal path between vertex 1 and 2. This is performed by describing the arrival times, as well as various other properties of the vertices in the collision edges, as a point process. The main technical result in [3] shows that this point process of arrival times of collision edges converges to an appropriate (inhomogeneous) Poisson point process. This result immediately implies Theorem 4.1. As these details will not be used in this paper, we refrain from giving more details and refer the interested reader to [3] instead.

4.3. **Extension to the uniform random graph.** The proof of Theorem 3.1 follows rather directly from that of Theorem 4.1, by conditioning on simplicity. By [8] or [20], under Condition 2.1,

$$\lim_{n\to\infty} \mathbb{P}(\mathrm{CM}_n(\boldsymbol{d}) \text{ simple}) = \mathrm{e}^{-\nu/2-\nu^2/4}. \tag{4.2}$$

The proof of Theorem 4.1 reveals that in order to find the minimal weight path between vertices U_1 and U_2, we only need to investigate $O_{\mathbb{P}}(\sqrt{n}\log n)$ edges. Therefore, the event of simplicity of the CM will be mainly determined by the remaining $n - O_{\mathbb{P}}(\sqrt{n}\log n)$ *uninspected* edges, and is therefore asymptotically independent of (H_n, L_n). Together with the property that the CM conditioned on being simple is equal in distribution to a uniform random graph, this explains the content of Theorem 3.1.

Let $\mathrm{UG}_n(\boldsymbol{d})$ be a uniform random graph with degree sequence $\boldsymbol{d}$. By [7] (see also [8]), the law of $\mathrm{UG}_n(\boldsymbol{d})$ is the same as that of $\mathrm{CM}_n(\boldsymbol{d})$ conditioned on being simple, i.e., for every sequence of events $\mathcal{H}_n$ defined on graphs with vertex set $[n]$,

$$\mathbb{P}(\mathrm{UG}_n(\boldsymbol{d}) \in \mathcal{H}_n) = \mathbb{P}(\mathrm{CM}_n(\boldsymbol{d}) \in \mathcal{H}_n \mid \mathrm{CM}_n(\boldsymbol{d}) \text{ simple}) = \frac{\mathbb{P}(\mathrm{CM}_n(\boldsymbol{d}) \in \mathcal{H}_n, \mathrm{CM}_n(\boldsymbol{d}) \text{ simple})}{\mathbb{P}(\mathrm{CM}_n(\boldsymbol{d}) \text{ simple})}. \tag{4.3}$$

By (4.2), it suffices to investigate $\mathbb{P}(\mathrm{CM}_n(\boldsymbol{d}) \in \mathcal{H}_n, \mathrm{CM}_n(\boldsymbol{d}) \text{ simple})$. We take

$$\mathcal{H}_n = \Big\{\frac{H_n - \gamma_n \log n}{\sqrt{\beta \log n}} \leq x, L_n - \frac{1}{\alpha_n}\log n \leq y\Big\}, \tag{4.4}$$

Recall the point process convergence described in Section 4.2 using the exploration processes $\{\mathsf{SWT}^{(i)}(t)\}_{t\geq 0}$. The proof in [3] shows that, taking $\tilde{t}_n = t_n + B_n$, where $t_n = \log n/(2\alpha_n)$ and $B_n = \log\log\log n$, whp, we have found the minimal weight path before time $\tilde{t}_n$. The probability that we have found a self-loop or multiple edge at time $\tilde{t}_n$ is negligible, since, by that time we have found of order $\overline{m}_n = \sqrt{n}(\log n)^{1/4}$ vertices and paired of order $\overline{m}_n$ edges. Let $\tilde{d}_i(\tilde{t}_n)$ denote the number of unpaired half-edges incident to vertex i at time $\tilde{t}_n$. Since $\mathrm{CM}_n(\boldsymbol{d})$ is created by matching the half-edges uniformly at random, in order to create $\mathrm{CM}_n(\boldsymbol{d})$ after time $\tilde{t}_n$, we need to match the half-edges corresponding to $(\tilde{d}_i(\tilde{t}_n))_{i\in[n]}$. This corresponds to the configuration model on $[n]$ with degrees $(\tilde{d}_i(\tilde{t}_n))_{i\in[n]}$. Since we have found of order $\overline{m}_n = \sqrt{n}(\log n)^{1/4}$ vertices and paired of order $\overline{m}_n$ edges at time $\tilde{t}_n$, when $\boldsymbol{d}$ satisfies Condition 2.1, then so does $(\tilde{d}_i(\tilde{t}_n))_{i\in[n]}$ with the *same* limiting degree distribution D. As result, the probability that the configuration model on $[n]$ with degrees $(\tilde{d}_i(\tilde{t}_n))_{i\in[n]}$ is simple is asymptotically equal to $\mathrm{e}^{-\nu/2-\nu^2/4}(1+o(1))$, and we obtain that the event that $\mathrm{CM}_n(\boldsymbol{d})$ is simple is asymptotically independent of the event $\mathcal{H}_n$ in (4.4). Therefore, Theorem 3.1 follows from Theorem 4.1. ∎

4.4. **Extension to rank-1 inhomogenous random graphs.** In this section we prove Theorem 3.3. By Janson [21], when $W_n \xrightarrow{d} W$ and $\mathbb{E}[W_n^2] \to \mathbb{E}[W^2]$, the inhomogeneous random graphs with edge probabilities in (2.10), (2.12) or (2.13) are asymptotically equivalent, so it suffices to prove the claim for the *generalized random graph* for which $p_{ij} = w_iw_j/(\ell_n + w_iw_j)$. We denote the degree sequence in the generalized random graph by $\boldsymbol{d} = (d_i)_{i\in[n]}$, and note that $\boldsymbol{d}$ now is a *random* sequence. As explained in Section 2.1, conditionally on the degrees in the generalized random graph being equal to $\boldsymbol{d}$, the distribution of the resulting random graph is uniform over all random graphs with these degrees.

Denote by $\mathbb{P}_n, \mathbb{E}_n$ the conditional probability and expectation given $\boldsymbol{d}$. Then, Theorem 3.3 follows from Theorem 3.1, but with ν_n defined in (3.4), replaced by the random centering:

$$\frac{\sum_{i\in[n]} d_i(d_i-1)}{\sum_{i\in[n]} d_i},$$

precisely when Condition 2.1 holds in probability. By an argument similar to Remark 3.2, a full proof of Theorem 3.3 is hence given, when we show that:

(i) Condition 2.1 holds in probability and
(ii) for the degree sequence, as $n\to\infty$,

$$\left|\frac{\sum_{i\in[n]} d_i(d_i-1)}{\sum_{i\in[n]} d_i} - \nu_n\right| = o_{\mathbb{P}}((\log n)^{-1}), \tag{4.5}$$

where ν_n is as in (3.4).

Proof of Condition 2.1: We let $D_n = d_V$, where $V \in [n]$ is a uniformly chosen vertex. First, by (2.14), $D_n \xrightarrow{d} D$, where $D \sim F$ with F given in (2.15). Further,

$$\mathbb{E}_n[D_n] = \frac{1}{n}\sum_{i\in[n]} d_i, \qquad \mathbb{E}_n[D_n^2] = \frac{1}{n}\sum_{i\in[n]} d_i^2, \tag{4.6}$$

where $d_i = \sum_{j\in[n], j\neq i} I_{ij}$ and I_{ij} are independent Bernoulli variables with parameter $p_{ij} = w_iw_j/(\ell_n + w_iw_j)$. We will instead use the approximation $p_{ij} = w_iw_j/\ell_n$. Simple but cumbersome algebra extends this to previous case. Taking double expectations we get

$$\mathbb{E}[D_n] = \frac{1}{n}\sum_i w_i - \frac{1}{n\ell_n}\sum_i w_i^2 = \mathbb{E}[W] - O(n^{-1}). \tag{4.7}$$

Similarly, since by assumption $\max_{i\in[n]} w_i = o(\sqrt{n})$, it is easy to check that

$$\mathbb{E}[D_n^2] = \frac{1}{n}\sum_{i\in[n]} w_i + \frac{1}{n}\sum_{i\in[n]} w_i^2 - O(\frac{1}{\sqrt{n}}) = \mathbb{E}[W] + \mathbb{E}[W^2] - O(n^{-1/2}). \tag{4.8}$$

Further, in Proposition 4.2 below, we will show that $n^{-1}\sum_i d_i$ and $n^{-1}\sum_i d_i^2$ are strongly concentrated about their means. Thus, given Proposition 4.2, Condition 2.1(a-b) are satisfied.

In order to show Condition 2.1(c), we note that

$$\mathbb{E}_n[D_n^2\log(D_n/K_n)_+] = \frac{1}{n}\sum_{i\in[n]} d_i^2 \log{(d_i/K_n)_+}. \tag{4.9}$$

As before, $d_i = \sum_{j\in[n],j\neq i} I_{ij}$ and I_{ij} are independent Bernoulli variables with parameter $p_{ij} = w_iw_j/(\ell_n + w_iw_j)$. By standard Chernoff bounds [8], there exists a constant $a > 0$ such that, for every $\lambda \geq 2$,

$$\mathbb{P}(d_i \geq \lambda\mathbb{E}[d_i]) \leq \mathrm{e}^{-a\lambda\mathbb{E}[d_i]}. \tag{4.10}$$

Here,

$$\mathbb{E}[d_i] = \sum_{j\neq i} w_iw_j/(\ell_n + w_iw_j) \in w_i(\frac{\ell_n - \sqrt{n}}{\ell_n(1+o(1))}, 1) \subseteq w_i(1/2, 1), \tag{4.11}$$

where in the first inequality, we have used that $\max_{i\in[n]} w_i = o(\sqrt{n})$, and the second is true when n is sufficiently large. This suggests the split

$$\begin{aligned}\mathbb{E}_n[D_n^2\log(D_n/K_n)_+] \leq{}& \frac{1}{n}\sum_{i\in[n]} d_i^2 \log{(d_i/K_n)_+}\mathbb{1}_{\{d_i<2w_i\}} + \frac{1}{n}\sum_{i\in[n]} d_i^2 \log{(d_i/K_n)_+}\mathbb{1}_{\{w_i\leq 1\}} \\ &+ \frac{1}{n}\sum_{i\in[n]} \mathbb{1}_{\{d_i\geq 2w_i\}}\mathbb{1}_{\{w_i>1\}} d_i^2 \log{(d_i/K_n)_+}.\end{aligned} \tag{4.12}$$

The first term in (4.12) vanishes since

$$\frac{1}{n}\sum_{i\in[n]} d_i^2 \log{(d_i/K_n)_+}\mathbb{1}_{\{d_i<2w_i\}} \leq \frac{4}{n}\sum_{i\in[n]} w_i^2 \log{(2w_i/K_n)_+} = 4\mathbb{E}[W_n^2 \log{(2W_n/K_n)_+}], \tag{4.13}$$

and the fact that $\lim_{n\to\infty}\mathbb{E}[W_n^2\log{(W_n/K_n)_+}] = 0$. Using that for each $t > 0$,

$$d_i^2 \log{(d_i/K_n)_+} \leq \mathrm{e}^{(d_i-K_n)t} d_i^2 \log(d_i) \leq \mathrm{e}^{-tK_n} C_t \mathrm{e}^{2td_i}, \tag{4.14}$$

where $C_t = \max_{d\geq 1}\mathrm{e}^{-td}d^2\log d < \infty$, we can bound the expectation of second term in (4.12) as

$$\begin{aligned}&\frac{1}{n}\sum_{i\in[n]} d_i^2 \log{(d_i/K_n)_+}\mathbb{1}_{\{w_i\leq 1\}} \\ &\quad\leq \mathrm{e}^{-tK_n}C_t\frac{1}{n}\sum_{i\in[n]}\mathbb{E}[\mathrm{e}^{2td_i}]\mathbb{1}_{\{w_i\leq 1\}} = \mathrm{e}^{-tK_n}C_t\frac{1}{n}\sum_{i\in[n]}\prod_{j\neq i}(\mathrm{e}^{2t}p_{ij} + (1-p_{ij})\mathbb{1}_{\{w_i\leq 1\}} \\ &\quad\leq \mathrm{e}^{-tK_n}C_t\frac{1}{n}\sum_{i\in[n]}\mathbb{1}_{\{w_i\leq 1\}}\mathrm{e}^{\sum_{j\neq i}(\mathrm{e}^{2t}-1)p_{ij}} \leq \mathrm{e}^{-tK_n}C_t\frac{1}{n}\sum_{i\in[n]}\mathbb{1}_{\{w_i\leq 1\}}\mathrm{e}^{(\mathrm{e}^{2t}-1)w_i} \\ &\quad\leq \mathrm{e}^{-tK_n}C_t\mathrm{e}^{(\mathrm{e}^{2t}-1)} = o(1),\end{aligned} \tag{4.15}$$

where we use that $\mathrm{e}^{2t}p_{ij} + (1-p_{ij}) = (1 + (\mathrm{e}^{2t}-1)p_{ij}) \leq \mathrm{e}^{(\mathrm{e}^{2t}-1)p_{ij}}$ and $K_n \to \infty$.

The third term in (4.12) can be split as

$$\begin{aligned}&\frac{1}{n}\sum_{i\in[n]}\mathbb{1}_{\{w_i\geq 1, d_i\geq 2w_i\}} d_i^2\log{(d_i/K_n)_+} \\ &\quad\leq \frac{1}{n}\sum_{k=1}^{\infty}\sum_{i\in[n]}\mathbb{1}_{\{w_i\geq 1\}}\mathbb{1}_{\{d_i\in[2^kw_i, 2^{k+1}w_i)\}} d_i^2\log{(d_i/K_n)_+} \\ &\quad\leq \frac{1}{n}\sum_{i\in[n]}\mathbb{1}_{\{w_i\geq 1\}}\sum_{k=1}^{\infty}4^{k+1}\mathbb{1}_{\{d_i\geq 2^kw_i\}} w_i^2\log{(2^{k+1}w_i/K_n)_+}.\end{aligned} \tag{4.16}$$

By (4.10)-(4.11) with $\lambda = 2^k$, the mean of the above random variable is bounded by

$$\mathbb{E}\Big[\frac{1}{n}\sum_{i\in[n]} \mathbb{1}_{\{w_i\geq 1, d_i\geq 2w_i\}} d_i^2 \log{(d_i/K_n)_+}\Big] \tag{4.17}$$
$$\leq \frac{1}{n}\sum_{i\in[n]} \mathbb{1}_{\{w_i\geq 1\}} \sum_{k=1}^{\infty} 4^{k+1} \mathrm{e}^{-a2^{k-1}w_i} w_i^2 \log{(2^{k+1}w_i/K_n)_+}$$
$$\leq \frac{8}{K_n}\frac{1}{n}\sum_{i\in[n]} \mathbb{1}_{\{w_i\geq 1\}} \sum_{k=1}^{\infty} (w_i 2^k)^3 \mathrm{e}^{-a(w_i 2^k)/2} = \frac{8}{K_n}\frac{1}{n}\sum_{i\in[n]} \mathbb{1}_{\{w_i\geq 1\}} C,$$

where, in the second inequality, we use that $\log(x)_+ \leq x$, and in the last that, uniformly in $w \geq 1$,

$$\sum_{k=1}^{\infty} (w2^k)^3 \mathrm{e}^{-a(w2^k)/2} \leq C.$$

By Markov's inequality, this implies that $\mathbb{E}_n[D_n^2 \log(D_n/K_n)_+]$ converges to zero in probability. This completes checking that Condition 2.1 holds modulo the following result which simultaneously also proves (4.5):

Proposition 4.2 (Concentration of first two moment of degree sequence)**.** *As $n \to \infty$,*

$$\Big|\frac{1}{n}\sum_{i\in[n]} d_i - \frac{1}{n}\sum_{i\in[n]} w_i\Big| = o_{\mathbb{P}}((\log n)^{-1}), \quad \textit{and} \quad \Big|\frac{1}{n}\sum_{i\in[n]} d_i(d_i-1) - \frac{1}{n}\sum_{i\in[n]} w_i^2\Big| = o_{\mathbb{P}}((\log n)^{-1}).$$

Proof. First note, using the expressions for the expectations from (4.7) and (4.8), that it suffices to show that

$$(\log n)\left|Y_n^{(l)} - \mathbb{E}[Y_n^{(l)}]\right| \to 0, \tag{4.18}$$

as $n \to \infty$, for $l = 1, 2$, where $Y_n^{(l)} = n^{-1}\sum_i (d_i)_l$ and $(m)_l = m(m-1)\cdots(m-l+1)$. We first paraphrase the following standard concentration inequality for self-bounding functions (see [10, Chapter 6, Theorem 6.19 and 6.20]). The setting is as follows. Fix a measurable space χ and let $\mathbf{I} = (I_j)_{j\in[N]}$ be a collection of N independent χ-valued random variables. Let $f\colon \chi^N \to \mathbb{R}_+$ be a non-negative function. Call such a function (a, b) *strongly self-bounding* for two constants a, b, if there exist a collection of functions $(f_i)_{i\in N}$, with $f_i\colon \chi^{N-1} \to \mathbb{R}_+$ such that, for every point $\boldsymbol{x} \in \chi^N$,

$$0 \leq f(\boldsymbol{x}) - f_i(\boldsymbol{x}^i) \leq 1, \qquad \sum_{i=1}^{N} (f(\boldsymbol{x}) - f_i(\boldsymbol{x}^i)) \leq af(\boldsymbol{x}) + b. \tag{4.19}$$

Here $\boldsymbol{x}^i = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_N)$. Now consider the random variable $Y = f(\mathbf{I})$. Then one has the following concentration inequality:

Theorem 4.3. *For any $\lambda > 0$, and with $c = (3a-1)/6$,*

$$\mathbb{P}(Y \geq \mathbb{E}[Y] + \lambda) \leq \exp\Big(\frac{-\lambda^2}{2(a\mathbb{E}[Y] + b + a\lambda/2)}\Big), \mathbb{P}(Y \leq \mathbb{E}[Y] - \lambda) \leq \exp\Big(\frac{-\lambda^2}{2(a\mathbb{E}[Y] + b + c_-\lambda)}\Big),$$

where $c_- = \min(c, 0)$.

Let us now analyze the random variable $Y = \sum_i d_i/2$ (here the factor 2 is used to bring it to the above setting). We take: $\chi = \{0, 1\}$, $N = \binom{n}{2}$ and for $x \in \chi^N$,

$$f(x) = \sum_{i\in N} x_i. \tag{4.20}$$

To simplify notation, for any edge $e = \{i, j\}$, write $I_e \in \chi$ for the Bernoulli random variable signifying presence or absence of the edge and write $[n]_2$ for the collection of all possible edges. Note that our random variable $Y = f(\mathbf{I})$, satisfies:

$$Y = f(\mathbf{I}) = \sum_{e\in[n]_2} I_e = \sum_{j\in[n]} d_j/2,$$

since in any graph the sum of the degrees is equal to twice the number of edges. Now define $f_e(\mathbf{I}^e) := f(\mathbf{I}) - I_e$, namely the sum of degrees (divided by 2), when we leave out edge e. Then,

$$0 \le f(\mathbf{I}) - f_e(\mathbf{I}^e) \le 1, \qquad \sum_{e\in[n]_2} (f(\mathbf{I}) - f_e(\mathbf{I}^e)) = \sum_e I_e = f(\mathbf{I}). \tag{4.21}$$

So we can take $a = 1, b = 0$ and $c = 1/3$ and Theorem 4.3 implies that

$$\mathbb{P}(|\sum_i d_i - \mathbb{E}(\sum_i d_i)| > 2\lambda) \le 2\exp\Big(-\frac{\lambda^2}{2(\mathbb{E}(\sum_i d_i) + \lambda/2)}\Big).$$

Taking $2\lambda = n^{1-\alpha}$ this implies that

$$\Big|\frac{1}{n}\sum_i d_i - \mathbb{E}\Big[\frac{1}{n}\sum_i d_i\Big]\Big| = o_{\mathbb{P}}(n^{-\alpha}),$$

for any $\alpha < 1/2$. This proves (4.18) for $l = 1$.

Now let us consider $l = 2$. We observe that $d_i(d_i - 1) = \sum_{j,k} I_{ij}I_{ik}$, where the sum runs over all $j \neq k$ and where both j and k are distinct from i, Hence

$$\sum_{i\in[n]} d_i(d_i - 1) = \sum_{i,j,k} I_{ij}I_{ik}, \tag{4.22}$$

where the sum on the right-hand side is over all *distinct* $i, j, k \in [n]$. Note that in this representation the same term $I_{ij}I_{ik}$ appears twice, because the order can be interchanged. We define

$$f(\mathbf{I}) = \sum_{i,j,k} I_{ij}I_{ik}, \quad \text{and} \quad f_e(\mathbf{I}^e) = \sum_{i,j,k} I_{ij}I_{ik} - \sum_{i,j,k}^{e} I_{ij}I_{ik}, \tag{4.23}$$

where in the $\sum^e$ all terms with either $\{i, j\} = e$ or $\{i, k\} = e$ are left out. Obviously, writing $e = \{\underline{e}, \bar{e}\}$,

$$0 \le f(\mathbf{I}) - f_e(\mathbf{I}^e) \le (d_{\underline{e}} - 1) + (d_{\bar{e}} - 1), \tag{4.24}$$

so that

$$\sum_{e\in[n]_2} (f(\mathbf{I}) - f_e(\mathbf{I}^e)) \le \sum_{e\in[n]_2} (d_{\underline{e}} + d_{\bar{e}} - 2) = 2f(\mathbf{I}). \tag{4.25}$$

By Condition 2.1(c), $\max_i d_i = o_{\mathbb{P}}(\sqrt{n})$, and a simple truncation argument implies that it is enough to show that

$$\Big|\frac{1}{n}\sum_{i\in[n]} \bar{d}_i(\bar{d}_i - 1) - \mathbb{E}\Big[\frac{1}{n}\sum_{i\in[n]} \bar{d}_i(\bar{d}_i - 1)\Big]\Big| = o_{\mathbb{P}}\left((\log n)^{-1}\right)$$

where $\bar{d}_i = \max(d_i, \sqrt{n})$. Hence without loss of generality we may assume that $\max_i d_i \le \sqrt{n}$. Define the functions

$$g(\mathbf{I}) = \frac{f(\mathbf{I})}{2\sqrt{n}}, \qquad g_e(\mathbf{I}^e) = \frac{f_e(\mathbf{I}^e)}{2\sqrt{n}}.$$

It is now straightforward that g is a self-bounding function via the functions $(g_e)_{e\in[n]_2}$ (namely satisfies (4.19) with $a = 2, b = 0$). Using Theorem 4.3,

$$\mathbb{P}\Big(\Big|\sum_{i\in[n]} d_i(d_i - 1) - \mathbb{E}\Big[\sum_{i\in[n]} d_i(d_i - 1)\Big]\Big| > 2\lambda\sqrt{n}\Big) \le 2\exp\Big(\frac{-\lambda^2\sqrt{n}}{\mathbb{E}(2\sum_i d_i(d_i - 1)) + 2\lambda\sqrt{n}}\Big).$$

By taking $2\lambda\sqrt{n} = n^{1-\alpha}$, the above upper bound shows that for any $\alpha < 1/4$,

$$\Big|\frac{1}{n}\sum_{i\in[n]} d_i(d_i - 1) - \mathbb{E}\Big(\frac{1}{n}\sum_{i\in[n]} d_i(d_i - 1)\Big)\Big| = o_{\mathbb{P}}(n^{-\alpha}).$$

This completes the proof. . ∎

Acknowledgements. The research of SB is partially supported by NSF-DMS grant 1105581, NSF-DMS 1310002 and NSF SES 1357622 and would like to thank Eurandom and TU/e for travel support where part of this work was done. The work of RvdH was supported in part by Netherlands Organisation for Scientific Research (NWO).

References

[1] H. Amini, M. Draief, and M. Lelarge, *Flooding in weighted sparse random graphs*, SIAM Journal on Discrete Mathematics **27** (2013), no. 1, 1–26.
[2] K. B. Athreya and P. E. Ney, *Branching processes*, Dover Publications Inc., Mineola, NY, (2004). Reprint of the 1972 original [Springer, New York; MR0373040]. MR2047480
[3] S. Bhamidi, R. v. d. Hofstad, and G Hooghiemstra, *Universality for first passage percolation on sparse uniform and generalized random graphs*. Preprint (2014).
[4] S. Bhamidi, R. v. d. Hofstad, and G. Hooghiemstra, *Extreme value theory, Poisson-Dirichlet distributions, and first passage percolation on random networks*, Adv. in Appl. Probab. **42** (2010), no. 3, 706–738. MR2779556 (2012c:60226)
[5] S. Bhamidi, R. v. d. Hofstad, and G Hooghiemstra, *First passage percolation on random graphs with finite mean degrees*, Ann. Appl. Probab. **20** (2010), no. 5, 1907–1965. MR2724425 (2011m:60289)
[6] S. Bhamidi, R. v. d. Hofstad, and G Hooghiemstra, *First passage percolation on the Erdős-Rényi random graph*, Combin. Probab. Comput. **20** (2011), no. 5, 683–707. MR2825584
[7] B. Bollobás, *A probabilistic proof of an asymptotic formula for the number of labelled regular graphs*, European J. Combin. **1** ((1980)), no. 4, 311–316. MR595929 (82i:05045)
[8] B. Bollobás, *Random graphs*, Second edition, Cambridge Studies in Advanced Mathematics, vol. **73**, Cambridge University Press, Cambridge, (2001). MR1864966 (2002j:05132)
[9] B. Bollobás, S. Janson, and O. Riordan, *The phase transition in inhomogeneous random graphs*, Random Structures Algorithms **31** (2007), no. 1, 3–122. MR2337396
[10] S. Boucheron, G. Lugosi, and P. Massart, *Concentration inequalities: A nonasymptotic theory of independence*, Oxford University Press, 2013.
[11] T. Britton, M. Deijfen, and A. Martin-Löf, *Generating simple random graphs with prescribed degree distribution*, J. Stat. Phys. **124** ((2006)), no. 6, 1377–1397. MR2266448
[12] F. Chung and L. Lu, *The average distances in random graphs with given expected degrees*, Proc. Natl. Acad. Sci. USA **99** ((2002)), no. 25, 15879–15882 (electronic). MR1944974 (2003k:05124)
[13] F. Chung and L. Lu, *Connected components in random graphs with given expected degree sequences*, Ann. Comb. **6** ((2002)), no. 2, 125–145. MR1955514 (2003k:05123)
[14] F. Chung and L. Lu, *The average distance in a random graph with given expected degrees*, Internet Math. **1** ((2003)), no. 1, 91–113. MR2076728 (2005e:05122)
[15] F. Chung and L. Lu, *Complex graphs and networks*, CBMS Regional Conference Series in Mathematics, vol. **107**, Published for the Conference Board of the Mathematical Sciences, Washington, DC, (2006). MR2248695 (2007i:05169)
[16] F. Chung and L. Lu, *The volume of the giant component of a random graph with given expected degrees*, SIAM J. Discrete Math. **20** ((2006)), 395–411.
[17] T. E. Harris, *The Theory of Branching processes*, Springer-Verlag, Berlin, (1963).
[18] P. Jagers, *Branching Processes with Biological Applications*, John Wiley & Sons, (1975).
[19] P. Jagers and O. Nerman, *The growth and composition of branching populations*, Advances in Applied Probability **16** (1984), no. 2, 221–259.
[20] S. Janson, *The probability that a random multigraph is simple*, Combinatorics, Probability and Computing **18** (2009), no. 1-2, 205–225.
[21] S. Janson, *Asymptotic equivalence and contiguity of some random graphs*, Random Structures Algorithms **36** (2010), no. 1, 26–45. MR2591045
[22] S. Janson and M. Luczak, *A new approach to the giant component problem*, Random Structures Algorithms **34** (2009), no. 2, 197–216. MR2490288 (2010d:05140)
[23] M. Molloy and B. Reed, *A critical point for random graphs with a given degree sequence*, Random Structures Algorithms **6** (1995), no. 2-3, 161–179. MR1370952 (97a:05191)
[24] M. Molloy and B. Reed, *The size of the giant component of a random graph with a given degree sequence*, Combinatorics Probability and Computing **7** (1998), no. 3, 295–305. MR1664335 (2000c:05130)