

# EXTREME VALUE THEORY, POISSON–DIRICHLET DISTRIBUTIONS, AND FIRST PASSAGE PERCOLATION ON RANDOM NETWORKS

SHANKAR BHAMIDI,\* *University of North Carolina*

REMCO VAN DER HOFSTAD,\*\* *Eindhoven University of Technology*

GERARD HOOGHIEMSTRA,\*\*\* *Delft University of Technology*

## Abstract

We study first passage percolation (FPP) on the configuration model (CM) having power-law degrees with exponent  $\tau \in [1, 2)$  and exponential edge weights. We derive the distributional limit of the minimal weight of a path between typical vertices in the network and the number of edges on the minimal-weight path, both of which can be computed in terms of the Poisson–Dirichlet distribution. We explicitly describe these limits via construction of infinite limiting objects describing the FPP problem in the densely connected core of the network. We consider two separate cases, the *original CM*, in which each edge, regardless of its multiplicity, receives an independent exponential weight, and the *erased CM*, for which there is an independent exponential weight between any pair of direct neighbors. While the results are qualitatively similar, surprisingly, the limiting random variables are quite different. Our results imply that the flow carrying properties of the network are markedly different from either the mean-field setting or the locally tree-like setting, which occurs as  $\tau > 2$ , and for which the hopcount between typical vertices scales as  $\log n$ . In our setting the hopcount is tight and has an explicit limiting distribution, showing that information can be transferred remarkably quickly between different vertices in the network. This efficiency has a down side in that such networks are remarkably fragile to directed attacks. These results continue a general program by the authors to obtain a complete picture of how random disorder changes the inherent geometry of various random network models; see Aldous and Bhamidi (2010), Bhamidi (2008), and Bhamidi, van der Hofstad and Hooghiemstra (2009).

*Keywords:* Configuration model; random graph; first passage percolation; hopcount; extreme value theory; Poisson–Dirichlet distribution; scale-free network

2010 Mathematics Subject Classification: Primary 60C05; 05C80; 90B15

## 1. Introduction

First passage percolation (FPP) was introduced in [11] to model the flow of fluid through random media. This model has evolved into one of the fundamental problems studied in modern probability theory, not just for its own sake but also due to the fact that it plays a crucial role in the analysis of many other problems in statistical physics, in areas such as the contact process,

---

Received 1 June 2009; revision received 5 January 2010.

\* Postal address: Department of Statistics and Operations Research, University of North Carolina, #304 Hanes Hall, Chapel Hill, NC 27599, USA.

\*\* Postal address: Department of Mathematics and Computer Science, Eindhoven University of Technology, PO Box 513, 5600 MB Eindhoven, The Netherlands. Email address: rhofstad@win.tue.nl

\*\*\* Postal address: DIAM, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands. Email address: g.hooghiemstra@ewi.tudelft.nl

the voter model, electrical resistance problems, and in fundamental stochastic models from evolutionary biology; see, e.g. [8]. The basic model for FPP on a (random) graph is defined as follows. We start with a connected graph on  $n$  vertices. Each edge is given some random weight, assumed to be nonnegative, and independent and identically distributed (i.i.d.) across the edges. The weight on an edge has the interpretation of the *length* or *cost* of traversing this edge. Fixing two vertices in the network, we are then interested in the length or weight of the minimal-weight path between these two vertices and the asymptotics of these statistics as the size of the network tends to  $\infty$ .

Most of the classical theorems about FPP deal with the  $d$ -dimensional integer lattice, where the connected network is the  $[-r, r]^d$  box in the integer lattice and one is interested in asymptotics of various quantities as  $n = (2r + 1)^d \rightarrow \infty$ . In this context, probabilists are interested in proving shape theorems, namely, for fixed distance  $t$ , showing that  $\mathcal{C}_t/t$  converges to a deterministic limiting set as  $t \rightarrow \infty$ , where  $\mathcal{C}_t$  is the cluster of all vertices within distance  $t$  from the origin. See, e.g. [12] for a survey of results in this context.

In the modern context such problems have taken on a new significance. The last few years have witnessed an explosion in the amount of empirical data on networks, including data transmission networks such as the Internet, biochemical networks such as gene regulatory networks, spatial flow routing networks such as power transmission networks, and transportation networks such as road and rail networks. This has stimulated an intense cross-disciplinary effort in formulating network models to understand the structure and evolution of such real-world networks. Understanding FPP in the context of these random models seems to be of paramount importance, with the minimal weight between typical vertices representing the cost of transporting flow between these vertices, while the hopcount, which is defined to be the number of edges on the minimal-weight path between two typical vertices, represents the amount of time it takes for flow to be transported between these vertices.

The *configuration model* (CM) was introduced in [4]. Molloy and Reed [18] were the first to use specified degree sequences. This model has arisen in many applied contexts, ranging from combinatorics, computer science, statistical physics, and epidemiology, and seems to be one of the most widely used models in the modern networking community. We will defer a formal definition of this model to Section 2 and discuss related work in Section 4.

We will consider FPP on the CM where the exponent  $\tau$  of the degree distribution satisfies  $\tau \in [1, 2)$  and each edge is given a random exponential edge weight. FPP for the case  $\tau > 2$  was analyzed in [6], where the hopcount seems to exhibit a remarkably universal behavior. More precisely, in this regime, the hopcount always scales as  $\log n$  and central limit theorems (CLTs) with matching asymptotic means and variances hold. While these graphs are sparse and locally tree-like, what is remarkable is that the same fact also holds in the case of the most well-connected graph, namely the complete graph, for which the hopcount satisfies a CLT similar to the model with degree exponent  $\tau > 2$ , with asymptotic mean and variance equal to  $\log n$  as  $n \rightarrow \infty$ . See, e.g. [5], [13], and the references therein.

When the degree exponent  $\tau$  is in the interval  $[1, 2)$ , we will find that CLTs do *not* hold, and that the hopcount is tight due to the remarkable shape of such networks, which may be thought of as a collection of interconnected star networks, the centers of the stars corresponding to the vertices with highest degrees. We will consider two models of the network topology, one in which we look at the original CM and a second more realistic model called the *erased* model in which we will delete all self-loops and merge all multiple edges from the original CM. In the resulting graph, each edge receives an independent exponential weight with rate 1. Thus, for the erased CM, the direct weight between two vertices connected by an edge is an exponential

random variable with rate 1, while, for the original CM, it is also exponential, but with rate equal to the number of edges between the pair of vertices. When  $\tau > 2$ , there is no essential difference between the original and the erased CM; see, e.g. [6].

In both cases, we will see that the hopcount is tight and that a limit distribution exists. More surprisingly, in the erased CM, this limiting distribution puts mass only on the *even* integers. We also exhibit a nice constructive picture of how this arises, which uses the powerful machinery of Poisson–Dirichlet distributions. We further find the distributional limit of the weight of the minimal-weight path joining two typical vertices.

Since the hopcount remains tight, this model is remarkably efficient in transporting or routing flow between vertices in the network. However, a downside of this property of the network is its extreme fragility with respect to directed attacks on the network. More precisely, we will show that there exists a simple algorithm deleting a bounded number of vertices such that the chance of disconnecting any two typical vertices is close to 1 as  $n \rightarrow \infty$ . At the same time we will also show that these networks are relatively stable against random attacks.

This paper is organized as follows. In Section 2 we introduce the model and some notation. In Section 3 we state our main results. In Section 4 we describe connections to the literature and discuss our results. In Section 5 we give the proof in the original CM, and in Section 6 we prove the results in the erased CM.

## 2. Notation and definitions

In this section we describe the construction of the random graph model and recall some limiting results for i.i.d. random variables with infinite mean. We use the notation  $f(n) = O(g(n))$  as  $n \rightarrow \infty$  if  $|f(n)| \leq Cg(n)$  for some constant  $C$ , and  $f(n) = o(g(n))$  as  $n \rightarrow \infty$  if  $|f(n)|/g(n) \rightarrow 0$ . For two sequences of random variables  $X_n$  and  $Y_n$ , we write  $X_n = O_P(Y_n)$  as  $n \rightarrow \infty$  when  $\{X_n/Y_n\}_{n \geq 1}$  is a tight sequence of random variables. Furthermore, we write  $X_n = \Theta_P(Y_n)$  if  $X_n = O_P(Y_n)$  and  $Y_n = O_P(X_n)$ , and  $X_n = o_P(Y_n)$  when  $|X_n|/Y_n$  goes to 0 in probability ( $|X_n|/Y_n \xrightarrow{P} 0$ ). Equality in distribution is denoted by the symbol ' $\sim$ '. Throughout this paper, for a sequence of events  $\{F_n\}_{n \geq 1}$ , we say that  $F_n$  occurs with high probability (w.h.p.) if  $\lim_{n \rightarrow \infty} P(F_n) = 1$ .

*Graphs.* We will typically be working with random graphs on  $n$  vertices, which have a giant component consisting of  $n - o(n)$  vertices. Edges are given a random edge weight (sometimes alternatively referred to as cost) which in this study will always be assumed to be independent, exponentially distributed random variables with mean 1. Choose two vertices uniformly at random in the network, let  $W_n$  be the random variable denoting the total weight of the minimum weight path between the two vertices, and let  $H_n$  be the number of edges on this path, namely the *hopcount*.

*Construction of the configuration model.* We are interested in constructing a random graph on  $n$  vertices. Given a *degree sequence*, namely a sequence of  $n$  positive integers  $\mathbf{D} = (D_1, D_2, \dots, D_n)$  with the total degree

$$L_n = \sum_{i=1}^n D_i$$

assumed to be even, the CM on  $n$  vertices with degree sequence  $\mathbf{D}$  is constructed as follows.

Start with  $n$  vertices and  $D_j$  stubs adjacent to vertex  $j$ . The graph is constructed by pairing up each stub to some other stub to form edges. Number the stubs from 1 to  $L_n$  in some arbitrary

order. Then, at each step, two stubs (that are not yet paired) are chosen uniformly at random among all the *free* stubs and are paired to form a single edge in the graph. These stubs are no longer free and removed from the list of free stubs. We continue with this procedure of choosing and pairing two stubs until all the stubs are paired.

*Degree distribution.* The above denotes the construction of the CM when the degree distribution is given and the total degree is even. Here we specify how we construct the actual degree sequence  $\mathbf{D}$ . We will assume that each of the random variables  $D_1, D_2, \dots, D_n$  are i.i.d. random variables with distribution function  $F$ . (If the sum of stubs  $L_n$  is not even then we use the degree sequence with  $D_n$  replaced with  $D_n + 1$ . This will not effect our calculations.)

We will assume that the degree distribution  $F$  with atoms  $f_1, f_2, \dots$  satisfies

$$1 - F(x) = x^{-(\tau-1)}L(x) \tag{2.1}$$

for some function  $x \mapsto L(x)$  that is slowly varying at  $\infty$ . Here, the parameter  $\tau$ , which we refer to as the *degree exponent*, is assumed to be in the interval  $[1, 2)$ , so that  $E[D_i] = \infty$ . In some cases, we will make stronger assumptions than (2.1). As described above, we will deal with two variants of the FPP model.

*Original model.* We assign to each edge a random and i.i.d. exponential edge weight with rate 1. In what follows, the weighted random graph so generated will be referred to as the *original* model and we will denote the random network so obtained by  $\mathcal{G}_n^{or}$ .

*Erased model.* This model is constructed as follows. Generate a CM as before and then erase all self-loops and merge all multiple edges between any two vertices into a single edge between these vertices. *After* this, put independent exponential weights with rate 1 on the (remaining) edges. While the graph distances are not affected by the erasure, we will see that the hopcount has a different limiting distribution. We denote the random network on  $n$  vertices so obtained by  $\mathcal{G}_n^{er}$ .

### 2.1. Poisson–Dirichlet distribution

Before describing our results, we need to make a brief detour into extreme value theory for heavy-tailed random variables. As in [23], where the graph distances in the CM with  $\tau \in [1, 2)$  are studied, the relative sizes of the order statistics of the degrees play a crucial role in the proof. In order to describe the limiting behavior of the order statistics, we need some definitions.

We define a (random) probability distribution  $P = \{P_i\}_{i \geq 1}$  as follows. Let  $\{E_i\}_{i=1}^\infty$  be i.i.d. exponential random variables with rate 1, and define  $\Gamma_i = \sum_{j=1}^i E_j$ . Let  $\{D_i\}_{i=1}^\infty$  be an i.i.d. sequence of random variables with distribution function  $F$  in (2.1), and let  $D_{(n:n)} \geq D_{(n-1:n)} \geq \dots \geq D_{(1:n)}$  be the order statistics of  $\{D_i\}_{i=1}^n$ . In the rest of this paper we label vertices according to their degree, so that vertex 1 has maximal degree, etc.

We recall from [23, Lemma 2.1] that, by (2.1), there exists a sequence  $u_n$ , with  $u_n = n^{1/(\tau-1)}l(n)$ , where  $l$  is slowly varying, such that

$$u_n^{-1}(L_n, \{D_{(n+1-i:n)}\}_{i=1}^\infty) \xrightarrow{D} \left( \sum_{j=1}^\infty \Gamma_j^{-1/(\tau-1)}, \{\Gamma_i^{-1/(\tau-1)}\}_{i=1}^\infty \right), \tag{2.2}$$

where ‘ $\xrightarrow{D}$ ’ denotes convergence in distribution. Here  $u_n$  can be defined as

$$u_n = \sup\{u : 1 - F(u) \geq 1/n\}. \tag{2.3}$$

We abbreviate  $\xi_i = \Gamma_i^{-1/(\tau-1)}$  and  $\eta = \sum_{j=1}^\infty \xi_j$ , and let

$$P_i = \xi_i/\eta, \quad i \geq 1, \tag{2.4}$$

so that  $P = \{P_i\}_{i \geq 1}$  is a random probability distribution. The sequence  $\{P_i\}_{i \geq 1}$  is called the *Poisson–Dirichlet distribution* (see, e.g. [20]). A lot is known about this random probability distribution  $P$ . For example, Pitman and Yor [20, Equation (6)] proved that, for any  $f : [0, 1] \rightarrow \mathbb{R}$ , and with  $\alpha = \tau - 1 \in (0, 1)$ ,

$$\mathbb{E} \left[ \sum_{i=1}^\infty f(P_i) \right] = \frac{1}{\Gamma(\alpha)\Gamma(1-\alpha)} \int_0^1 f(u)u^{-\alpha-1}(1-u)^{\alpha-1} du, \tag{2.5}$$

where  $\Gamma$  denotes the gamma function [1, Section 6.1.1]. This in particular implies that

$$\mathbb{E} \left[ \sum_{i=1}^\infty P_i^2 \right] = \frac{\Gamma(\alpha)\Gamma(2-\alpha)}{\Gamma(\alpha)\Gamma(1-\alpha)} = 1 - \alpha = 2 - \tau. \tag{2.6}$$

### 3. Results

In this section we state the main results of the paper, distinguishing between the original CM and the erased CM.

#### 3.1. Analysis of shortest-weight paths for the original CM

Before describing the results, we need to construct a limiting infinite object  $\mathcal{K}_\infty^{\text{or}}$  in terms of the Poisson–Dirichlet distribution  $\{P_i\}_{i \geq 1}$  given in (2.4), and the sequence of random variables  $\xi_i$  and their sum  $\eta$  which arise in the representation of this distribution. This will be an infinite graph with vertex set  $\mathbb{Z}^+ = \{1, 2, \dots\}$  and edge set consisting of all unordered pairs  $\{i, j\}$ , with  $i, j \in \mathbb{Z}^+$ . Conditioned on the sequence  $\{\xi_i\}_{i \geq 1}$ , the edge weights are independent, exponentially distributed random variables, with rate  $\xi_i \xi_j / \eta$  on edge  $\{i, j\}$ .

Let  $W_{ij}^{\text{or}}$  and  $H_{ij}^{\text{or}}$  denote the weight and number of edges of the minimal-weight path in  $\mathcal{K}_\infty^{\text{or}}$  between the vertices  $i, j \in \mathbb{Z}^+$ . Our results will show that, in fact, the FPP problem on  $\mathcal{K}_\infty^{\text{or}}$  is well defined (see Proposition 5.1). Conditioned on  $\{P_i\}_{i \geq 1}$ , we define  $I^{\text{or}}$  and  $J^{\text{or}}$  as two vertices chosen independently at random from the vertex set  $\mathbb{Z}^+$  with probability mass function  $\{P_i\}_{i \geq 1}$ , and as above,  $W_{I^{\text{or}}J^{\text{or}}}$  and  $H_{I^{\text{or}}J^{\text{or}}}$  denote the weight and hopcount of the minimal-weight path between these two randomly selected vertices in  $\mathcal{K}_\infty^{\text{or}}$ . Finally, recall that we used  $\mathcal{G}_n^{\text{or}}$  to represent the original random network on  $n$  vertices with exponential edge weights as constructed in Section 2, where we have not deleted self loops or multiple edges. We are now in a position to describe our limiting results for the original CM.

**Theorem 3.1.** (FPP asymptotics for the original CM.) *Consider the random network  $\mathcal{G}_n^{\text{or}}$ , with the degree distribution  $F$  satisfying (2.1) for some  $\tau \in (1, 2)$ .*

- (a) *Let  $W_n^{\text{or}}$  be the weight of the minimal-weight path between two uniformly chosen vertices in the network. Then*

$$W_n^{\text{or}} \xrightarrow{D} V_1^{\text{or}} + V_2^{\text{or}}, \tag{3.1}$$

where  $V_i^{\text{or}}$ ,  $i = 1, 2$ , are independent random variables with  $V_i^{\text{or}} \sim E_i/D_i$ , where  $E_i$  is exponential with rate 1 and the  $D_i$  are i.i.d. with distribution  $F$ , distributed independently of  $E_i$ . More precisely, as  $n \rightarrow \infty$ ,

$$u_n(W_n^{\text{or}} - (V_1^{\text{or}} + V_2^{\text{or}})) \xrightarrow{D} W_{I^{\text{or}}J^{\text{or}}}^{\text{or}},$$

where  $u_n$  is defined by (2.3).

(b) Let  $H_n^{or}$  be the number of edges in the minimal-weight path between two uniformly chosen vertices in the network. Then

$$H_n^{or} \xrightarrow{D} 2 + H_{J^{or}}^{or}.$$

Writing  $\pi_k = P(H_{J^{or}}^{or} = k - 2)$ , we have  $\pi_k > 0$  for each  $k \geq 2$ , when  $\tau \in (1, 2)$ . The probability distribution  $\pi$  depends only on  $\tau$ , and not on any other detail of the degree distribution  $F$ . Moreover,

$$\pi_2 = 2 - \tau. \tag{3.2}$$

Theorem 3.1 implies that, for  $\tau \in (1, 2)$ , the hopcount is tight, as is the case for the typical graph distance obtained by taking the weights to be equal to 1 almost surely (a.s.); see [23]. However, while for unit edge weights and  $\tau \in (1, 2)$ , the limiting hopcount is at most 3, for i.i.d. exponential weights, the limiting hopcount can take all integer values greater than or equal to 2.

### 3.2. Analysis of shortest-weight paths for the erased CM

The results in the erased CM hold under a more restrictive condition on the degree distribution  $F$ . More precisely, we assume that there exists a constant  $0 < c < \infty$  such that

$$1 - F(x) = cx^{-(\tau-1)}(1 + o(1)), \quad x \rightarrow \infty, \tag{3.3}$$

and we will often make use of the upper bound  $1 - F(x) \leq c_2x^{-(\tau-1)}$ , valid for all  $x \geq 0$  and some constant  $c_2 > 0$ .

Before we can describe our limit result for the erased CM, we need an explicit construction of a limiting infinite network  $\mathcal{K}_\infty^{er}$  using the Poisson–Dirichlet distribution described in (2.4). Fix a realization  $\{P_i\}_{i \geq 1}$ . Conditional on this sequence, let  $f(P_i, P_j)$  be the probability,

$$f(P_i, P_j) = P(\mathcal{E}_{ij}),$$

of the following event  $\mathcal{E}_{ij}$ .

- Generate a random variable  $D \sim F$ , where  $F$  is the degree distribution. Conduct  $D$  independent multinomial trials where we select cell  $i$  with probability  $P_i$  at each stage. Then  $\mathcal{E}_{ij}$  is the event that both cells  $i$  and  $j$  are selected.

More precisely, for  $0 \leq s, t \leq 1$ ,

$$f(s, t) = 1 - E[(1 - s)^D] - E[(1 - t)^D] + E[(1 - s - t)^D]. \tag{3.4}$$

Now consider the following construction  $\mathcal{K}_\infty^{er}$  of a random network on the vertex set  $\mathbb{Z}^+$ , where every vertex is connected to every other vertex by a single edge. Furthermore, each edge  $(i, j)$  has a random weight  $l_{ij}$  where, given  $\{P_i\}_{i \geq 1}$ , the collection  $\{l_{ij}\}_{1 \leq i < j < \infty}$  are conditionally independent with distribution

$$P(l_{ij} > x) = \exp\left(-\frac{1}{2}f(P_i, P_j)x^2\right). \tag{3.5}$$

Let  $W_{ij}^{er}$  and  $H_{ij}^{er}$  denote the weight and number of edges of the minimal-weight path in  $\mathcal{K}_\infty^{er}$  between the vertices  $i, j \in \mathbb{Z}^+$ . Our analysis will, in particular, show that the FPP on  $\mathcal{K}_\infty^{er}$  is well defined (see Proposition 6.4 below).

Finally, construct the random variables  $D^{\text{er}}$  and  $I^{\text{er}}$  as follows. Let  $D \sim F$ , and consider a multinomial experiment with  $D$  independent trials where at each trial, we choose cell  $i$  with probability  $P_i$ . Let  $D^{\text{er}}$  be the number of *distinct* cells so chosen, and suppose that the cells chosen are  $\mathcal{A} = \{a_1, a_2, \dots, a_{D^{\text{er}}}\}$ . Then let  $I^{\text{er}} \in \mathbb{Z}_+$  be a cell chosen uniformly at random amongst this random set  $\mathcal{A}$ . In terms of these constructs, we are in a position to describe the limiting distribution of the hopcount in the erased CM.

**Theorem 3.2.** (FPP asymptotics for the erased CM.) *Consider the random network  $\mathcal{G}_n^{\text{er}}$ , with the degree distribution  $F$  satisfying (3.3) for some  $\tau \in (1, 2)$ .*

- (a) *Let  $W_n^{\text{er}}$  be the weight of the minimal-weight path between two uniformly chosen vertices in the network. Then*

$$W_n^{\text{er}} \xrightarrow{D} V_1^{\text{er}} + V_2^{\text{er}}, \tag{3.6}$$

where  $V_i^{\text{er}}, i = 1, 2$ , are independent random variables with  $V_i^{\text{er}} \sim E_i/D_i^{\text{er}}$ , where  $E_i$  is exponential with rate 1, and  $D_1^{\text{er}}$  and  $D_2^{\text{er}}$  are, conditionally on  $\{P_i\}_{i \geq 1}$ , independent random variables distributed as  $D^{\text{er}}$ , independently of  $E_1$  and  $E_2$ . More precisely, as  $n \rightarrow \infty$ ,

$$\sqrt{n}(W_n^{\text{er}} - (V_1^{\text{er}} + V_2^{\text{er}})) \xrightarrow{D} W_{I^{\text{er}} J^{\text{er}}}^{\text{er}}, \tag{3.7}$$

where  $I^{\text{er}}$  and  $J^{\text{er}}$  are two copies of the random variable  $I^{\text{er}}$  described above, which are conditionally independent given  $P = \{P_i\}_{i \geq 1}$ .

- (b) *Let  $H_n^{\text{er}}$  be the number of edges in the minimal-weight path between two uniformly chosen vertices in the network. Then*

$$H_n^{\text{er}} \xrightarrow{D} 2 + 2H_{I^{\text{er}} J^{\text{er}}}^{\text{er}}. \tag{3.8}$$

*In particular, the limiting probability measure of the hopcount is supported only on the even integers.*

**Remark 3.1.** In the erased model, the connection between the limiting probabilities,

$$\{P(H_{I^{\text{er}} J^{\text{er}}}^{\text{er}} = k)\}_{k \geq 1},$$

and the Poisson–Dirichlet distribution,  $\{P_k\}_{k \geq 1}$ , becomes cumbersome, in part because by erasing multiple edges we change the degrees. Therefore, we have decided not to introduce quantities similar to the probabilities  $\{\pi_k\}_{k \geq 1}$  introduced in Theorem 3.1. For more details on the limiting probabilities, we refer the reader to Section 6.

We now present an intuitive explanation of the results in Theorem 3.2, starting with (3.6). We let  $A_1$  and  $A_2$  denote two uniformly chosen vertices; note that they can be identical with probability  $1/n$ . We further note that both vertices  $A_1$  and  $A_2$  have random degrees which are close to independent copies of  $D$ . From (2.2) we see that the maximal degree scales as  $u_n = n^{1/(\tau-1)}l(n)$ , where  $l$  is slowly varying at  $\infty$ . We informally refer to the vertices with degrees  $\Theta_P(n^{1/(\tau-1)})$  as *super vertices* (see (6.1) below for a precise definition), and the other vertices as *normal vertices*. We frequently make use of the fact that normal vertices are, w.h.p., exclusively attached to super vertices; see, e.g. [23]. Discounting multiplicity, the number of unique super vertices to which  $A_i, i = 1, 2$ , is attached to is equal to  $D_i^{\text{er}}$ , as described above. The minimal-weight edge between  $A_i, i = 1, 2$ , and any of its neighbors is hence equal in distribution to the minimum of a total of  $D_i^{\text{er}}$  independent, exponentially distributed random variables with rate 1. The shortest-weight path between two super vertices can pass through

intermediate normal vertices, of which there are  $\Theta_P(n)$ . This implies that the minimal weight between any pair of super vertices is of order  $\Theta_P(1)$ , so that the main contribution to  $W_n^{er}$  in (3.6) is from the two minimal edges coming out of the vertices  $A_i, i = 1, 2$ . This indicates, at least at an intuitive level, how (3.6) arises.

We proceed with the intuitive explanation of (3.7). We use the fact that, w.h.p., the vertices  $A_i, i = 1, 2$ , are only attached to super vertices. Thus, in (3.7), we investigate the shortest-weight paths between super vertices. Observe that we deal with the erased CM, so between any pair of direct neighbors there exists only *one* edge having an exponentially distributed weight with mean 1. As before, we number the super vertices by  $i = 1, 2, \dots$ , starting from the largest degree. We denote the number of common neighbors of the super vertices  $i$  and  $j$  by  $N_{ij}^{er}$ , which we will show is  $\Theta_P(n)$ .

Each element in  $N_{ij}^{er}$  corresponds to a unique two-edge path between the super vertices  $i$  and  $j$ . Therefore, the weight of the minimal two-edge path between the super vertices  $i$  and  $j$  has distribution  $w_{ij}^{(n)} \equiv \min_{s \in N_{ij}^{er}} (E_{is} + E_{sj})$ . Note that  $\{E_{is} + E_{sj}\}_{s \in N_{ij}^{er}}$  is a collection of  $N_{ij}^{er}$  i.i.d. gamma(2,1) random variables. More precisely,  $N_{ij}^{er}$  behaves as  $nf(P_i^{(n)}, P_j^{(n)})$ , where  $P_i^{(n)} = D_{(n+1-i:n)}/L_n$ . Indeed, when we consider an *arbitrary* vertex with degree  $D \sim F$ , the probability, conditioned on  $\{P_i^{(n)}\}_{i \geq 1}$ , that this vertex is both connected to super vertex  $i$  and super vertex  $j$  equals

$$1 - (1 - P_i^{(n)})^D - (1 - P_j^{(n)})^D + (1 - P_i^{(n)} - P_j^{(n)})^D.$$

Thus, the expected number of vertices connected to both super vertices  $i$  and  $j$  is, conditionally on  $\{P_i^{(n)}\}_{i \geq 1}, N_{ij}^{er} \approx nf(P_i^{(n)}, P_j^{(n)})$ , and  $f(P_i^{(n)}, P_j^{(n)})$  weakly converges to  $f(P_i, P_j)$ .

We conclude that the minimal two-edge path between super vertex  $i$  and super vertex  $j$  is the minimum of  $nf(P_i^{(n)}, P_j^{(n)})$  gamma(2,1) random variables  $Y_s$ , which are close to being independent. Extreme value theory implies that

$$\lim_{n \rightarrow \infty} P\left(\sqrt{n} \min_{1 \leq s \leq \beta n} Y_s > x\right) = e^{-\beta x^2/2}$$

for any  $\beta > 0$ , which, with  $\beta = \beta_{ij} = f(P_i^{(n)}, P_j^{(n)}) \approx f(P_i, P_j)$ , explains the weights  $l_{ij}$  defined in (3.5), and also explains intuitively why (3.7) holds.

The convergence in (3.8) is explained in a similar manner. Observe that in (3.8) the first 2 on the right-hand side originates from the two edges that connect  $A_1$  and  $A_2$  to the minimal-weight super vertex. Furthermore, the factor 2 in front of  $H_n^{er}$  is due to the fact that shortest-weight paths between super vertices are concatenations of two-edge paths with random weights  $l_{ij}$ . We will further show that two-edge paths, consisting of an alternate sequence of super and normal vertices, form, w.h.p., the minimal-weight paths between super vertices.

This completes the intuitive explanation of Theorem 3.2.

**Remark 3.2.** (*The critical cases  $\tau = 1$  and  $\tau = 2$ .*) In Theorem 3.1 and Theorem 3.2 we considered FPP on the CM with  $\tau \in (1, 2)$ . Furthermore, FPP on the CM for  $\tau > 2$  has been studied in [6]. A natural question that arises is: what happens for  $\tau = 1$  and  $\tau = 2$ , respectively?

For  $\tau = 1$ , it is well known that, with probability 1,  $D_{(n:n)}/L_n \rightarrow 1$ ; see, e.g. [21]. This means that, for  $\tau = 1$ , the CM contains one super vertex, whose degree is, with probability 1, asymptotically equivalent to the sum of the degrees of all vertices. Our two uniformly chosen vertices  $A_1$  and  $A_2$  are hence, w.h.p., exclusively connected to this single super vertex. Hence, in the original model, the hopcount converges to the value 2 in probability and the weight of

the minimal-weight path converges in distribution to the right-hand side of (3.1). We conclude that Theorem 3.1 remains valid for  $\tau = 1$ . For the erased model and  $\tau = 1$ , assuming the asymptotics in (2.1), we have  $H_n^{\text{er}} \xrightarrow{D} 2$ , since, as we can easily conclude from the definition,  $I^{\text{er}} = J^{\text{er}} = 1$  with high probability. Note that we again used the fact that the numbering of the vertices starts from the largest degree. Also, (3.6) remains valid, with the interpretation that  $V_i^{\text{er}} = E_i$ ,  $i = 1, 2$ , where the  $E_i$  are independent, exponentially distributed random variables with rate 1. This follows since  $D_1^{\text{er}} = D_2^{\text{er}} = 1$  with high probability. Hence, with the obvious changes, Theorem 3.2 also remains valid for  $\tau = 1$ .

What happens for  $\tau = 2$  is less clear to us. A natural conjecture would be that the asymptotic mean of the hopcount can interpolate between its asymptotics  $\log n$  for  $\tau > 2$  and  $\Theta_{\mathbb{P}}(1)$  for  $\tau \in (1, 2)$ . It would be of interest to investigate whether the CLT remains valid as long as the mean hopcount tends to  $\infty$ .

### 3.3. Robustness and fragility

The above results show that the hopcount  $H_n$  in both models converges in distribution as  $n \rightarrow \infty$ . Interpreting the hopcount as the amount of time it takes for messages to get from one typical vertex to another typical vertex, the above shows that the CM with  $\tau \in (1, 2)$  is remarkably efficient in routing flow between vertices. We will now show that there exists a downside to this efficiency. The theorem is stated for the more natural erased CM, but one could formulate a corresponding theorem for the original CM under the assumptions of Theorem 3.1.

**Theorem 3.3.** (Robustness and fragility.) *Consider the random weighted network  $\mathcal{G}_n^{\text{er}}$ , where the degree distribution satisfies (3.3) for some  $\tau \in (1, 2)$ . Then, the following properties hold.*

- (a) *Robustness: suppose that an adversary attacks the network by randomly and independently deleting each vertex with probability  $1 - p$  and leaving each vertex with probability  $p$ . Then, for any  $p > 0$ , there exists a unique giant component of size  $\Theta_{\mathbb{P}}(n)$ .*
- (b) *Fragility: suppose that an adversary attacks the network by deleting vertices of maximal degree. Then, for any  $\varepsilon > 0$ , there exists an integer  $K_\varepsilon < \infty$  such that deleting the  $K_\varepsilon$  maximal degree vertices implies that, for two vertices  $A_1$  and  $A_2$  chosen uniformly at random from  $\mathcal{G}_n^{\text{er}}$ ,*

$$\limsup_{n \rightarrow \infty} \mathbb{P}(A_1 \leftrightarrow A_2) \leq \varepsilon, \quad (3.9)$$

where  $A_1 \leftrightarrow A_2$  means that there exists a path connecting vertex  $A_1$  and  $A_2$  after deletion of the maximal vertices. Thus, we can disconnect the network by deleting  $O_{\mathbb{P}}(1)$  vertices.

**Remark.** As in much of percolation theory, we could ask for the size of the giant component in part (a) above when we randomly delete vertices. See Section 7, where we find the size of the giant component as  $n \rightarrow \infty$ , and give the idea of the proofs for this behavior.

## 4. Discussion and related literature

In this section we discuss the literature and state some further open problems and conjectures.

*The configuration model.* The CM is one of the most popular models of random networks, in part because one can generate random graphs with any prescribed degree distribution. This model is often used to understand how the degree distribution affects various models of dynamics on the network, such as epidemics, flows, etc. See in particular [16] and [17] for applications

to modeling of disease epidemics, [19] for a full survey of various questions from statistical physics, and [9, Section 3] for a more rigorous treatment of such questions.

For the CM, the graph distance is well understood. We refer the reader to [26] for  $\tau > 3$ , [27] and [22] for  $\tau \in (2, 3)$ , and [23] for  $\tau \in (1, 2)$ . In the latter paper, it was shown that the graph distance weakly converges, where the limit takes the values two or three, each with positive probability.

*FPP on random graphs.* Analysis of FPP in the context of modern random graph models started only recently (see [5], [13], [28], [24], and [25]). The particular case of the CM with degree distribution  $1 - F(x) = x^{1-\tau}L(x)$ , where  $\tau > 2$ , was studied in [6]. For  $\tau > 2$ , the hopcount remarkably scales as  $\Theta(\log n)$  and satisfies a CLT with asymptotic mean and variance both equal to  $\alpha \log n$  for some  $\alpha > 0$  (see [6]), despite the fact that, for  $\tau \in (2, 3)$ , the graph distance scales as  $\log \log n$ . The parameter  $\alpha = 2(\tau - 2)/(\tau - 1) \in (0, 1)$  for  $\tau \in (2, 3)$ , while  $\alpha = E[D(D - 1)]/E[D(D - 2)] > 1$  for  $\tau > 3$ , where  $D$  has the degree distribution  $F$ . Hence, the limiting distribution for the hopcount depends only on  $\tau$  for  $\tau \in (2, 3)$ , whereas, for  $\tau > 3$ , it depends only on the first two moments of  $F$ . By Theorems 3.1 and 3.2, the behavior for  $\tau \in (1, 2)$ , where the hopcount remains bounded and weakly converges, is rather different from the one for  $\tau > 2$ .

*Universality of  $\mathcal{K}_\infty^{\text{or}}$  and  $\mathcal{K}_\infty^{\text{er}}$ .* Although we have used exponential edge weights, we believe that one obtains the same result with ‘similar’ edge weight distributions with density  $g$  satisfying  $g(0) = 1$ . More precisely, in the hopcount result, the descriptions of  $\mathcal{K}_\infty^{\text{or}}$  and  $\mathcal{K}_\infty^{\text{er}}$  and the corresponding limiting distributions in Theorems 3.1–3.2 will remain unchanged. The only thing that will change is the distribution of  $(V_1^{\text{or}}, V_2^{\text{or}})$  and  $(V_1^{\text{er}}, V_2^{\text{er}})$ . In Section 8, Theorem 8.1, we state what happens when the weight density  $g$  satisfies  $g(0) = \zeta \in (0, \infty)$ . When the edge weight density  $g$  satisfies  $g(0) = 0$  or  $g(0) = \infty$ , then we expect that the hopcount remains tight, but the weight of the minimal path  $W_n$ , as well as the limiting FPP constructions for both the original and erased CM change.

*Robustness and fragility of random networks.* The issue of robustness and fragility of random network models has stimulated an enormous amount of research in the recent past. See [2] for one of the original statistical physics papers on this topic, and [7] for a rigorous derivation of this fact when the power-law exponent  $\tau = 3$  and the random graph is the preferential attachment model. The following universal property is believed to hold for a wide range of models.

- If the degree exponent  $\tau$  of the model is in  $(1, 3]$  then the network is robust against random attacks but fragile against directed attacks, while, for  $\tau > 3$ , under random deletion of vertices, there exists a critical (model dependent)  $p_c$  such that, for  $p < p_c$ , there is no giant component, while, for  $p > p_c$ , there is a giant component.

Proving these results in a wide degree of generality is a challenging program in modern applied probability.

*Load distributions on random networks.* Understanding the FPP model on these networks opens the door to the analysis of more complicated functionals, such as the load distribution on various vertices and edges of the network. Such constructs measure the ability of the network in dealing with congestion when transporting material from one part of the network to another. We discuss such questions in some more detail in Section 8.

*Organization of the proofs and conventions on notation.* The proofs in this paper are organized as follows. In Section 5 we prove the results for the original CM, while Section 6 contains the

proofs for the erased CM. Theorem 3.3 is proved in Section 7, and we close with a conclusion and discussion in Section 8.

In order to simplify notation, we drop the superscripts ‘er’ and ‘or’ so that, for example, the minimal-weight random variable  $W_n^{or}$  between two uniformly selected vertices will be denoted by  $W_n$  when proving facts about the original CM in Section 5, while  $W_n$  will be used to denote  $W_n^{er}$  when proving facts about the erased CM in Section 6.

### 5. Proofs in the original CM: Theorem 3.1

In this section we prove Theorem 3.1. As stated above, in the proofs below, we delete the superscript or from various random variables, such as  $I^{or}$  and  $J^{or}$ , for notational convenience. As part of the proof, we also prove that the FPP on  $\mathcal{K}_\infty^{or}$  is well defined, as formalized in the following proposition.

**Proposition 5.1.** (FPP on  $\mathcal{K}_\infty^{or}$  is well defined.) *For all  $i, j$  in  $\mathcal{K}_\infty^{or}$ , we have  $W_{ij}^{or} > 0$  for  $i \neq j$  and  $H_{ij}^{or} < \infty$ . In particular, this implies that  $H_{I^{or} J^{or}}^{or} < \infty$  a.s., where we recall that  $I^{or}$  and  $J^{or}$  are two random vertices in  $\mathbb{Z}_+$  chosen (conditionally) independently with distribution  $\{P_i\}_{i \geq 1}$ .*

Recall that we label vertices according to their degree. We let  $A_i, i = 1, 2$ , denote two uniformly chosen vertices. Since the CM with infinite mean degrees has, w.h.p., a giant component containing  $n - o(n)$  vertices, w.h.p.,  $A_1$  and  $A_2$  will be connected. We note that the edge incident to vertex  $A_i$  with minimal weight has weight given by  $V_i = E_i/D_{A_i}, i = 1, 2$ , where  $D_{A_i}$  denotes the degree of vertex  $A_i$ . As a result,  $V_i, i = 1, 2$ , has the same distribution as  $E_i/D_i$ , where the  $D_i$  are two independent random variables with distribution function  $F$ . Furthermore, by [23, Theorem 1.1], w.h.p., the vertices  $A_1$  and  $A_2$  are not directly connected. When  $A_1$  and  $A_2$  are not directly connected, then  $W_n \geq V_1 + V_2$ , and  $V_1$  and  $V_2$  are independent, as they depend on the exponential weights of *disjoint* sets of edges, while, by construction,  $D_{A_1}$  and  $D_{A_2}$  are independent. This proves the required lower bound in Theorem 3.1(a). For the upper bound, we further note that, by [23, Lemma 2.2], the vertices  $A_1$  and  $A_2$  are, w.h.p., exclusively connected to so-called *super vertices*, which are the  $m_n$  vertices with the largest degrees, for any  $m_n \rightarrow \infty$  arbitrarily slowly. Thus, the upper bound follows if any two of such super vertices are connected by an edge with weight which converges to 0 in distribution. Denote by  $M_{i,j}$  the minimal weight of all edges connecting the vertices  $i$  and  $j$ . Then, conditionally on the number of edges between  $i$  and  $j$ , we have  $M_{i,j} \sim \text{Exp}(N(i, j))$ , where  $N(i, j)$  denotes the number of edges between  $i$  and  $j$ , and where we use  $\text{Exp}(\lambda)$  to denote an exponential random variable with rate  $\lambda$ . We further define  $P_i^{(n)} = D_{(n+1-i:n)}/L_n$ , so that  $P^{(n)} = \{P_i^{(n)}\}_{i=1}^n$  converges in distribution to the Poisson–Dirichlet distribution  $P = \{P_i\}_{i \geq 1}$  in (2.2)–(2.4). We show that, conditionally on the degrees and w.h.p.,

$$N(i, j) = (1 + o_P(1))L_n P_i^{(n)} P_j^{(n)}. \tag{5.1}$$

Indeed, we note that

$$N(i, j) = \sum_{s=1}^{D_i} I_s(i, j),$$

where  $I_s(i, j)$  is the indicator that the  $s$ th stub of vertex  $i$  connects to  $j$ . We write  $P_n$  for the conditional distribution given the degrees, and  $E_n$  for the expectation with respect to  $P_n$ . It turns out that we can even prove Theorem 3.1 *conditionally on the degrees*, which is stronger

than Theorem 3.1 averaged over the degrees. For this, we note that, for  $1 \leq s_1 < s_2 \leq D_i$ ,

$$P_n(I_{s_1}(i, j) = 1) = \frac{D_j}{L_n - 1}, \quad P_n(I_{s_1}(i, j) = I_{s_2}(i, j) = 1) = \frac{D_j(D_j - 1)}{(L_n - 1)(L_n - 3)},$$

which implies, further using the fact that  $D_j = D_{(n+1-j:n)}$  and, thus,  $D_j/L_n \xrightarrow{D} P_j$ , that

$$\text{var}_n(N(i, j)) \leq C \frac{D_i^2 D_j}{L_n^2} = o_P\left(\frac{D_i^2 D_j^2}{L_n^2}\right) = o_P(\mathbb{E}_n[N(i, j)]^2).$$

As a result,  $N(i, j)$  is concentrated, and, thus, (5.1) follows.

In particular, we see that the vector  $\{N(i, j)/L_n\}_{i,j=1}^n$  converges in distribution to  $\{P_i P_j\}_{i,j=1}^\infty$ . Thus, for every  $i, j$ , and conditionally on the degrees,  $M_{i,j}$  is approximately equal to an exponential random variable with asymptotic mean  $L_n P_i P_j$ . Let  $J_i = J_i^{\text{or}}$  denote the super vertex to which  $A_i$ ,  $i = 1, 2$ , is connected through the minimal-weight edge. Then, conditioned on  $P = \{P_k\}_{k \geq 1}$ , we have, w.h.p.,

$$P(J_i = k \mid P) = P_k, \quad i = 1, 2,$$

and, by (5.1),

$$V_1 + V_2 \leq W_n \leq V_1 + V_2 + \text{Exp}(L_n P_{J_1} P_{J_2}) + o_P(1).$$

Since  $L_n/u_n$  has a distributional limit (see (2.2)) and  $P_{J_i} \in (0, 1)$ ,  $i = 1, 2$ , it follows that  $u_n(W_n - (V_1 + V_2))$  is a tight sequence. Below, we prove that, in fact,  $u_n(W_n - (V_1 + V_2))$  converges weakly to a nontrivial random variable.

Observe that  $H_n = 2$  precisely when  $J_1 = J_2$ , which occurs, by the conditional independence of  $J_1$  and  $J_2$  given  $P$ , with asymptotic probability

$$P_n(H_n = 2) = \sum_{i=1}^\infty P_i^2 + o_P(1).$$

Taking expectations, and using (2.6) together with the bounded convergence theorem, proves (3.2).

We now prove the weak convergence of  $H_n$  and of  $u_n(W_n - (V_1 + V_2))$  by constructing a shortest-weight tree in  $\mathcal{K}_\infty^{\text{or}}$ . We start building the shortest-weight tree from  $J_1$ , terminating when  $J_2$  appears for the first time in this tree. We denote the tree of size  $l$  by  $T_l$ , and note that  $T_1 = \{J_1\}$ . Now we have the following recursive procedure to describe the asymptotic distribution of  $T_l$ . We note that, for any set of vertices  $A$ , the edge with minimal weight outside of  $A$  is a uniform edge pointing outside of  $A$ . When we have already constructed  $T_{l-1}$ , and we fix  $i \in T_{l-1}$ ,  $j \notin T_{l-1}$ , then, by (5.1), there are approximately  $L_n P_i P_j$  edges linking  $i$  and  $j$ . Thus, the probability that vertex  $j \notin T_{l-1}$  is added to  $T_{l-1}$  is, conditionally on  $P$ , approximately equal to

$$p_{ij}(l) = \frac{L_n P_j \sum_{a \in T_{l-1}} P_a}{L_n \sum_{a \in T_{l-1}, b \notin T_{l-1}} P_a P_b} = \frac{P_j}{1 - P_{T_{l-1}}} \geq P_j, \tag{5.2}$$

where, for a set of vertices  $A$ , we write

$$P_A = \sum_{a \in A} P_a.$$

Denote by  $B_l$  the  $l$ th vertex chosen. We stop this procedure when  $B_l = J_2$  for the first time, and denote this stopping time by  $S$ , so that, w.h.p.,  $H_n = 2 + H(S)$ , where  $H(S)$  is the height of  $B_S$  in  $T_S$ . Also,  $u_n(W_n - (V_1 + V_2))$  converges in distribution to  $W_S$ , which is the weight of the path linking  $J_1$  and  $J_2$  in  $\mathcal{K}_\infty^{\text{er}}$ .

Note that the above procedure terminates in finite time, since  $P_{J_2} > 0$  and at each time, and, by (5.2), we pick  $J_2$  with probability at least  $P_{J_2}$ . This proves that  $H_n$  weakly converges, and that the distribution is given only in terms of  $P$ . Also, it proves that the FPP problem on  $\mathcal{K}_\infty^{\text{er}}$  is well defined, as formalized in Proposition 5.1.

Furthermore, since the distribution of  $P$  depends only on  $\tau \in [1, 2)$ , and not on any other details of the degree distribution  $F$ , the same follows for  $H_n$ . When  $\tau = 1$ , then  $P_1 = 1$  a.s., so that  $P_n(H_n = 2) = 1 + o_p(1)$ . When  $\tau \in (1, 2)$ , on the other hand,  $P_i > 0$  a.s. for each  $i \in \mathbb{N}$ , so that, by the above construction, it is not hard to see that  $\lim_{n \rightarrow \infty} P_n(H_n = k) = \pi_k(P) > 0$  a.s. for each  $k \geq 2$ . Thus, the same follows for  $\pi_k = \lim_{n \rightarrow \infty} P(H_n = k) = E[\pi_k(P)]$ . It would be of interest to compute  $\pi_k$  for  $k > 2$  explicitly, or even  $\pi_3$ , but this seems a difficult problem.

### 6. Proofs in the erased CM: Theorem 3.2

In this section we prove the various results in the erased setup. In the proofs, we omit the superscript ‘er’ from the limiting random variables, and write, e.g.  $I, J$ , for notational convenience and when no confusion can arise. We start by giving an overview of the proof.

#### 6.1. Overview of the proof of Theorem 3.2

In this section we formulate four key propositions, which, together, make the intuitive proof given below Theorem 3.2 precise, and which combine to a formal proof of Theorem 3.2.

As before, we label vertices by their (original) degree so that vertex  $i$  will be the vertex with the  $i$ th largest degree. Fix a sequence  $\varepsilon_n \rightarrow 0$  arbitrarily slowly. Then, we define the set of *super vertices*  $\mathcal{S}_n$  to be the set of vertices with largest degrees, namely,

$$\mathcal{S}_n = \{i : D_i > \varepsilon_n n^{1/(\tau-1)}\}. \tag{6.1}$$

We refer to  $\mathcal{S}_n^c$  as the set of *normal vertices*.

Recall the definition of the limiting infinite ‘complete graph’  $\mathcal{K}_\infty^{\text{er}}$  defined in Section 3.2 and, for any fixed  $k \geq 1$ , let  $(\mathcal{K}_\infty^{\text{er}})^k$  denote the projection of this object onto the first  $k$  vertices (so that we retain only the first  $k$  vertices  $1, 2, \dots, k$  and the corresponding edges between these vertices). Then the following proposition says that we can move between the super vertices via two-edge paths which have weight  $\Theta(1/\sqrt{n})$ . For notational convenience, we write  $[k] := \{1, 2, \dots, k\}$ .

**Proposition 6.1.** (Weak convergence of the FPP problem.) *Fix  $k$  and consider the subgraph of the CM formed by retaining the  $k$  vertices of maximal degree and all paths connecting any pair of these vertices by a single intermediary normal vertex (i.e. two-edge paths). For any pair of vertices  $i, j \in [k]$ , let  $l_{ij}^{(n)} = \sqrt{n}w_{ij}^{(2)}$ , where  $w_{ij}^{(2)}$  is the minimal weight of all two-edge paths between  $i$  and  $j$  (with  $w_{ij}^{(2)} = \infty$  if they are not connected by a two-edge path). Consider the complete graph  $\mathcal{K}_n^k$  on vertex set  $[k]$  with edge weights  $l_{ij}^{(n)}$ . Then*

$$\mathcal{K}_n^k \xrightarrow{D} (\mathcal{K}_\infty^{\text{er}})^k,$$

where ‘ $\xrightarrow{D}$ ’ denotes the usual finite-dimensional convergence of the  $\binom{k}{2}$  random variables  $l_{ij}^{(n)}$ .

The proof of Proposition 6.1 is deferred to Section 6.2. Proposition 6.1 implies that the FPP problem on the first  $k$  super vertices along the two-edge paths converges in distribution to the one on  $\mathcal{K}_\infty^{\text{er}}$  restricted to  $[k]$ . We now investigate the structure of the minimal weights from a uniform vertex, and the tightness of recentered minimal weight

**Proposition 6.2.** (Coupling of the minimal edges from uniform vertices.) *Let  $(A_1, A_2)$  be two uniform vertices, and let  $(V_1^{(n)}, V_2^{(n)})$  denote the minimal weight in the erased CM along the edges attached to  $(A_1, A_2)$ .*

- (a) *Let  $I^{(n)}$  and  $J^{(n)}$  denote the vertices to which  $A_i, i = 1, 2,$  are connected, and let  $(I^{\text{er}}, J^{\text{er}})$  be two random variables having the distribution specified right before Theorem 3.2, which are conditionally independent given  $\{P_i\}_{i \geq 1}$ . Then, we can couple  $(I^{(n)}, J^{(n)})$  and  $(I^{\text{er}}, J^{\text{er}})$  in such a way that*

$$P((I^{(n)}, J^{(n)}) \neq (I^{\text{er}}, J^{\text{er}})) = o(1). \tag{6.2}$$

- (b) *Let  $V_i = E_i/D_i^{\text{er}}$ , where  $D_i^{\text{er}}$  are two copies of the random variable  $D^{\text{er}}$  described right before Theorem 3.2, which are conditionally independent given  $\{P_i\}_{i \geq 1}$ . Then, we can couple  $(V_1^{(n)}, V_2^{(n)})$  to  $(V_1, V_2)$  in such a way that*

$$P((V_1^{(n)}, V_2^{(n)}) \neq (V_1, V_2)) = o(1). \tag{6.3}$$

*As a result, the recentered random variables  $\sqrt{n}(W_n - (V_1 + V_2))$  form a tight sequence.*

The proof of Proposition 6.2 is deferred to Section 6.3. The following proposition asserts that the hopcount and the recentered weight between the first  $k$  super vertices are tight random variables, and, in particular, they remain within the first  $[K]$  vertices with probability  $1 - o(1)$  as  $K \rightarrow \infty$ .

**Proposition 6.3.** (Tightness of the FPP problem and evenness of the hopcount.) *Fix  $k \geq 1$ . For any pair of vertices  $i, j \in [k]$ , let  $H_n(i, j)$  denote the number of edges of the minimal-weight path between  $i$  and  $j$ . Then,*

- (a)  *$H_n(i, j)$  is a tight sequence of random variables, which is such that  $P(H_n(i, j) \notin 2\mathbb{Z}^+) = o(1)$ ;*
- (b) *the probability that any of the minimal-weight paths between  $i, j \in [k]$ , at even times, leaves the  $K$  vertices of largest degree tends to 0 when  $K \rightarrow \infty$ ;*
- (c) *the hopcount  $H_n$ , being the number of edges of the minimal-weight path between two vertices chosen uniformly at random, is a tight sequence of random variables, which is such that  $P(H_n \notin 2\mathbb{Z}^+) = o(1)$ .*

The proof of Proposition 6.3 is deferred to Section 6.4. The statement is consistent with the intuitive explanation given right after Theorem 3.2: the minimal-weight paths between two uniform vertices consists of an *alternating* sequence of normal vertices and super vertices. We finally state that the infinite FPP on the erased CM is well defined.

**Proposition 6.4.** (The infinite FPP is well defined.) *Consider FPP on  $\mathcal{K}_\infty^{\text{er}}$  with weights  $\{l_{ij}\}_{1 \leq i < j < \infty}$  defined in (3.5). Fix  $k \geq 1$  and  $i, j \in [k]$ . Let  $\mathcal{A}_K$  be the event that there exists a path of weight at most  $W$  connecting  $i$  and  $j$ , which contains a vertex in  $\mathbb{Z}^+ \setminus [K]$ . Then, there exists a  $C > 0$  such that, for all sufficiently large  $K$ ,*

$$P(\mathcal{A}_K) \leq CWK^{-1} e^{CW\sqrt{\log K}}. \tag{6.4}$$

The proof of Proposition 6.4 is deferred to Section 6.5. With Propositions 6.1–6.4 at hand, we are able to prove Theorem 3.2.

*Proof of Theorem 3.2 subject to Propositions 6.1–6.4.* By Proposition 6.2(b) we can couple  $(V_1^{(n)}, V_2^{(n)})$  to  $(V_1, V_2)$  in such a way that  $(V_1^{(n)}, V_2^{(n)}) = (V_1, V_2)$  occurs w.h.p. Furthermore, with probability converging to 1 as  $k \rightarrow \infty$ ,  $I, J \leq k$ , which we assume from now on, while, by Proposition 6.2(b),  $\sqrt{n}(W_n - (V_1 + V_2))$  is a tight sequence of random variables.

By Proposition 6.3, the hopcount is a tight sequence of random variables, which is, w.h.p., even. Indeed, it consists of an alternating sequence of normal and super vertices. We call the path of super vertices the *two-edge path*. Then, Proposition 6.3 implies that the probability that any of the two-edge paths between any of the first  $k$  vertices leaves the first  $K$  vertices is small when  $K$  grows big. As a result, we can write  $H_n = 2 + 2H_{I^{(n)}J^{(n)}}^{(n)}$ , where  $H_{I^{(n)}J^{(n)}}^{(n)}$  is the number of two-edge paths in  $\mathcal{K}_n^{\text{er}}$ . By (6.2) we have, w.h.p.,  $H_{I^{(n)}J^{(n)}}^{(n)} = H_{IJ}^{(n)}$ .

By Proposition 6.1, the FPP on the  $k$  vertices of largest degree in the CM weakly converges to the FPP on the first  $k$  vertices of  $\mathcal{K}_\infty^{\text{er}}$ , for any  $k \geq 1$ . By Proposition 6.4, w.h.p., the shortest-weight path between any two vertices in  $[k]$  in  $\mathcal{K}_\infty^{\text{er}}$  does not leave the first  $K$  vertices, so that  $W_{IJ}$  and  $H_{IJ}$  are finite random variables, where  $W_{IJ}$  and  $H_{IJ}$  denote the weight and number of steps in the minimal path between  $I$  and  $J$  in  $\mathcal{K}_\infty^{\text{er}}$ . In particular, it follows that  $\sqrt{n}(W_n - (V_1^{(n)} + V_2^{(n)})) \xrightarrow{D} W_{IJ}$ , and that  $H_{ij}^{(n)} \xrightarrow{D} H_{ij}$  for every  $i, j \in [k]$ , which is the number of hops between  $i, j \in [k]$  in  $\mathcal{K}_\infty^{\text{er}}$ . Since, w.h.p.,  $(V_1, V_2) = (V_1^{(n)}, V_2^{(n)})$ ,  $\sqrt{n}(W_n - (V_1 + V_2))$  converges to the same limit. This completes the proof of Theorem 3.2 subject to Propositions 6.1–6.4.

**6.2. Weak convergence of the finite FPP problem to  $\mathcal{K}_\infty^{\text{er}}$ : proof of Proposition 6.1**

In this section we study the weak convergence of the FPP on  $\mathcal{K}_n^k$  to the one on  $(\mathcal{K}_\infty^{\text{er}})^k$ , by proving Proposition 6.1.

We start by proving some elementary results regarding the extrema of gamma random variables. We start with a particularly simple case, and after this, generalize it to the convergence of all weights of two-edge paths in  $\mathcal{K}_n^{\text{er}}$ .

**Lemma 6.1.** (Minima of gamma random variable.) (a) Fix  $\beta > 0$  and consider  $n\beta$  i.i.d. gamma(2,1) random variables  $Y_i$ . Let  $T_n = \min_{1 \leq i \leq n\beta} Y_i$  be the minimum of these random variables. Then, as  $n \rightarrow \infty$ ,

$$P(\sqrt{n}T_n > x) \rightarrow \exp(-\frac{1}{2}\beta x^2).$$

(b) Let  $\{X_i\}_{1 \leq i \leq m}$ ,  $\{Y_i\}_{1 \leq i \leq m}$ , and  $\{Z_i\}_{1 \leq i \leq m}$  be all independent collections of independent exponential mean 1 random variables. Let

$$\eta_m = \sqrt{m} \min_{1 \leq i \leq m} (X_i + Y_i), \quad \kappa_m = \sqrt{m} \min_{1 \leq i \leq m} (X_i + Z_i),$$

$$\text{and } \rho_m = \sqrt{m} \min_{1 \leq i \leq m} (Y_i + Z_i).$$

Then, as  $m \rightarrow \infty$ ,

$$(\eta_m, \kappa_m, \rho_m) \xrightarrow{D} (\zeta_1, \zeta_2, \zeta_3).$$

Here the  $\zeta_i$  are independent with distribution function  $1 - \exp(-x^2/2)$ ,  $x \geq 0$ .

We note that the independence claimed in part (b) is nontrivial, in particular, since the random variables  $(\eta_m, \kappa_m, \rho_m)$  are all defined in terms of the same exponential random variables. We will later see a more general version of this result.

*Proof of Lemma 6.5.* Part (a) is quite trivial, so we will leave the proof to the reader and focus on part (b). Note that, for any fixed  $x_0, y_0,$  and  $z_0$  all positive and  $X, Y,$  and  $Z$  all independent exponential random variables, we have

$$P\left(X + Y \leq \frac{x_0}{\sqrt{m}}\right) = \frac{x_0^2}{2m} + O(m^{-3/2}), \tag{6.5}$$

and similar estimates hold for  $P(X + Z \leq y_0/\sqrt{m})$  and  $P(Y + Z \leq z_0/\sqrt{m})$ . Furthermore, we make use of the fact that, for  $m \rightarrow \infty,$

$$P\left(X + Y \leq \frac{x_0}{\sqrt{m}}, X + Z \leq \frac{y_0}{\sqrt{m}}\right) = \Theta(m^{-3/2}),$$

since  $X + Y \leq x_0/\sqrt{m}$  and  $X + Z \leq y_0/\sqrt{m}$  imply that  $X, Y,$  and  $Z$  are all of order  $1/\sqrt{m}$ . Then, we rewrite

$$P(\eta_m > x_0, \kappa_m > y_0, \rho_m > z_0) = P\left(\sum_{i=1}^m I_i = 0, \sum_{i=1}^m J_i = 0, \sum_{i=1}^m L_i = 0\right),$$

where  $I_i = \mathbf{1}_{\{X_i + Y_i < x_0/\sqrt{m}\}}, J_i = \mathbf{1}_{\{X_i + Z_i < y_0/\sqrt{m}\}},$  and  $L_i = \mathbf{1}_{\{Y_i + Z_i < z_0/\sqrt{m}\}},$  where we write  $\mathbf{1}_A$  for the indicator of the event  $A$ . This implies, in particular, that

$$\begin{aligned} P(\eta_m > x_0, \kappa_m > y_0, \rho_m > z_0) &= (P(I_1 = 0, J_1 = 0, L_1 = 0))^m \\ &= (1 - P(\{I_1 = 1\} \cup \{J_1 = 1\} \cup \{L_1 = 1\}))^m \\ &= \left[1 - \left(\frac{x_0^2}{2m} + \frac{y_0^2}{2m} + \frac{z_0^2}{2m} - \Theta(m^{-3/2})\right)\right]^m \\ &= e^{-(x_0^2/2 + y_0^2/2 + z_0^2/2)}(1 + o(1)), \end{aligned} \tag{6.6}$$

as  $m \rightarrow \infty,$  where we use the fact that

$$\begin{aligned} &|P(\{I_1 = 1\} \cup \{J_1 = 1\} \cup \{L_1 = 1\}) - P(I_1 = 1) - P(J_1 = 1) - P(L_1 = 1)| \\ &\leq P(I_1 = J_1 = 1) + P(I_1 = L_1 = 1) + P(J_1 = L_1 = 1) \\ &= \Theta(m^{-3/2}). \end{aligned}$$

This proves the result.

The next lemma generalizes the statement of Lemma 6.1 in a substantial way.

**Lemma 6.2.** (Minima of gamma random variables on the complete graph.) *Fix  $k \geq 1$  and  $n \geq k$ . Let  $\{E_{s,t}\}_{1 \leq s < t \leq n}$  be an i.i.d. sequence of exponential random variables with mean 1. For each  $i \in [k],$  let  $\mathcal{N}_i \subseteq [n] \setminus [k]$  denote deterministic sets of indices. Let  $\mathcal{N}_{ij} = \mathcal{N}_i \cap \mathcal{N}_j,$  and assume that, for each  $i, j \in [k],$*

$$|\mathcal{N}_{ij}|/n \rightarrow \beta_{ij} > 0. \tag{6.7}$$

Let

$$\eta_{ij}^{(n)} = \sqrt{n} \min_{s \in \mathcal{N}_{ij}} (E_{i,s} + E_{s,j}).$$

Then, for each  $k$ ,

$$\{\eta_{ij}^{(n)}\}_{1 \leq i < j \leq k} \xrightarrow{D} \{\eta_{ij}\}_{1 \leq i < j \leq k},$$

where the random variables  $\{\eta_{ij}\}_{1 \leq i < j \leq k}$  are independent random variables with distribution

$$P(\eta_{ij} > x) = \exp\left(-\frac{1}{2}\beta_{ij}x^2\right).$$

When the  $\mathcal{N}_i$  denote random sets of indices which are independent of the exponential random variables, then the same result holds when the convergence in (6.7) is replaced with convergence in distribution, where the limits  $\beta_{ij}$  satisfy that  $\beta_{ij} > 0$  holds a.s. and the limits  $\{\eta_{ij}\}_{1 \leq i < j \leq k}$  are conditionally independent given  $\{\beta_{ij}\}_{1 \leq i < j \leq k}$ .

*Proof.* We follow the proof of Lemma 6.1 as closely as possible. For  $i \in [k]$  and  $s \in [n] \setminus [k]$ , we define  $X_{i,s} = E_{i,s}$  when  $s \in \mathcal{N}_i$ , and  $X_{i,s} = +\infty$  when  $s \notin \mathcal{N}_i$ . Since the sets of indices  $\{\mathcal{N}_i\}_{i \in [k]}$  are independent from the exponential random variables, the variables  $\{X_{i,s}\}_{i \in [k], s \in [n] \setminus [k]}$  are, conditionally on  $\{\mathcal{N}_i\}_{i \in [k]}$ , independent random variables. Then, since  $\mathcal{N}_{ij} = \mathcal{N}_i \cap \mathcal{N}_j$ ,

$$\eta_{ij}^{(n)} = \sqrt{n} \min_{s \in \mathcal{N}_{ij}} (E_{i,s} + E_{j,s}) = \sqrt{n} \min_{s \in [n] \setminus [k]} (X_{i,s} + X_{j,s}).$$

Let  $\{x_{ij}\}_{1 \leq i < j \leq k}$  be a vector with positive coordinates. We note that

$$P(\eta_{ij}^{(n)} > x_{ij} \text{ for all } i, j \in [k]) = P\left(\sum_{s \in [n] \setminus [k]} J_{ij,s} = 0 \text{ for all } i, j \in [k]\right),$$

where  $J_{ij,s} = \mathbf{1}_{\{X_{i,s} + X_{j,s} < x_{ij}/\sqrt{n}\}}$ . We note that the random vectors  $\{J_{ij,s}\}_{s \in [n] \setminus [k]}$  are conditionally independent given  $\{\mathcal{N}_i\}_{i \in [k]}$ , so that

$$P(\eta_{ij}^{(n)} > x_{ij} \text{ for all } i, j \in [k]) = \prod_{s \in [n] \setminus [k]} P(J_{ij,s} = 0 \text{ for all } i, j \in [k]).$$

Now, note that  $J_{ij,s} = 0$  a.s. when  $s \notin \mathcal{N}_{ij}$ , while, for  $s \in \mathcal{N}_{ij}$ , we have, similarly to (6.5),

$$P(J_{ij,s} = 1) = \frac{x_{ij}^2}{2n} + O(n^{-3/2}).$$

Therefore, we can summarize these two claims by

$$P(J_{ij,s} = 1) = \mathbf{1}_{\{s \in \mathcal{N}_{ij}\}} \left(\frac{x_{ij}^2}{2n} + \Theta(n^{-3/2})\right).$$

Similarly to the argument in (6.6), we have

$$\begin{aligned} P(J_{ij,s} = 0 \text{ for all } i, j \in [k]) &= 1 - \sum_{1 \leq i < j \leq k} P(J_{ij,s} = 1) + \Theta(n^{-3/2}) \\ &= \exp\left\{-\sum_{1 \leq i < j \leq k} \mathbf{1}_{\{s \in \mathcal{N}_{ij}\}} \left(\frac{x_{ij}^2}{2n} + \Theta(n^{-3/2})\right)\right\}. \end{aligned}$$

We conclude that

$$\begin{aligned} P(\eta_{ij}^{(n)} > x_{ij} \text{ for all } i, j \in [k]) &= \prod_{s \in [n] \setminus [k]} P(J_{ij,s} = 0 \text{ for all } i, j \in [k]) \\ &= \exp \left\{ - \sum_{s \in [n] \setminus [k]} \sum_{1 \leq i < j \leq k} \mathbf{1}_{\{s \in \mathcal{N}_{ij}\}} \left( \frac{x_{ij}^2}{2n} + \Theta(n^{-3/2}) \right) \right\} \\ &= \exp \left\{ - \sum_{1 \leq i < j \leq k} \frac{x_{ij}^2 \beta_{ij}}{2} \right\} (1 + o(1)), \end{aligned}$$

as required.

We will apply Lemma 6.2 to  $\mathcal{N}_i$  being the direct neighbors in  $[n] \setminus [k]$  of vertex  $i \in [k]$ . Thus, by Lemma 6.2, in order to prove the convergence of the weights, it suffices to prove the convergence of the number of *joint* neighbors of the super vertices  $i$  and  $j$ , simultaneously, for all  $i, j \in [k]$ . That is the content of the following lemma.

**Lemma 6.3.** (Weak convergence of  $N_{ij}^{er}/n$ .) *The random vector  $\{N_{ij}^{er}/n\}_{1 \leq i < j \leq n}$  converges in distribution in the product topology to  $\{f(P_i, P_j)\}_{1 \leq i < j < \infty}$ , where  $f(P_i, P_j)$  is defined in (3.4), and  $\{P_i\}_{i \geq 1}$  has the Poisson–Dirichlet distribution.*

*Proof.* We first prove that the random vector  $\{N_{ij}^{er}/n - f(P_i^{(n)}, P_j^{(n)})\}_{1 \leq i < j \leq n}$  converges in probability in the product topology to 0, where  $P_i^{(n)} = D_{(n+1-i:n)}/L_n$  is the normalized  $i$ th largest degree. For this, we note that

$$N_{ij}^{er} = \sum_{s=1}^n I_s(i, j),$$

where  $I_s(i, j)$  is the indicator that  $s \in [n]$  is a neighbor of both  $i$  and  $j$ . Now, weak convergence in the product topology is equivalent to the weak convergence of  $\{N_{ij}^{er}/n\}_{1 \leq i < j < K}$  for any  $K \in \mathbb{Z}^+$  (see [15, Theorem 4.29]). For this, we use a second moment method. We first note that  $|N_{ij}^{er}/n - N_{\leq b_n}^{er}(i, j)/n| \leq (1/n) \sum_{s=1}^n \mathbf{1}_{\{D_s \geq b_n\}} \xrightarrow{P} 0$ , where  $b_n \rightarrow \infty$  and

$$N_{\leq b_n}^{er}(i, j) = \sum_{s=1}^n I_s(i, j) \mathbf{1}_{\{D_s \leq b_n\}}.$$

Take  $b_n = n$  and note that, when  $i, j \leq K$ , the vertices  $i$  and  $j$  both have degree of order  $n^{1/(\tau-1)}$ , which is at least  $n$  w.h.p. Thus, the sum over  $s$  in  $N_{\leq n}^{er}(i, j)$  involves different vertices than  $i$  and  $j$ . Next, we note that

$$\begin{aligned} E_n \left[ \frac{N_{\leq n}^{er}(i, j)}{n} \right] &= \frac{1}{n} \sum_{s=1}^n \mathbf{1}_{\{D_s \leq n\}} P_n(I_s(i, j) = 1) \\ &= \frac{1}{n} \sum_{s=1}^n \mathbf{1}_{\{D_s \leq n\}} [1 - (1 - P_i^{(n)})^{D_s} - (1 - P_j^{(n)})^{D_s} + (1 - P_i^{(n)} - P_j^{(n)})^{D_s}] + o_P(1), \end{aligned}$$

in a similar way as in (3.4). By dominated convergence we have, for every  $s \in [0, 1]$ ,

$$\frac{1}{n} \sum_{s=1}^n \mathbf{1}_{\{D_s \leq n\}} (1-s)^{D_s} \xrightarrow{\text{a.s.}} \mathbb{E}[(1-s)^D],$$

which implies that

$$\mathbb{E}_n \left[ \frac{N_{\leq n}^{\text{er}}(i, j)}{n} \right] - f(P_i^{(n)}, P_j^{(n)}) \xrightarrow{P} 0.$$

Furthermore, the indicators  $\{I_s(i, j)\}_{s=1}^n$  are close to independent, so that  $\text{var}_n(N_{\leq n}^{\text{er}}(i, j)/n) = o_P(1)$ , where  $\text{var}_n$  denotes the variance with respect to  $P_n$ . The weak convergence claimed in Lemma 6.3 follows directly from the above results, as well as from the weak convergence of the order statistics in (2.2) and the continuity of  $(s, t) \mapsto f(s, t)$ .

The following corollary completes the proof of the convergence of the rescaled minimal-weight two-edge paths in  $\mathcal{G}_n^{\text{er}}$ .

**Corollary 6.1.** (Conditional independence of weights.) *Let  $l_{ij}^{(n)} = \sqrt{n}w_{ij}^{(2)}$ , where  $w_{ij}^{(2)}$  is the minimal weight of all two-edge paths between the vertices  $i$  and  $j$  (with  $w_{ij}^{(2)} = \infty$  if they are not connected by a two-edge path). Fix  $k \geq 1$ . Then*

$$\left( \{l_{ij}^{(n)}\}_{1 \leq i < j \leq k}, \left\{ \frac{D_i}{L_n} \right\}_{1 \leq i \leq n} \right) \xrightarrow{D} (\{l_{ij}\}_{1 \leq i < j \leq k}, \{P_i\}_{i \geq 1}),$$

where, given  $\{P_i\}_{i \geq 1}$ , the random variables  $\{l_{ij}\}_{1 \leq i < j \leq k}$  are conditionally independent with distribution

$$P(l_{ij} > x) = \exp\left(-\frac{1}{2} f(P_i, P_j)x^2\right).$$

*Proof.* The convergence of  $\{D_m/L_n\}_{1 \leq m \leq n}$  follows from Section 2.1. Then we apply Lemma 6.2. We let  $\mathcal{N}_i$  denote the set of neighbors in  $[n] \setminus [k]$  of the super vertex  $i \in [k]$ . Then,  $|\mathcal{N}_{ij}| = |\mathcal{N}_i \cap \mathcal{N}_j| = N_{ij}^{\text{er}}$ , so that (6.7) is equivalent to the convergence in distribution of  $N_{ij}^{\text{er}}/n$ . The latter is proved in Lemma 6.3, with  $\beta_{ij} = f(P_i, P_j)$ . Since  $P_i > 0$  a.s. for each  $i \in [k]$ , we obtain  $\beta_{ij} > 0$  a.s. for all  $i, j \in [k]$ . Therefore, Lemma 6.2 applies, and completes the proof of the claim.

Now we are ready to prove Proposition 6.1.

*Proof of Proposition 6.1.* By Corollary 6.1 we see that the weights in the FPP problem  $\mathcal{K}_n^k$  converge in distribution to the weights in the FPP on  $(\mathcal{K}_\infty^{\text{er}})^k$ . Since the weights  $W_{ij}^{(n)}$  of the minimal two-edge paths between  $i, j \in [k]$  are continuous functions of the weights  $\{l_{ij}^{(n)}\}_{1 \leq i < j \leq k}$ , it follows that  $\{W_{ij}^{(n)}\}_{1 \leq i < j \leq k}$  converges in distribution to  $\{W_{ij}\}_{1 \leq i < j \leq k}$ . Since the weights are continuous random variables, this also implies that the hopcounts  $\{H_{ij}^{(n)}\}_{1 \leq i < j \leq k}$  in  $\mathcal{K}_n^k$  converge in distribution to the hopcounts  $\{H_{ij}\}_{1 \leq i < j \leq k}$  in  $(\mathcal{K}_\infty^{\text{er}})^k$ . This proves Proposition 6.1.

**6.3. Coupling of the minimal edges from uniform vertices: proof of Proposition 6.2**

In this section we prove Proposition 6.2. As mentioned earlier, the vertices  $A_i, i = 1, 2$ , are, w.h.p., only attached to super vertices. Let  $I^{(n)}$  and  $J^{(n)}$  denote the vertices to which  $A_i, i = 1, 2$ , are connected and of which the edge weights are minimal. Then, by the discussion below (3.5),  $(I^{(n)}, J^{(n)})$  converges in distribution to the random vector  $(I, J)$  having the distribution specified right before Theorem 3.2, and where the two components are conditionally independent, given  $\{P_i\}_{i \geq 1}$ .

Furthermore, denote the weight of the edges attaching  $(A_1, A_2)$  to  $(I^{(n)}, J^{(n)})$  by  $(V_1^{(n)}, V_2^{(n)})$ . Then,  $(V_1^{(n)}, V_2^{(n)}) \xrightarrow{D} (V_1^{er}, V_2^{er})$  defined in Theorem 3.2. This in particular proves (3.6) since the weight between any two super vertices is  $o_P(1)$ . Furthermore, since  $(I^{(n)}, J^{(n)})$  are discrete random variables that weakly converge to  $(I, J)$ , we can couple  $(I^{(n)}, J^{(n)})$  and  $(I, J)$  in such a way that (6.2) holds.

Let  $(D_{A_1}^{er(n)}, D_{A_2}^{er(n)})$  denote the erased degrees of the vertices  $(A_1, A_2)$  in  $\mathcal{G}_n^{er}$ . The following lemma states that these erased degrees converge in distribution.

**Lemma 6.4.** (Convergence in distribution of erased degrees.) *Under the conditions of Theorem 3.2, as  $n \rightarrow \infty$ ,*

$$(D_{A_1}^{er(n)}, D_{A_2}^{er(n)}) \xrightarrow{D} (D_1^{er}, D_2^{er}), \tag{6.8}$$

which are two copies of the random variable  $D^{er}$  described right before Theorem 3.2, and which are conditionally independent given  $\{P_i\}_{i \geq 1}$ .

*Proof.* Conditionally on  $A_1 \neq A_2$ , which occurs with probability  $1 - 1/n$ , the degrees before erasure,  $(D_{A_1}, D_{A_2})$ , are i.i.d. with distribution function  $F$ . Therefore, in particular,  $D_{A_1}$  and  $D_{A_2}$  are bounded by  $K$  with probability  $1 - o(1)$  as  $K \rightarrow \infty$ . We next investigate the effect of erasure. We condition on  $\{P_i^{(n)}\}_{i=1}^{m_n}$ , the rescaled  $m_n$  largest degrees, and note that, by (2.2),  $\{P_i^{(n)}\}_{i=1}^{m_n} = \{D_i/L_n\}_{i=1}^{m_n}$  converges in distribution to  $\{P_i\}_{i \geq 1}$ . We let  $m_n \rightarrow \infty$  arbitrarily slowly, and note that, w.h.p., the  $(D_{A_1}, D_{A_2})$  half-edges incident to the vertices  $(A_1, A_2)$  are exclusively connected to vertices in  $[m_n]$ . The convergence in (6.8) follows when

$$\begin{aligned} P((D_{A_1}^{er(n)}, D_{A_2}^{er(n)}) = (k_1, k_2) \mid \{P_i^{(n)}\}_{i=1}^{m_n}, (D_{A_1}, D_{A_2}) = (j_1, j_2)) \\ = G_{k_1, j_1}(\{P_i^{(n)}\}_{i=1}^{m_n}) G_{k_2, j_2}(\{P_i^{(n)}\}_{i=1}^{m_n}) + o_P(1), \end{aligned} \tag{6.9}$$

for an appropriate function  $G_{k,j} : \mathbb{R}_+^{\mathbb{N}} \rightarrow [0, 1]$ , which, for every  $k, j$ , is continuous in the product topology. (By convention, for a vector with finitely many coordinates  $\{x_i\}_{i=1}^m$ , we let  $G_{k_1, j_1}(\{x_i\}_{i=1}^m) = G_{k_1, j_1}(\{x_i\}_{i=1}^{\infty})$ , where  $x_i = 0$  for  $i > m$ .)

Indeed, from (6.9) it follows that, by dominated convergence,

$$\begin{aligned} P((D_{A_1}^{er(n)}, D_{A_2}^{er(n)}) = (k_1, k_2)) \\ = E[P((D_{A_1}^{er(n)}, D_{A_2}^{er(n)}) = (k_1, k_2) \mid \{P_i^{(n)}\}_{i=1}^{m_n}, (D_{A_1}, D_{A_2}))] \\ = E[G_{k_1, D_1}(\{P_i^{(n)}\}_{i=1}^{m_n}) G_{k_2, D_2}(\{P_i^{(n)}\}_{i=1}^{m_n})] + o(1) \\ \rightarrow E[G_{k_1, D_1}(\{P_i\}_{i \geq 1}) G_{k_2, D_2}(\{P_i\}_{i \geq 1})], \end{aligned}$$

where the last convergence follows from the weak convergence of  $\{P_i^{(n)}\}_{i=1}^{m_n}$  and the assumed continuity of  $G$ . The above convergence, in turn, is equivalent to (6.8), when  $G_{k,j}(\{P_i\}_{i \geq 1})$  denotes the probability that  $k$  distinct cells are chosen in a multinomial experiment with  $j$  independent trials where, at each trial, we choose cell  $i$  with probability  $P_i$ . It is not hard to see that, for each  $k, j$ ,  $G_{k,j}$  is indeed a continuous function in the product topology.

To see (6.9), we note that, conditionally on  $\{P_i^{(n)}\}_{i=1}^{m_n}$ , the vertices to which the  $D_{A_i} = j_i$  stubs attach are close to independent, so that it suffices to prove that

$$P(D_{A_1}^{er(n)} = k_1 \mid \{P_i^{(n)}\}_{i=1}^{m_n}, D_{A_1} = j_1) = G_{k_1, j_1}(\{P_i^{(n)}\}_{i=1}^{m_n}) + o_P(1).$$

The latter follows, since, again conditionally on  $\{P_i^{(n)}\}_{i=1}^{m_n}$ , each stub chooses to connect to vertex  $i$  with probability  $D_i/L_n = P_i^{(n)}$ , and the different stubs choose close to independently. This completes the proof of Lemma 6.4.

By Lemma 6.4 we can also couple  $(D_{A_1}^{er(n)}, D_{A_2}^{er(n)})$  to  $(D_1^{er}, D_2^{er})$  in such a way that

$$P((D_{A_1}^{er(n)}, D_{A_2}^{er(n)}) \neq (D_1^{er}, D_2^{er})) = o(1).$$

Now,  $(V_1^{(n)}, V_2^{(n)})$  is equal in distribution to  $(E_1/D_{A_1}^{er(n)}, E_2/D_{A_2}^{er(n)})$ , where  $(E_1, E_2)$  are two independent exponential random variables with rate 1 independent of  $(D_{A_1}^{er(n)}, D_{A_2}^{er(n)})$ . Let  $V_i = V_i^{er} = E_i/D_i^{er}$ , where we use the *same* exponential random variables. Then  $(V_1, V_2)$  has the right distribution, and the above coupling also provides a coupling of  $(V_1^{(n)}, V_2^{(n)})$  to  $(V_1, V_2)$  such that (6.3) holds.

By the above couplings we have

$$\sqrt{n}(W_n - (V_1^{(n)} + V_2^{(n)})) = \sqrt{n}(W_n - (V_1 + V_2)) \quad \text{w.h.p.}$$

By construction,  $\sqrt{n}(W_n - (V_1^{(n)} + V_2^{(n)})) \geq 0$  a.s., so that also, w.h.p.,

$$\sqrt{n}(W_n - (V_1 + V_2)) \geq 0.$$

Furthermore,  $\sqrt{n}(W_n - (V_1^{(n)} + V_2^{(n)})) \leq l_{I^{(n)}, J^{(n)}}^{(n)}$ , which is the weight of the minimal two-edge path between the super vertices  $I^{(n)}$  and  $J^{(n)}$  normalized by  $\sqrt{n}$ . Now, by (6.2),

$$(I^{(n)}, J^{(n)}) = (I, J) \quad \text{w.h.p.}$$

Thus, w.h.p.,  $l_{I^{(n)}, J^{(n)}}^{(n)} = l_{I, J}^{(n)}$ , which, by Proposition 6.1, converges in distribution to  $l_{IJ}$ , which is a finite random variable. As a result,  $l_{I^{(n)}, J^{(n)}}^{(n)}$  is a tight sequence of random variables, and, therefore, so is  $\sqrt{n}(W_n - (V_1^{(n)} + V_2^{(n)}))$ . This completes the proof of Proposition 6.2.

**6.4. Tightness of the FPP problem and evenness of the hopcount: proof of Proposition 6.3**

In this section we prove that the only possible minimal-weight paths between the super vertices are two-edge paths. All other paths are much too costly to be used. We start by stating and proving a technical lemma about expectations of degrees conditioned to be at most  $x$ . It is here that we make use of the condition in (3.3).

**Lemma 6.5.** (Bounds on restricted moments of  $D$ .) *Let  $D$  be a random variable with distribution function  $F$  satisfying (3.3) for some  $\tau \in (1, 2)$ . Then, there exists a constant  $C$  such that, for every  $x \geq 1$ ,*

$$E[D\mathbf{1}_{\{D \leq x\}}] \leq Cx^{2-\tau}, \quad E[D^{\tau-1}\mathbf{1}_{\{D \leq x\}}] \leq C \log x, \quad E[D^\tau\mathbf{1}_{\{D \leq x\}}] \leq Cx. \quad (6.10)$$

*Proof.* We note that, for every  $a > 0$ , using partial integration,

$$E[D^a\mathbf{1}_{\{D \leq x\}}] = - \int_{(0,x]} y^a d(1 - F(y)) \leq a \int_0^x y^{a-1}[1 - F(y)] dy \leq c_2 a \int_0^x y^{a-\tau} dy. \quad (6.11)$$

The proof is completed by considering the three cases separately and computing in each case the integral on the right-hand side of (6.11).

In a second lemma, we treat a bound on the probability that the vertices  $i_1, i_2, \dots, i_m$ , with  $m$  even, are normal vertices and are connected by an edge. In its statement, we use the event  $\mathcal{H}_j$  to denote that the vertex  $i_j$  is normal, i.e.  $\mathcal{H}_j = \{D_{i_j} < \varepsilon_n n^{1/(\tau-1)}\}$ .

**Lemma 6.6.** (Paths of normal vertices are unlikely.) *For every  $a > 0$ , there exists a constant  $C = C(a)$  such that, for all  $i_1, \dots, i_m \in [n]$  and  $m$  even,*

$$\mathbb{E} \left[ \prod_{j=1}^{m-1} \left( \frac{D_{i_j} D_{i_{j+1}}}{L_n} \wedge 1 \right) \mathbf{1}_{\mathcal{F}_m} \mathbf{1}_{\{L_n > an^{1/(\tau-1)}\}} \right] \leq \left( C \frac{\log n}{n} \right)^{m/2}, \tag{6.12}$$

where  $\mathcal{F}_m = \bigcap_{j=1}^m \mathcal{H}_j$  denotes the event that all vertices  $i_1, \dots, i_m$  are normal.

*Proof.* We start with an inequality for  $m = 2$ , where we abbreviate  $n^{1/(\tau-1)}$  by  $u_n$  so that  $u_n^{\tau-1} = n$ . We use the inequalities in (6.10) and the upper bound  $1 - F(x) \leq c_2 x^{1-\tau}$ . We have

$$\begin{aligned} \mathbb{E} \left[ \left( \frac{D_{i_1} D_{i_2}}{au_n} \wedge 1 \right) \mathbf{1}_{\mathcal{H}_1} \mid D_{i_1} \right] &= \mathbf{1}_{\mathcal{H}_1} \left\{ \mathbb{P} \left( D_{i_2} > \frac{au_n}{D_{i_1}} \mid D_{i_1} \right) \right. \\ &\quad \left. + \frac{D_{i_1}}{au_n} \mathbb{E} [D_{i_2} \mathbf{1}_{\{D_{i_2} \leq au_n/D_{i_1}\}} \mid D_{i_1}] \right\} \\ &\leq \mathbf{1}_{\mathcal{H}_1} \left\{ c_2 \left( \frac{au_n}{D_{i_1}} \right)^{1-\tau} + \frac{D_{i_1}}{au_n} C \left( \frac{au_n}{D_{i_1}} \right)^{2-\tau} \right\} \\ &\leq \mathbf{1}_{\mathcal{H}_1} \left\{ \tilde{C} \left( \frac{au_n}{D_{i_1}} \right)^{1-\tau} \right\}. \end{aligned}$$

Taking expectations on both sides, and using the fact that  $\mathcal{F}_2 \subset \mathcal{H}_1$ , yields, using again one of the inequalities in (6.10),

$$\begin{aligned} \mathbb{E} \left[ \left( \frac{D_{i_1} D_{i_2}}{au_n} \wedge 1 \right) \mathbf{1}_{\mathcal{F}_2} \right] &\leq \tilde{C} a^{1-\tau} n^{-1} \mathbb{E} [(D_{i_1})^{\tau-1} \mathbf{1}_{\mathcal{H}_1}] \\ &\leq \tilde{C} a^{1-\tau} n^{-1} C \log \{\varepsilon_n n^{1/(\tau-1)}\} \\ &\leq C(a) \frac{\log n}{n}. \end{aligned}$$

The bound for general even  $m \geq 4$  is verified as follows. We note that in (6.12) we take the expectation of the product of  $m - 1$  terms. We bound the second term, i.e.  $(D_{i_2} D_{i_3} / L_n \wedge 1)$ , by the obvious bound 1, and do the same with the fourth up to the  $(m - 2)$ th term. At the same time we bound  $\mathbf{1}_{\mathcal{F}_m}$  by  $\mathbf{1}_{\{\bigcap_{k=1}^{m/2} \mathcal{H}_{2k-1}\}}$ . After replacing  $L_n$  by  $au_n$ , we end up with an upper bound consisting of the independent product of  $m/2$  terms of the form

$$\mathbb{E} \left[ \left( \frac{D_{i_j} D_{i_{j+1}}}{au_n} \wedge 1 \right) \mathbf{1}_{\mathcal{H}_j} \right], \quad j = 1, 3, \dots, m - 1,$$

each of which can be bounded by  $C \log n/n$  as above.

The following lemma shows that paths of odd lengths are unlikely.

**Lemma 6.7.** (Shortest-weight paths on super vertices are of even length.) *Let the distribution function  $F$  of the degrees of the CM satisfy (3.3). Let  $\mathcal{B}^{(n)}$  be the event that there exists a path between two super vertices consisting of all normal vertices and having an odd number of edges with total weight  $w_n/\sqrt{n}$ . Then, for some constant  $C$ ,*

$$\mathbb{P}(\mathcal{B}^{(n)} \cap \{L_n > au_n\}) \leq C \frac{\varepsilon_n^{-2(\tau-1)}}{\sqrt{n \log n}} e^{C w_n \sqrt{\log n}}.$$

*Proof.* We will show that the probability that there exists a path between two super vertices consisting of all normal vertices and having an odd number of edges with total weight  $w_n/\sqrt{n}$  is small. For this, we use the first moment method and show that the expected number of such paths goes to 0 as  $n \rightarrow \infty$ . Fix two super vertices which will be the end points of the path and an even number,  $m \geq 0$ , of normal vertices with indices  $i_1, i_2, \dots, i_m$ . Note that when a path between two super vertices consists of an even number of vertices, then the path has an odd number of edges.

Let  $\mathcal{B}_m^{(n)}$  be the event that there exists a path between two super vertices consisting of exactly  $m$  intermediate normal vertices with total weight  $w_n/\sqrt{n}$ . We start by investigating the  $m = 0$  case, so that the super vertices are directly connected. Recall that  $\mathcal{S}_n$  denotes the set of super vertices, i.e. vertices whose degree is at least  $\varepsilon_n n^{1/(\tau-1)}$ . Note that the number of super vertices  $|\mathcal{S}_n|$  is a binomial random variable, which is concentrated as soon as its mean tends to  $\infty$ . Thus,  $|\mathcal{S}_n| = O_P(E[|\mathcal{S}_n|])$  since

$$E[|\mathcal{S}_n|] = n P(D_1 > \varepsilon_n n^{1/(\tau-1)}) = O(\varepsilon_n^{-(\tau-1)}) \rightarrow \infty.$$

Hence, there are  $O_P(\varepsilon_n^{-(\tau-1)})$  super vertices and, thus,  $O_P(\varepsilon_n^{-2(\tau-1)})$  edges between them. The probability that any one of them is smaller than  $w_n/\sqrt{n}$  is of order  $\varepsilon_n^{-2(\tau-1)} w_n/\sqrt{n}$ , and it follows that  $P(\mathcal{B}_0^{(n)}) \leq \varepsilon_n^{-2(\tau-1)} w_n/\sqrt{n}$ .

Let  $M_m^{(n)}$  be the total number of paths connecting two specific super vertices and which are such that the total weight on the paths is at most  $w_n/\sqrt{n}$ , so that

$$P(\mathcal{B}_m^{(n)}) = P(M_m^{(n)} \geq 1) \leq E[M_m^{(n)}].$$

In the following argument, for convenience, we let  $\{D_i\}_{i=1}^n$  denote the i.i.d. vector of degrees (i.e. below  $D_i$  is not the  $i$ th largest degree, but rather a copy of the random variable  $D \sim F$  independently of the other degrees).

Let  $\vec{i} = (i_1, i_2, \dots, i_m)$ , and denote by  $p_{m,n}(\vec{i})$  the probability that the  $m$  vertices  $i_1, i_2, \dots, i_m$  are normal and are such that there is an edge between  $i_s$  and  $i_{s+1}$  for  $s = 1, \dots, m - 1$ . Furthermore, note that, with  $S_{m+1} = \sum_{i=1}^{m+1} E_i$ , where the  $E_i$  are independent exponential random variables with rate 1, we have, for any  $u \in [0, 1]$ ,

$$P(S_{m+1} \leq u) = \int_0^u \frac{x^m e^{-x}}{m!} dx \leq \frac{u^{m+1}}{(m+1)!}. \tag{6.13}$$

Together with the fact that there are  $O_P(\varepsilon_n^{-(\tau-1)})$  super vertices, this implies that

$$P(\mathcal{B}_m^{(n)}) \leq E[M_m^{(n)}] \leq \frac{C E[|\mathcal{S}_n|^2] w_n^{m+1}}{(m+1)! n^{(m+1)/2}} \sum_{\vec{i}} p_{m,n}(\vec{i}) \leq \frac{C \varepsilon_n^{-2(\tau-1)} w_n^{m+1}}{(m+1)! n^{(m+1)/2}} \sum_{\vec{i}} p_{m,n}(\vec{i}), \tag{6.14}$$

since (6.13) implies that the probability that the sum of  $m + 1$  exponentially distributed random variables is smaller than  $u_n = w_n/\sqrt{n}$  is at most  $u_n^{m+1}/(m+1)!$ .

By the construction of the CM we have

$$p_{m,n}(\vec{i}) \leq E \left[ \prod_{j=1}^{m-1} \left( \frac{D_{i_j} D_{i_{j+1}}}{L_n - 2j + 1} \wedge 1 \right) \mathbf{1}_{\mathcal{F}_m} \right] \leq E \left[ \prod_{j=1}^{m-1} \left( \frac{D_{i_j} D_{i_{j+1}}}{L_n} \wedge 1 \right) \mathbf{1}_{\mathcal{F}_m} \right] (1 + o(1)),$$

where  $\mathcal{F}_m$  is the event that  $D_{i_j} < \varepsilon_n n^{1/(\tau-1)}$  for all  $1 \leq j \leq m$ . It follows from Lemma 6.6 that on the event  $L_n > au_n$ , and, for even  $m$ ,

$$p_{m,n}(\vec{i}) \leq \frac{(C \log n)^{m/2}}{n^{m/2}}.$$

Using this estimate in (6.14), and summing over all even  $m$ , using the notation that  $m \in 2\mathbb{Z}^+$ , shows that

$$\begin{aligned}
 P(\mathcal{B}^{(n)} \cap \{L_n > au_n\}) &\leq \sum_{m \in 2\mathbb{Z}^+} P(\mathcal{B}_m^{(n)} \cap \{L_n > au_n\}) \\
 &\leq \sum_{m \in 2\mathbb{Z}^+} \frac{C\varepsilon_n^{-2(\tau-1)}w_n^{m+1}}{(m+1)!n^{(m+1)/2}}(Cn \log n)^{m/2} \\
 &= C\varepsilon_n^{-2(\tau-1)}n^{-1/2} \sum_{k=1}^{\infty} \frac{w_n^{2k+1}(C \log n)^k}{(2k+1)!} \\
 &\leq C \frac{\varepsilon_n^{-2(\tau-1)}}{n^{1/2}\sqrt{\log n}} e^{Cw_n\sqrt{\log n}}.
 \end{aligned}
 \tag{6.15}$$

Since the event  $\{L_n > au_n\}$  happens with probability  $1 - o(1)$  when  $a \downarrow 0$ , Lemma 6.7, with the correct choice of  $\varepsilon_n$ , shows that  $P(H_n \notin 2\mathbb{Z}^+) = o(1)$ . In order to prove Theorem 3.2, we further show that the shortest-weight paths between any two specific super vertices alternate between super vertices and normal vertices. We will prove this statement in Lemma 6.9 below, by showing that the probability that a vertex with index at least  $K$  is used at an even place is, for large  $K$ , quite small. This shows in particular that, w.h.p., at all even places we have super vertices. In the following lemma, we collect the properties of the degrees and erased degrees that we will make use of in the sequel. In its statement, we define

$$\mathcal{G}^{(n)} = \mathcal{G}_1^{(n)} \cap \mathcal{G}_2^{(n)} \cap \mathcal{G}_3^{(n)},$$

where, for  $a \in (0, 1)$  and  $C, C^{\text{er}} > 0$ , we let

$$\begin{aligned}
 \mathcal{G}_1^{(n)} &= \{L_n n^{-1/(\tau-1)} \in [a, a^{-1}]\}, \\
 \mathcal{G}_2^{(n)} &= \left\{ C^{-1} \left(\frac{n}{i}\right)^{1/(\tau-1)} \leq D_{(n+1-i:n)} \leq C \left(\frac{n}{i}\right)^{1/(\tau-1)} \text{ for all } i \in [n] \right\}, \\
 \mathcal{G}_3^{(n)} &= \left\{ D_i^{\text{er}} \leq C^{\text{er}} \left(\frac{n}{i}\right) \text{ for all } i \in [n] \right\}.
 \end{aligned}$$

The event  $\mathcal{G}^{(n)}$  is the *good event* that we will work with. We will first show that, if we take  $a > 0$  sufficiently small and  $C, C^{\text{er}}$  sufficiently large, then  $P(\mathcal{G}^{(n)})$  is close to 1.

**Lemma 6.8.** (The good event has high probability.) *For every  $\varepsilon > 0$ , there exist  $a > 0$  sufficiently small and  $C, C^{\text{er}}$  sufficiently large such that*

$$P(\mathcal{G}^{(n)}) \geq 1 - \varepsilon.$$

*Proof.* We split

$$P((\mathcal{G}^{(n)})^c) = P((\mathcal{G}_1^{(n)})^c) + P(\mathcal{G}_1^{(n)} \cap (\mathcal{G}_2^{(n)})^c) + P(\mathcal{G}_1^{(n)} \cap \mathcal{G}_2^{(n)} \cap (\mathcal{G}_3^{(n)})^c),$$

and bound each term separately. We can make  $P((\mathcal{G}_1^{(n)})^c) \leq \varepsilon/3$  by choosing  $a > 0$  sufficiently small by the weak convergence in (2.5).

To bound  $P(\mathcal{G}_1^{(n)} \cap (\mathcal{G}_2^{(n)})^c)$ , we note that  $D_{(n+1-l:n)} > C(n/l)^{1/(\tau-1)}$  is equivalent to the statement that the number of values  $i$  such that  $D_i > C(n/l)^{1/(\tau-1)}$  is at least  $l$ . Since  $\{D_i\}_{i=1}^n$

is an i.i.d. sequence, this number has a binomial distribution with parameters  $n$  and success probability

$$q_{l,n} = \left[ 1 - F\left(C\left(\frac{n}{l}\right)^{1/(\tau-1)}\right) \right] \leq \frac{c_2 C^{-(\tau-1)} l}{n},$$

by (3.3). When the mean of this binomial, which is  $c_2 C^{-(\tau-1)} l$ , is much smaller than  $l$ , which is equivalent to  $C > 0$  being large, the probability that this binomial exceeds  $l$  is exponentially small in  $l$ :

$$P\left(D_{(n+1-l:n)} > C\left(\frac{n}{l}\right)^{1/(\tau-1)}\right) \leq e^{-I(C)l},$$

where  $I(C) \rightarrow \infty$  when  $C \rightarrow \infty$ . Thus, by taking  $C$  sufficiently large, we can make the probability that there exists an  $l$  for which  $D_{(n+1-l:n)} > C(n/l)^{1/(\tau-1)}$  small. In more detail,

$$\begin{aligned} P\left(\text{there exists } l: D_{(n+1-l:n)} > C\left(\frac{n}{l}\right)^{1/(\tau-1)}\right) &\leq \sum_{l \in [n]} P\left(D_{(n+1-l:n)} > C\left(\frac{n}{l}\right)^{1/(\tau-1)}\right) \\ &\leq \sum_{l \in [n]} e^{-I(C)l} \\ &\leq \frac{\varepsilon}{3}, \end{aligned}$$

when we make  $C > 0$  sufficiently large. In a similar way, we can show that the probability that there exists  $l$  such that  $D_{(n+1-l:n)} \leq C^{-1}(n/l)^{1/(\tau-1)}$  is small when  $C > 0$  is large.

In order to bound  $P(\mathcal{G}_1^{(n)} \cap \mathcal{G}_2^{(n)} \cap (\mathcal{G}_3^{(n)})^c)$ , we need to investigate the random variable  $D_i^{\text{er}}$ . We claim that there exists an  $R = R(a, C, C^{\text{er}})$  with  $R(a, C, C^{\text{er}}) \rightarrow \infty$  as  $C^{\text{er}} \rightarrow \infty$  for each fixed  $a, C > 0$ , such that

$$P(D_i^{\text{er}} \geq C^{\text{er}} j^{\tau-1} \mid D_i = j, \mathcal{G}_1^{(n)} \cap \mathcal{G}_2^{(n)}) \leq e^{-Rj^{\tau-1}}. \tag{6.16}$$

Fix  $C^{\text{er}} > 0$ . In order for  $D_i^{\text{er}} \geq C^{\text{er}} j^{\tau-1}$  to occur, we must have at least  $C^{\text{er}} j^{\tau-1}/2$  of the neighbors of vertex  $i$  with index at least  $C^{\text{er}} j^{\tau-1}/2$ , where we recall that vertex  $i$  is such that  $D_i = D_{(n+1-i:n)}$  is the  $i$ th largest degree. The  $j$  neighbors of vertex  $i$  are close to being independent, and the probability that any of them connects to a vertex with index at least  $k$  is, conditionally on the degrees  $\{D_i\}_{i=1}^n$ , equal to

$$\sum_{l \geq k} \frac{D_{(n+1-l:n)}}{L_n}.$$

When  $\mathcal{G}_1^{(n)} \cap \mathcal{G}_2^{(n)}$  holds, then  $D_{(n+1-l:n)}/L_n \leq (C/a)l^{-1/(\tau-1)}$ , so that

$$\sum_{l \geq k} \frac{D_{(n+1-l:n)}}{L_n} \leq c\left(\frac{C}{a}\right)k^{-(2-\tau)/(\tau-1)}.$$

As a result, we can bound the number of neighbors of vertex  $i$  by a binomial random variable with  $p = c'k^{-(2-\tau)/(\tau-1)}$ , where  $k = C^{\text{er}} j^{\tau-1}/2$ , i.e.

$$P(D_i^{\text{er}} \geq C^{\text{er}} j^{\tau-1} \mid D_i = j, \mathcal{G}_1^{(n)} \cap \mathcal{G}_2^{(n)}) \leq P\left(\text{Bin}\left(j, c\left(\frac{C}{a}\right)j^{-(2-\tau)}\right) \geq \frac{C^{\text{er}} j^{\tau-1}}{2}\right).$$

Next, we note that the mean of the above binomial random variable is given by

$$c \left( \frac{C}{a} \right) j^{1-(2-\tau)} = c \left( \frac{C}{a} \right) j^{\tau-1}.$$

A concentration result for binomial random variables [14], yields, for sufficiently large  $C > 0$ ,

$$P(D_i^{\text{er}} \geq C^{\text{er}} j^{\tau-1} \mid D_i = j, \mathcal{G}_1^{(n)} \cap \mathcal{G}_2^{(n)}) \leq e^{-Rj^{\tau-1}}.$$

This proves (6.16). Taking  $C^{\text{er}} > 0$  sufficiently large, we obtain

$$\begin{aligned} & P(\mathcal{G}_1^{(n)} \cap \mathcal{G}_2^{(n)} \cap (\mathcal{G}_3^{(n)})^c) \\ & \leq \sum_{i=1}^n \sum_j P(D_i^{\text{er}} \geq C^{\text{er}} j^{\tau-1} \mid D_i = j, \mathcal{G}_1^{(n)} \cap \mathcal{G}_2^{(n)}) P(D_i = j, \mathcal{G}_1^{(n)} \cap \mathcal{G}_2^{(n)}) \\ & \leq \sum_{i=1}^n e^{-RCn/i} \\ & \leq \frac{\varepsilon}{3}, \end{aligned}$$

where we used the fact that  $D_i \geq C^{-1}(n/i)^{1/(\tau-1)}$  since the event  $\mathcal{G}_2^{(n)}$  occurs, and where, in the last step, we used the fact that, for each  $a, C > 0$ , we can make  $R = R(a, C, C^{\text{er}})$  large by taking  $C^{\text{er}}$  sufficiently large.

Now we are ready to prove the tightness of the FPP problem. In the statement below, we let  $\mathcal{A}_{m,K}^{(n)}(i, j)$  be the event that there exists a path of length  $2m$  connecting  $i$  and  $j$  of weight at most  $W/\sqrt{n}$  that leaves  $[K]$ , and we write, for fixed  $k$ ,

$$\mathcal{A}_{m,K}^{(n)} = \bigcup_{i,j \in [k]} \mathcal{A}_{m,K}^{(n)}(i, j), \quad \mathcal{A}_K^{(n)} = \bigcup_{m=1}^{\infty} \mathcal{A}_{m,K}^{(n)}.$$

**Lemma 6.9.** (Tightness of even shortest-weight paths on the super vertices.) *Fix  $k, K \in \mathbb{Z}^+$  and  $i, j \in [k]$ . Then, there exists a  $C > 0$  such that*

$$P(\mathcal{A}_{m,K}^{(n)} \cap \mathcal{G}^{(n)}) \leq CWK^{-1} e^{CW\sqrt{\log K}}.$$

*Proof.* We follow the same line of argument as in the proof of Lemma 6.7, but we need to be more careful in estimating the expected number of paths between the super vertices  $i$  and  $j$ . For  $m \geq 2$  and  $\vec{l} = (i_1, \dots, i_{m-1})$ , let  $q_{m,n}(\vec{l})$  be the expected number of paths with  $2m$  edges ( $2m$  step paths) such that the position of the path at time  $2k$  is equal to  $i_k$ , where, by convention,  $i_0 = i$  and  $i_m = j$ . Then, similarly as in (6.14), but note that now  $q_{m,n}(\vec{l})$  is an expectation and not a probability, we have

$$P(\mathcal{A}_{m,K}^{(n)} \cap \mathcal{G}^{(n)}) \leq \frac{CW^{2m}}{(2m)! n^m} \sum_{\vec{l}} q_{m,n}(\vec{l}).$$

Observe that

$$q_{m,n}(\vec{l}) \leq E \left[ \prod_{s=1}^m N_{i_{s-1}i_s}^{\text{er}} \mathbf{1}_{\mathcal{G}^{(n)}} \right]. \tag{6.17}$$

We further note that, by Lemma 6.8 and on  $\mathcal{G}^{(n)}$ ,

$$N_{ij} \leq D_i^{\text{er}} \wedge D_j^{\text{er}} \leq D_{i \vee j}^{\text{er}} \leq \frac{C^{\text{er}} n}{i \vee j},$$

where we abbreviate, for  $x, y \in \mathbb{R}$ ,  $x \wedge y = \min\{x, y\}$  and  $x \vee y = \max\{x, y\}$ . Thus, by (6.17) we arrive at

$$q_{m,n}(\vec{i}) \leq \prod_{s=1}^m \frac{C^{\text{er}} n}{i_s \vee i_{s-1}},$$

and, hence, after possibly enlarging  $C^{\text{er}}$ ,

$$P(\mathcal{A}_{m,K}^{(n)} \cap \mathcal{G}^{(n)}) \leq \frac{CW^{2m}}{(2m)!n^m} \sum_{\vec{i}} \prod_{s=1}^m \frac{C^{\text{er}} n}{i_s \vee i_{s-1}} = \frac{(C^{\text{er}})^m W^{2m}}{(2m)!} \sum_{\vec{i}} \prod_{s=1}^m \frac{1}{i_s \vee i_{s-1}}, \tag{6.18}$$

where the sum over  $\vec{i}$  is such that there exists at least one  $s$  such that  $i_s > K$ , because the path is assumed to leave  $[K]$ . We now bound (6.18). Let  $1 \leq t \leq m$  be such that  $i_t = \max_{s=1}^m i_s$ , so that  $i_t > K$ . Then, using the fact that both  $i_s$  and  $i_{s-1}$  are smaller than  $i_s \vee i_{s-1}$ , we can bound

$$\begin{aligned} \prod_{s=1}^m \frac{1}{i_s \vee i_{s-1}} &= \left( \prod_{s=1}^{t-1} \frac{1}{i_s \vee i_{s-1}} \right) \frac{1}{i_{t-1} \vee i_t} \frac{1}{i_t \vee i_{t+1}} \left( \prod_{s=t+2}^m \frac{1}{i_s \vee i_{s-1}} \right) \\ &\leq \frac{1}{i_t^2} \prod_{s=1}^{t-1} \frac{1}{i_s} \prod_{s=t+2}^m \frac{1}{i_{s-1}}. \end{aligned}$$

Thus,

$$\begin{aligned} \sum_{\vec{i}} \prod_{s=1}^m \frac{1}{i_s \vee i_{s-1}} &= \sum_{t=1}^m \sum_{i_t > K} \frac{1}{i_t^2} \sum_{i_1, \dots, i_{t-1} \leq i_t} \prod_{s=1}^{t-1} \frac{1}{i_s} \sum_{i_{t+1}, \dots, i_m \leq i_t} \prod_{s=t+2}^m \frac{1}{i_{s-1}} \\ &= m \sum_{u > K} \frac{1}{u^2} h_u^{m-2}, \end{aligned}$$

where  $h_u = \sum_{v=1}^u 1/v$ . We arrive at

$$P(\mathcal{A}_{m,K}^{(n)} \cap \mathcal{G}^{(n)}) \leq \frac{(C^{\text{er}})^m W^{2m}}{(2m)!} m \sum_{u > K} \frac{1}{u^2} h_u^{m-2}. \tag{6.19}$$

By Boole’s inequality and (6.19), we obtain, after replacing  $C^{\text{er}}$  by  $C$ ,

$$\begin{aligned} P(\mathcal{A}_K^{(n)} \cap \mathcal{G}^{(n)}) &\leq \sum_{m=2}^{\infty} P(\mathcal{A}_{m,K}^{(n)} \cap \mathcal{G}^{(n)}) \\ &\leq \sum_{m=2}^{\infty} \frac{C^m W^{2m}}{(2m)!} m \sum_{u > K} \frac{1}{u^2} h_u^{m-2} \\ &\leq W \sum_{u > K} \frac{1}{u^2} \sum_{m=2}^{\infty} h_u^{(2m-1)/2} \frac{C^{2m-1} W^{2m-1}}{(2m-1)!} \\ &\leq CW \sum_{u > K} \frac{1}{u^2} e^{CW\sqrt{h_u}} \\ &\leq CWK^{-1} e^{CW\sqrt{\log K}}, \end{aligned}$$

where we used the fact that

$$\begin{aligned} \sum_{u>K} \frac{1}{u^2} e^{CW\sqrt{h_u}} &\leq \int_{\log K}^\infty e^{-y+CW\sqrt{y+c}} dy \\ &\leq \int_{\log K}^\infty \exp\left\{-y\left(1 - \frac{2CW}{\sqrt{\log K}}\right)\right\} dy \\ &\leq K^{-1} e^{C'W\sqrt{\log K}} \end{aligned}$$

for some  $c, C' > 0$ . This completes the proof of Lemma 6.9.

Now we are ready to complete the proof of Proposition 6.3.

*Proof of Proposition 6.3.* We write  $H_n(i, j)$  and  $W_n(i, j)$  for the number of edges and weight of the shortest-weight path between the super vertices  $i, j \in [k]$ .

(a) The fact that  $P(H_n(i, j) \notin 2\mathbb{Z}^+) = o(1)$  for any super vertices  $i, j$ , follows immediately from Lemma 6.7, which implies that the even length path between  $i$  and  $j$  is a two-edge path w.h.p. The tightness of  $H_n(i, j)$  follows from part (b), which we prove next.

(b) By Proposition 6.1, the rescaled weight  $\sqrt{n} W_n(i, j) \leq l_{ij}^{(n)}$  is a tight sequence of random variables, so that, for large  $W$ , it is at most  $W$  with probability converging to 1 when  $W \rightarrow \infty$ . Fix  $\varepsilon > 0$  arbitrarily. Then, fix  $W > 0$  sufficiently large such that the probability that  $\sqrt{n} W_n(i, j) > W$  is at most  $\varepsilon/3$ ,  $K > 0$  such that  $CWK^{-1} e^{CW\sqrt{\log K}} < \varepsilon/3$ , and, use Lemma 6.8 to see that we can choose  $a, C$ , and  $C^{er}$  such that  $P(\mathcal{G}^{(n)}) \geq 1 - \varepsilon/3$ . Then, by Lemma 6.9, the probability that this two-edge path leaves  $[K]$  is at most  $\varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon$ . This completes the proof of part (b).

(c) The proof that  $P(H_n \notin 2\mathbb{Z}^+) = o(1)$  follows from part (a) since, for large  $k$ , w.h.p.,  $A_1$  and  $A_2$  are exclusively attached to super vertices in  $[k]$ . The tightness of  $H_n$  also follows from this argument and part (a).

**6.5. The FPP on  $\mathcal{K}_\infty^{er}$  is well defined: proof of Proposition 6.4**

In this section we prove Proposition 6.4. For this, we start by investigating  $f(P_i, P_j)$  for large  $i, j$ . The main result is contained in the following lemma.

**Lemma 6.10.** (Asymptotics for  $f(P_i, P_j)$  for large  $i, j$ .) *Let  $\eta$  be a stable random variable with parameter  $\tau - 1 \in (0, 1)$ . Then, there exists a constant  $c > 0$  such that, as  $i \wedge j \rightarrow \infty$ ,*

$$f(P_i, P_j) \leq \frac{c\eta^{1-\tau}}{i \vee j} \quad a.s. \tag{6.20}$$

*Proof.* We note that, by (2.2), (2.4), and the strong law of large numbers, as  $i \rightarrow \infty$ ,

$$\eta P_i i^{1/(\tau-1)} = (i / \Gamma_i)^{1/(\tau-1)} \xrightarrow{a.s.} 1.$$

Furthermore, by (3.4),

$$f(s, t) = 1 - E[(1 - s)^D] - E[(1 - t)^D] + E[(1 - s - t)^D] \leq 1 - E[(1 - s)^D] \leq cs^{\tau-1},$$

since, for  $\alpha = \tau - 1$ , and with  $D$  in the domain of attraction of an  $\alpha$ -stable random variable, we have, as  $u \downarrow 0$ ,

$$E[e^{-uD}] = e^{-cu^\alpha(1+o(1))} = 1 - cu^\alpha(1 + o(1)).$$

Combing these asymptotics proves (6.20).

*Proof of Proposition 6.4.* Let  $\mathcal{A}_{m,K}$  be the event that there exists a path of length  $m$  and weight at most  $W$  connecting  $i$  and  $j$  and which contains a vertex in  $\mathbb{Z}^+ \setminus [K]$ . Then, by Boole’s inequality and the conditional independence of the weights  $\{l_{ij}\}_{1 \leq i < j < \infty}$ , we obtain

$$P(\mathcal{A}_{m,K}) \leq \sum_{\vec{i}} P\left(\sum_{s=1}^m l_{i_{s-1}i_s} \leq W\right),$$

where, as in the proof of Lemma 6.9, the sum over  $\vec{i}$  is over  $\vec{i} = (i_1, \dots, i_{m-1})$ , where, by convention,  $i_0 = i$  and  $i_m = j$ , and  $\max_{s=1}^m i_s \geq K$ . Now, by the conditional independence of  $\{l_{ij}\}_{1 \leq i < j < \infty}$ ,

$$\begin{aligned} &P\left(\sum_{s=1}^m l_{i_{s-1}i_s} \leq W \mid \{P_i\}_{i \geq 1}\right) \\ &= \int_{x_1 + \dots + x_m \leq W} \prod_{s=1}^m f(P_{i_{s-1}}, P_{i_s}) x_s e^{-f(P_{i_{s-1}}, P_{i_s}) x_s^2 / 2} dx_1 \dots dx_m \\ &\leq \prod_{s=1}^m f(P_{i_{s-1}}, P_{i_s}) \int_{x_1 + \dots + x_m \leq W} x_1 \dots x_m dx_1 \dots dx_m \\ &= \frac{W^{2m}}{(2m)!} \prod_{s=1}^m f(P_{i_{s-1}}, P_{i_s}), \end{aligned}$$

by [10, Equation 4.634]. We have  $f(P_i, P_j) \leq c\eta^{1-\tau}(i \vee j)^{-1}$  a.s., by Lemma 6.10. The random variable  $\eta$  has a stable distribution, and is therefore bounded below by  $a$  with probability  $1 - o(1)$  when  $a \rightarrow 0$ . The arising bound is identical to bound (6.18) derived in the proof of Lemma 6.9, and we can follow the proof to obtain (6.4).

### 7. Robustness and fragility: proof of Theorem 3.3

We start by proving Theorem 3.3(a), for which we note that whatever the value of  $p \in (0, 1)$ , w.h.p., not all super vertices will be deleted. The number of undeleted vertices that are connected to a kept super vertex will be  $\Theta_p(n)$ , which proves the claim. In fact, we now argue that a stronger result holds. We note that the size of the giant component is the same whether we consider  $\mathcal{G}_n^{\text{er}}$  or  $\mathcal{G}_n^{\text{or}}$ . It is easy to prove the following result.

**Theorem 7.1.** (Giant component after random attack.) *Consider either  $\mathcal{G}_n^{\text{or}}$  or  $\mathcal{G}_n^{\text{er}}$  and leave each vertex with probability  $p$  or delete it with probability  $(1 - p)$ . The resulting graph (of vertices which are left) has a unique giant component  $\mathcal{C}_n(p)$ . Furthermore, with  $|\mathcal{C}_n(p)|$  denoting the number of vertices in  $\mathcal{C}_n(p)$ ,*

$$\frac{E[|\mathcal{C}_n(p)|]}{n} \longrightarrow p E[1 - (1 - p)^{D^{\text{er}}}] = \lambda(p), \quad \text{var}\left(\frac{|\mathcal{C}_n(p)|}{n}\right) \rightarrow \beta(p) > 0. \quad (7.1)$$

Unlike for other random graph models, (7.1) suggests that  $|\mathcal{C}_n(p)|/n \xrightarrow{D} Z_p$ , where  $Z_p$  is a *nondegenerate* random variable. We will however not attempt to prove the latter statement here.

*Sketch of the proof of Theorem 7.1.* Note that we have the identity

$$\frac{E[|\mathcal{C}_n(p)|]}{n} = P(1 \in \mathcal{C}_n(p)),$$

where 1 is a uniformly chosen vertex in  $\mathcal{G}_n^{\text{er}}$ . For large  $n$ , the vertex 1 being in the giant component is *essentially* equivalent to the following two conditions.

- (i) Vertex 1 is not deleted; this happens with probability  $p$ .
- (ii) Vertex 1 is attached to  $D^{\text{er}}$  super vertices. If one of those super vertices is not deleted then the component of this super vertex is of order  $n$  and thus has to be the giant component. Thus, at least one of the super vertices to which 1 is attached should remain undeleted; conditionally on  $D^{\text{er}}$ , this happens with probability  $1 - (1 - p)^{D^{\text{er}}}$ .

Combining (i) and (ii) gives the result. A calculation, using similar ideas as in the proof of Lemma 6.8, suggests that  $\lambda(p) = \Theta(p^2)$  when  $p \downarrow 0$ . Furthermore, the giant component is unique, since any pair of super vertices which are kept are connected to each other, and are each connected to  $\Theta_p(n)$  other vertices.

To prove the convergence of the variance, we note that

$$\text{var}(|\mathcal{C}_n(p)|) = \sum_{i,j} [\text{P}(i, j \in \mathcal{C}_n(p)) - \text{P}(i \in \mathcal{C}_n(p)) \text{P}(j \in \mathcal{C}_n(p))].$$

Thus,

$$\text{var}\left(\frac{|\mathcal{C}_n(p)|}{n}\right) = \text{P}(1, 2 \in \mathcal{C}_n(p)) - \text{P}(1 \in \mathcal{C}_n(p)) \text{P}(2 \in \mathcal{C}_n(p)),$$

where 1 and 2 are two independent uniform vertices in  $[n]$ . Now,

$$\text{P}(1, 2 \in \mathcal{C}_n(p)) = p^2 \text{P}(1, 2 \in \mathcal{C}_n(p) \mid 1, 2 \text{ kept}) + o(1),$$

and

$$\begin{aligned} \text{P}(1, 2 \in \mathcal{C}_n(p) \mid 1, 2 \text{ kept}) &= 1 - \text{P}(\{1 \notin \mathcal{C}_n(p)\} \cup \{2 \notin \mathcal{C}_n(p)\} \mid 1, 2 \text{ kept}) \\ &= 1 - \text{P}(1 \notin \mathcal{C}_n(p) \mid 1, 2 \text{ kept}) - \text{P}(2 \notin \mathcal{C}_n(p) \mid 1, 2 \text{ kept}) \\ &\quad + \text{P}(1, 2 \notin \mathcal{C}_n(p) \mid 1, 2 \text{ kept}) \\ &= 1 - \text{P}(1 \notin \mathcal{C}_n(p) \mid 1 \text{ kept}) - \text{P}(2 \notin \mathcal{C}_n(p) \mid 2 \text{ kept}) \\ &\quad + \text{P}(1, 2 \notin \mathcal{C}_n(p) \mid 1, 2 \text{ kept}) + o(1), \end{aligned}$$

so that

$$\begin{aligned} \text{var}\left(\frac{|\mathcal{C}_n(p)|}{n}\right) &= p^2 \text{P}(1, 2 \notin \mathcal{C}_n(p) \mid 1, 2 \text{ kept}) \\ &\quad - p^2 \text{P}(1 \notin \mathcal{C}_n(p) \mid 1 \text{ kept}) \text{P}(2 \notin \mathcal{C}_n(p) \mid 2 \text{ kept}) + o(1). \end{aligned}$$

Then, we compute that

$$\text{P}(1 \notin \mathcal{C}_n(p) \mid 1 \text{ kept}) = \text{E}[(1 - p)^{D_1^{\text{er}}}],$$

while

$$\text{P}(1, 2 \notin \mathcal{C}_n(p) \mid 1, 2 \text{ kept}) = \text{E}[(1 - p)^{D_1^{\text{er}} + D_2^{\text{er}} - N_{12}^{\text{er}}}],$$

where  $D_1^{\text{er}}$  and  $D_2^{\text{er}}$  are conditionally independent given  $\{P_i\}_{i \geq 1}$ , and  $N_{12}^{\text{er}}$  denotes the number of joint neighbors of 1 and 2, and where we used the fact that the total number of super vertices to which 1 and 2 are connected is equal to  $D_1^{\text{er}} + D_2^{\text{er}} - N_{12}^{\text{er}}$ . As a result,

$$\text{var}\left(\frac{|\mathcal{C}_n(p)|}{n}\right) = p^2 (\text{E}[(1 - p)^{D_1^{\text{er}} + D_2^{\text{er}} - N_{12}^{\text{er}}}] - \text{E}[(1 - p)^{D_1^{\text{er}}}] \text{E}[(1 - p)^{D_2^{\text{er}}}] ) + o(1),$$

which identifies

$$\beta(p) = p^2(E[(1 - p)^{D_1^{er} + D_2^{er} - N_{12}^{er}}] - E[(1 - p)^{D_1^{er}}]E[(1 - p)^{D_2^{er}}]).$$

To see that  $\beta(p) > 0$ , we note that  $N_{12}^{er} > 0$  with positive probability, so that

$$\begin{aligned} E[(1 - p)^{D_1^{er} + D_2^{er} - N_{12}^{er}}] &> E[(1 - p)^{D_1^{er} + D_2^{er}}] \\ &= E[E[(1 - p)^{D_1^{er} + D_2^{er}} \mid \{P_i\}_{i \geq 1}]] \\ &= E[E[(1 - p)^{D_1^{er}} \mid \{P_i\}_{i \geq 1}]^2], \end{aligned}$$

by the conditional independence of  $D_1^{er}$  and  $D_2^{er}$ . Thus,  $\beta(p) > 0$  by the Cauchy–Schwarz inequality, as claimed.

To prove Theorem 3.3(b), we again use the fact that a uniform vertex is, w.h.p., only connected to a super vertex. Thus, there exists  $K_\epsilon$  such that by deleting the  $K_\epsilon$  vertices with largest degree we isolate  $A_1$  with probability at least  $\epsilon$ . This proves (3.9).

### 8. Conclusion

We conclude with a discussion about various extensions of the above results together with some further results without proof. Throughout the discussion, we use  $\mathcal{G}_n$  to denote either of  $\mathcal{G}_n^{or}$  and  $\mathcal{G}_n^{er}$ , where the choice depends on the context under consideration.

*Load distribution.* Understanding how random disorder changes the geometry of the network is crucial for understanding asymptotics of more complicated constructs such as the load distribution. More precisely, for any pair of vertices  $i, j \in \mathcal{G}_n$ , let  $\pi(i, j)$  denote the minimal-weight path between the two vertices. For any vertex  $v \in \mathcal{G}_n$ , the *load* on the vertex is defined as

$$L_n(v) = \sum_{i \neq j} \mathbf{1}_{\{v \in \pi(i, j)\}}.$$

For any fixed  $x$ , define the function  $G_n(x)$  as

$$G_n(x) = \#\{v : L_n(v) > x\}.$$

Understanding such functions is paramount to understanding the flow carrying properties of the network and are essential for the study of *betweenness centrality* of vertices in a network. For example, in social networks such measures are used to rate the relative importance of various individuals in the network, while in data networks such as the World Wide Web, such measures are used to rank the relative importance of web pages. An actual theoretical analysis of such questions is important, but seems difficult in many relevant situations. It would be of interest to find asymptotics of such functions in terms of the infinite objects  $\mathcal{K}_\infty^{or}$  and  $\mathcal{K}_\infty^{er}$  constructed in this paper. See also [3] for an analysis of such questions in the mean-field setting.

*Universality for edge weights.* In this study, to avoid technical complications, we assumed that each edge weight in  $\mathcal{G}_n$  has an exponential distribution. One natural question is: how far do these results depend on this assumption? It is well known in probabilistic combinatorial optimization that in a wide variety of contexts, when considering problems such as those in this paper, the actual distribution of the edge weights is not that important, what is important is the value of the density at 0. More precisely, consider  $\mathcal{G}_n^{er}$  (i.e. the erased CM) where each edge is given an i.i.d. edge weight having a continuous distribution with density  $g$  and let

$g(0) = \zeta \in (0, \infty)$ . Similar to  $\mathcal{K}_\infty^{\text{er}}$  defined in Section 3.2, define  $\mathcal{K}_\infty^{\text{er}}(\zeta)$  to be the infinite graph on the vertex set  $\mathbb{Z}^+$  where each edge  $l_{ij}$  has the distribution

$$P(l_{ij} > x) = \exp\left(-\frac{1}{2}f(P_i, P_j)\zeta^2 x^2\right). \quad (8.1)$$

Equation (8.1) can be proved along similar lines as in the proof of Lemma 6.1, and we leave this to the reader.

Let  $I^{\text{er}}$  be as defined in Section 3.2. Then we have the following modification of Theorem 3.2 which can be proved along the same lines.

**Theorem 8.1.** (Extension to other densities.) *Theorem 3.2 continues to hold with the modification that the quantities  $W_{ij}^{\text{er}}$  and  $H_{ij}^{\text{er}}$  arising in the limits are replaced by the corresponding quantities in  $\mathcal{K}_\infty^{\text{er}}(\zeta)$  instead of  $\mathcal{K}_\infty^{\text{er}}$ ,  $V_i^{\text{er}}$  is distributed as the minimum of  $D^{\text{er}}$  random variables having density  $g$ , while the distributions of  $I^{\text{er}}$  and  $J^{\text{er}}$  remain unchanged.*

A more challenging extension would be to densities for which either  $g(0) = 0$ , or for which  $\lim_{x \downarrow 0} g(x) = \infty$ . In this case, we believe the behavior to be entirely different from the one in Theorems 3.2 and 8.1, and it would be of interest to investigate whether a similar limiting FPP process arises.

### Acknowledgements

The research of SB was supported by NSF grant DMS 0704159, NSERC, PIMS Canada, and a UNC research council grant. SB would like to thank the hospitality of Eurandom where much of this work was done. The work of RvdH was supported in part by the Netherlands Organisation for Scientific Research (NWO). We thank the anonymous referee for closely reading the manuscript and suggesting changes that have made the paper more readable.

### References

- [1] ABRAMOWITZ, M. AND STEGUN, I. A. (eds) (1992). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. (Reprint of the 1972 edition.) Dover Publications, New York.
- [2] ALBERT, R., JEONG, H. AND BARABÁSI, A.-L. (2000). Error and attack tolerance of complex networks. *Nature* **406**, 378–382.
- [3] ALDOUS, D. J. AND BHAMIDI, S. (2010). Flows through random networks. To appear in *Random Structures and Algorithms*.
- [4] BENDER, E. A. AND CANFIELD, E. R. (1978). The asymptotic number of labeled graphs with given degree sequences. *J. Combinatorial Theory A* **24**, 296–307.
- [5] BHAMIDI, S. (2008). First passage percolation on locally treelike networks. I. Dense random graphs. *J. Math. Phys.* **49**, 125218, 27pp.
- [6] BHAMIDI, S., VAN DER HOFSTAD, R. AND HOOGHIEMSTRA, G. (2010). First passage percolation on sparse random graphs with finite mean degrees. To appear in *Ann. Appl. Prob.*
- [7] BOLLOBÁS, B. AND RIORDAN, O. (2004). Robustness and vulnerability of scale-free random graphs. *Internet Math.* **1**, 1–35.
- [8] DURRETT, R. (1988). *Lecture Notes on Particle Systems and Percolation*. Wadsworth & Brooks/Cole, Pacific Grove, CA.
- [9] DURRETT, R. (2007). *Random Graph Dynamics*. Cambridge University Press.
- [10] GRADSHTEYN, I. S. AND RYZHIK, I. M. (1965). *Table of Integrals, Series, and Products*, 4th edn. Academic Press, New York.
- [11] HAMMERSLEY, J. M. AND WELSH, D. J. A. (1965). First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory. In *Proc. Internat. Res. Semin.* (Statist. Lab., University of California, Berkeley), Springer, New York, pp. 61–110.
- [12] HOWARD, C. D. (2004). Models of first-passage percolation. In *Probability on Discrete Structures*, Springer, Berlin, pp. 125–173.

- [13] JANSON, S. (1999). One, two and three times  $\log n/n$  for paths in a complete graph with random weights. *Combin. Prob. Comput.* **8**, 347–361.
- [14] JANSON, S. (2002). On concentration of probability. In *Contemporary Combinatorics* (Bolyai Soc. Math. Stud. **10**), ed. B. Bollobás, János Bolyai Mathematical Society, Budapest, pp. 289–301.
- [15] KALLENBERG, O. (2002). *Foundations of Modern Probability*, 2nd edn. Springer, New York.
- [16] MEYERS, L. A., NEWMAN, M. E. J. AND POURBOHLOUL, B. (2006). Predicting epidemics on directed contact networks. *J. Theoret. Biol.* **240**, 400–418.
- [17] MEYERS, L. A., *et al.* (2005). Network theory and SARS: predicting outbreak diversity. *J. Theoret. Biol.* **232**, 71–81.
- [18] MOLLOY, M. AND REED, B. (1995). A critical point for random graphs with a given degree sequence. In *Proc. 6th Internat. Sem. on Random Graphs and Probabilistic Methods in Combinatorics and Computer Science*, John Wiley, New York, pp. 161–179.
- [19] NEWMAN, M. E. J. (2003). The structure and function of complex networks. *SIAM Rev.* **45**, 167–256.
- [20] PITMAN, J. AND YOR, M. (1997). The two-parameter Poisson–Dirichlet distribution derived from a stable subordinator. *Ann. Prob.* **25**, 855–900.
- [21] PRUITT, W. E. (1987). The contribution to the sum of the summand of maximum modulus. *Ann. Prob.* **15**, 885–896.
- [22] REITTU, H. AND NORROS, I. (2004). On the power-law random graph model of massive data networks. *Performance Evaluation* **55**, 3–23.
- [23] VAN DEN ESKER, H., VAN DER HOFSTAD, R., HOOGHIEMSTRA, G. AND ZNAMENSKI, D. (2005). Distances in random graphs with infinite mean degrees. *Extremes* **8**, 111–141.
- [24] VAN DER HOFSTAD, R., HOOGHIEMSTRA, G. AND VAN MIEGHEM, P. (2001). First-passage percolation on the random graph. *Prob. Eng. Inf. Sci.* **15**, 225–237.
- [25] VAN DER HOFSTAD, R., HOOGHIEMSTRA, G. AND VAN MIEGHEM, P. (2002). The flooding time in random graphs. *Extremes* **5**, 111–129.
- [26] VAN DER HOFSTAD, R., HOOGHIEMSTRA, G. AND VAN MIEGHEM, P. (2005). Distances in random graphs with finite variance degrees. *Random Structures Algorithms* **27**, 76–123.
- [27] VAN DER HOFSTAD, R., HOOGHIEMSTRA, G. AND ZNAMENSKI, D. (2007). Distances in random graphs with finite mean and infinite variance degrees. *Electron. J. Prob.* **12**, 703–766.
- [28] WÄSTLUND, J. (2006). Random assignment and shortest path problems. In *4th Colloq. Mathematics and Computer Science Algorithms, Trees, Combinatorics and Probabilities*, Assoc. Discrete Math. Theoret. Comput. Sci., Nancy, pp. 31–38. Available at <http://www.dmtcs.org/dmtcs-ojs/index.php/proceedings/issue/view/84/showToc>.