J Stat Phys (2018) 173:1082–1109
https://doi.org/10.1007/s10955-018-1957-5

Weighted Distances in Scale-Free Configuration Models

Erwin Adriaans[1] · **Júlia Komjáthy**[1]

Received: 28 September 2017 / Accepted: 6 January 2018 / Published online: 18 January 2018

Abstract In this paper we study first-passage percolation in the configuration model with empirical degree distribution that follows a power-law with exponent $\tau \in (2, 3)$. We assign independent and identically distributed (i.i.d.) weights to the edges of the graph. We investigate the weighted distance (the length of the shortest weighted path) between two uniformly chosen vertices, called typical distances. When the underlying age-dependent branching process approximating the local neighborhoods of vertices is found to produce infinitely many individuals in finite time—called explosive branching process—Baroni, Hofstad and the second author showed in Baroni et al. (J Appl Probab 54(1):146–164, 2017) that typical distances converge in distribution to a bounded random variable. The order of magnitude of typical distances remained open for the $\tau \in (2, 3)$ case when the underlying branching process is not explosive. We close this gap by determining the first order of magnitude of typical distances in this regime for arbitrary, not necessary continuous edge-weight distributions that produce a non-explosive age-dependent branching process with infinite mean power-law offspring distributions. This sequence tends to infinity with the amount of vertices, and, by choosing an appropriate weight distribution, can be tuned to be any growing function that is $O(\log \log n)$, where n is the number of vertices in the graph. We show that the result remains valid for the the erased configuration model as well, where we delete loops and any second and further edges between two vertices.

Keywords Random networks · Configuration model · Scale-free · Power-law degrees · Typical distances · First passage percolation

Mathematics Subject Classification Primary 60C05 · 05C80 · 90B15

✉ Júlia Komjáthy
j.komjathy@tue.nl

Erwin Adriaans
e.l.a.adriaans@student.tue.nl

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

1 Introduction

Every logistic company wants to be the fastest, cheapest and deliver on time. In order to achieve this, the routes they are driving should be (near-) optimal, meaning they should be the least costly and fastest for them in order to be competitive. This is just an example where weighted distances in a network play an important role. Other examples include the spreading of epidemics through society, the spreading of rumours, videos and advertisement through (online) social networks, and several other processes spreading on the internet.

The recent interest in understanding complex networks and processes on these networks motivates the study of more and more elaborate models for these (weighted) networks. The analysis of processes on these models often reveal finer topological aspects of the models themselves. And, vice versa, the organisation and topology of a network affect the behaviour of different processes on the network. Many real-life networks turn out to share some common properties, one of them being that the degree distribution follows a power-law [23,37], examples include the world-wide web [6], the movie-actor collaboration network [5], the network of citations of scientific publications [38], and many more. Another common property is the small-world phenomenon, popularized by Millgram [35] as: "everyone on this planet is separated from anyone else by only six people". Mathematically speaking, a network exhibits the small-world property if the minimal amount of connections to go from one node to another is of order $\log(n)$ or $\log\log(n)$ for ultra- small worlds, with n the amount of nodes in the network. This effect is not only seen in social networks, but also in neurological networks like the brain [1,17] or food webs [36]. A third common property is clustering as pointed out by Watts and Strogatz [40]. High clustering means that two vertices in the graph are more likely to be connected to one another when they have a common neighbor. This is a common feature in e.g. social networks.

The natural way to model a network from a mathematical point of view is to see this as a graph, where nodes are represented by vertices and their connections by edges. Since real-life networks are large, models often involve randomness to determine the presence of edges between the vertices. Random graph models that incorporate the (first two) above mentioned properties often serve as null-models for the analysis of real-life networks. Examples include variation of inhomogeneous random graphs such as the Chung-Lu or Norros-Reitu model [18,39], the configuration model [9,15], and the preferential attachment model [3]. Spatial variants are introduced to incorporate clustering, e.g. hyperbolic random graphs [14], geometric inhomogeneous random graphs [16], scale-free percolation [20], spatial preferred attachment [2,30], etc.

When modeling the spread of information in a network, edge weights to the edges can be added that represent the passage time of the information through the edge. The weighted distance is then the weight of the path with smallest total weight, corresponding to the passage time of the information from one vertex to the other. When the edge-weights are i.id., the study of the resulting weighted graph is often called *first-passage percolation* (FPP). Introduced by Hammersley and Welsh [24] for the grid $\mathbb{Z}^d$, FPP can be seen as a flow, starting from a vertex, flowing through the edges at a rate equal to the respective edge-weights, the weighted distance corresponding to the time it takes the front of the flow to reach the other vertex.

First passage percolation has been studied on the Erdős-Rényi random graph see [12], on inhomogeneous random graphs see [33]. FPP on the configuration model with finite mean degrees for exponential edge-weights is treated in [11], with finite variance degrees (i.e., power law exponent at least 3) and arbitrary edge-weight distributions in [13], and for infinite variance degrees (power-law exponent $\in (2, 3)$) for a class of edge-weights [8].

In particular, [8] determines the weighted distance when the edge-weights fall into what they call the *explosive* class. In this case, weighted distances converge in distribution (see Theorem 2.9 below), which heuristically means that regardless of how large the size of the network gets, the average weighted distance in the networks *stays bounded*. This explains the observed phenomena of extremely fast information spread in e.g. online networks such as meme spreading or viral spreading. The other class, where the weight distribution is 'non-explosive' is further studied in [7], for the special case when the edge weights are of the form $1 + X$. For this case [7] shows that the weighted distance is tight around the typical graph distance (that is, $2 \log \log n / |\log(\tau - 2)|$, where τ is the power-law exponent), if and only if the extra weight X falls into the explosive class.

In this paper we investigate the missing case, i.e., FPP on the configuration model with infinite variance degrees (power-law exponent $\in (2, 3)$) and i.i.d. edge-weights that fall into the 'non-explosive' class. We determine the first order of weighted distance in the highest generality, thus, together with [7] providing an (almost) full picture of weighted distances in the $\tau \in (2, 3)$ case. We also extend our results to the *erased configuration model*, when self-loops are deleted and only one edge of every multiple edge is kept.

Structure In the next section we introduce the configuration model and state our results, as well as discuss related results and open problems. In Sect. 3 we develop a coupling to branching processes (BPs), and state and prove some ingredient lemmas about the degrees and weighted distances within these BPs. In Sect. 4 we develop a crucial tool to prove the upper bound of the main result, degree-dependent percolation. In Sect. 5, we prove the main result and extend it to the erased configuration model.

Notation We say that a sequence of events $(\mathcal{E}_n)_{n\in\mathbb{N}}$ holds with high probability (whp) if $\lim_{n\to\infty} \mathbb{P}(\mathcal{E}_n) = 1$. For a sequence of random variables $(X_n)_{n\geq 1}$, we say than X_n converges in probability to a random variable X, shortly $X_n \xrightarrow{\mathbb{P}} X$, if for all $\varepsilon > 0$, $\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0$. Similarly, we say that X_n converges in distribution to a random variable X, shortly $X_n \xrightarrow{d} X$, if $\lim_{n\to\infty} \mathbb{P}(X_n \leq x) \to \mathbb{P}(X \leq x)$ for all $x \in \mathbb{R}$ where $\mathbb{P}(X \leq x)$ is continuous. For a non-decreasing right-continuous function $F(x)$ the generalised inverse of F is defined as $F^{(-1)}(x) := \inf\{y \in \mathbb{R} : F(y) \geq x\}$. For an edge $e = (x, y)$ we write L_e for the associated edge-length on e. We write lhs and rhs for the left-hand side and right-hand side, respectively.

2 Model and Results

In this section we introduce the weighted configuration model and present our results. Then we discuss related research and describe some open problems.

2.1 The Model

We consider the configuration model $\mathrm{CM}_n(\boldsymbol{d})$ on n vertices with degree sequence $\boldsymbol{d} = \{d_1, \ldots, d_n\}$. Let $\mathcal{H}_n := \sum_{v\in[n]} d_v$, the sum of the degrees with $[n] := \{1, 2, \ldots, n\}$. If $\mathcal{H}_n$ is odd we add an additional half-edge to vertex n, this does not further influence the analysis and we refrain from discussing this issue further. Given the degree sequence, the model is constructed as follows: To every vertex $v \in [n]$ we assign d_v half-edges, then we take a uniform random matching of the half-edges, where any two matched half-edges form an edge of the graph. The resulting random graph is denoted by $\mathrm{CM}_n(\boldsymbol{d})$. After constructing the edges, we assign each edge e an i.i.d. edge-length L_e from distribution L. We denote the

resulting weighted random graph by $\mathrm{CM}_n(\boldsymbol{d}, \boldsymbol{L})$, with $\boldsymbol{L} = (L_e)_{e\in[\mathcal{H}_n/2]}$. We assume that the empirical distribution function of the degrees, defined as $F_n(x) := \frac{1}{n}\sum_{v\in[n]} \mathbb{1}_{\{d_v\le x\}}$, satisfies the conditions for a power-law distribution, as given in the following assumption.

Assumption 2.1 (*Power-law tail behavior*). There exist $\tau \in (2, 3)$, $\gamma \in (0, 1)$, $C > 0$ and $\alpha > 1/2$ such that for all $x \in [0, n^\alpha)$,

$$\frac{1}{x^{\tau-1}}\mathrm{e}^{-C(\log x)^\gamma} \le 1 - F_n(x) \le \frac{1}{x^{\tau-1}}\mathrm{e}^{C(\log x)^\gamma}, \tag{2.1}$$

Additionally, we assume that $\min_{v\in[n]} d_v \ge 2$.

Under Assumptions 2.1, [32], there is a giant component of size $n(1 - o(1))$, thus two uniformly chosen vertices lie whp in the same connected component. Let D_n denote a random variable with distribution function F_n, the degree of an uniformly chosen vertex in $[n]$. We define B_n as the (size biased version of D_n)-1.

$$\mathbb{P}(B_n = k) := \frac{k+1}{\mathcal{H}_n}\sum_{v\in[n]} \mathbb{1}_{\{d_v=k+1\}} = \frac{k+1}{\mathbb{E}[D_n]}\mathbb{P}(D_n = k+1). \tag{2.2}$$

We write F_{B_n} for the distribution function of B_n. As shown in [29], F_{B_n} also satisfies a similar bound as (2.1), namely, for some $C^\star > 0$,

$$\frac{1}{x^{\tau-2}}\mathrm{e}^{-C^\star(\log x)^\gamma} \le 1 - F_{B_n}(x) \le \frac{1}{x^{\tau-2}}\mathrm{e}^{C^\star(\log x)^\gamma}. \tag{2.3}$$

To be able to relate models with different values of n to each other, we pose an additional assumption.

Assumption 2.2 (*Limiting distributions*). There exist distribution functions $F_D(x)$, $F_B(x)$ such that for some $\kappa > 0$,

$$\max\{\mathrm{d_{TV}}(F_n, F_D), \mathrm{d_{TV}}(F_{B_n}, F_B)\} \le n^{-\kappa},$$

where $\mathrm{d_{TV}}(F, G) := \frac{1}{2}\sum_{x\in\mathbb{N}} |F(x+1) - F(x) - (G(x+1) - G(x))|$ is the total variation distance between two (discrete) probability measures.

We denote the random variables following the distribution F_D and F_B of Assumption 2.2 by D and B. Clearly Assumption 2.2 implies $D_n \xrightarrow{d} D$ and $B_n \xrightarrow{d} B$. Since F_D and F_B are independent of n, it is elementary to show that they satisfy (2.1) and (2.3) for all $x \in \mathbb{N}$.

The goal of this paper to study the weighted distances in $\mathrm{CM}_n(\boldsymbol{d}, \boldsymbol{L})$.

Definition 2.3 (*Graph- and Weighted distance, Hopcount*). Let u and v be two vertices in $\mathrm{CM}_n(\boldsymbol{d}, \boldsymbol{L})$. Then the *graph distance* $d_G(u, v)$ is the number of edges used by the shortest path between u and v. The *weighted distance* or *L-distance* between u and v is defined as

$$d_L(u, v) := \min_{\pi:u\to v}\sum_{e\in\pi} L_e, \tag{2.4}$$

where the minimum is taken over all paths connecting u to v present in $\mathrm{CM}_n(\boldsymbol{d}, \boldsymbol{L})$. We set $d_L(u, v) = 0$ if $u = v$ and $d_L(u, v) = \infty$ if u and v are not connected. We define $d_H(u, v)$, the *hopcount*, as the number of edges on the optimal path realising $d_L(u, v)$. Finally, for two sets of vertices A, B, $\mathrm{d}_L(A, B) := \min_{x\in A, y\in B} \mathrm{d}_L(x, y)$.

The following theorem states the main result of this paper:

Theorem 2.4 (Weighted distances). *Consider the weighted configuration model* $\mathrm{CM}_n(\boldsymbol{d}, \boldsymbol{L})$ *satisfying Assumptions 2.1–2.2 and let u and v be two uniformly chosen vertices from* $[n]$. *Suppose that the distribution function* $F_L(x)$ *of L satisfies that*

$$\sum_{k=1}^{\infty} F_L^{(-1)}(\mathrm{e}^{-\mathrm{e}^k}) = \infty \tag{2.5}$$

Then, for the weighted distance,

$$d_L(u,v) \Big/ 2 \sum_{i=1}^{\left\lfloor \frac{\log\log n}{|\log(\tau-2)|} \right\rfloor} F_L^{(-1)}\left(\mathrm{e}^{-\left(\frac{1}{\tau-2}\right)^i}\right) \xrightarrow{\mathbb{P}} 1. \tag{2.6}$$

For the hopcount, for all $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}\Big(d_H(u,v) \Big/ 2\frac{\log\log n}{|\log(\tau-2)|} \le 1-\varepsilon\Big) = 0, \tag{2.7}$$

and whp, there exist at least one path of length at most $1+\varepsilon$ *times the denominator in (2.6), with number of edges at most* $(1+\varepsilon)2\log\log n/|\log(\tau-2)|$.

Convergence in distribution of the hopcount around $2\log\log n/|\log(\tau-2)|$ remains an open question, since the upper bound does not follow from our techniques. Namely, we cannot exclude the possibility of a much longer path with optimal total edge-length.

The weighted erased configuration model is defined as follows. After $\mathrm{CM}_n(\boldsymbol{d}, \boldsymbol{L})$ is constructed, we remove all self-loops and, if there are multiple-edges between two vertices, one of the edges is chosen uniformly at random independent of the edge weights and the other edges are deleted. The resulting graph is called the weighted erased configuration model, shortly $\mathrm{ECM}_n(\boldsymbol{d}, \boldsymbol{L})$. Let us denote the L-distance in this graph by $d_L^e(u,v)$ and the hopcount by $d_H^e(u,v)$.

Theorem 2.5 (Weighted distances in the erased configuration model). *Consider the erased configuration model* $\mathrm{ECM}_n(\boldsymbol{d}, \boldsymbol{L})$ *satisfying Assumptions 2.1–2.2 and let u and v be uniformly chosen from* $[n]$. *Suppose that the distribution function* $F_L(x)$ *of L satisfies (2.5). Then the results of Theorem 2.4 remain valid for* $d_L^e(u,v)$ *and* $d_H^e(u,v)$ *as well.*

Remark 2.6 (I.i.d. degrees). Using concentration techniques it can be shown that Assumptions 2.1 and 2.2 are satisfied whp when the degrees are i.i.d. coming from a background distribution function $F(x)$ satisfying (2.1) for all $x \in \mathbb{N}$, see [13].

Remark 2.7 [Explosive vs non-explosive edge-weight distributions] Given a particular distribution L for the edge weights, convergence versus divergence of the sum in (2.5) is elementary to check, and it depends on the behaviour of F_L around 0. The steeper F_L at the origin, the smaller the sum in (2.5): the sum converges e.g. if $F_L(x)$ increases as a polynomial in x around 0, which is the case for exponential, uniform, Gamma distributions. Distributions with support separated away from 0 always give a divergent sum, and distributions with inverse $F_L^{(-1)}(z) = O(1/\log\log(1/z))$ also diverge. This corresponds to the family of distributions $F_L(t) = \exp\{-C\exp\{-c/t^\beta\}\}$, that give explosion for $\beta < 1$ but non-explosion for $\beta \ge 1$.

Note that (2.5) does not require that F_L is continuous. By setting the edge weights to be deterministic and equal to 1 in Theorem 2.4, we obtain the following corollary. Stronger results about the graph distance were already obtained in [28,29].

Corollary 2.8 (Graph distances). *Consider the configuration model* $\mathrm{CM}_n(\boldsymbol{d})$ *satisfying Assumptions 2.1–2.2 and let u and v be uniformly chosen vertices from* $[n]$. *Then*

$$d_G(u,v)/2\frac{\log\log n}{|\log(\tau-2)|} \xrightarrow{\mathbb{P}} 1. \tag{2.8}$$

A counterpart of Theorem 2.4 is [8, Theorem], which we cite here for comparison.

Theorem 2.9 (Weighted distances with explosive edge-weights [8]). *Consider the weighted configuration model* $\mathrm{CM}_n(\boldsymbol{d},\boldsymbol{L})$ *satisfying Assumptions 2.1–2.2 and let u and v be two uniformly chosen vertices from* $[n]$. *Suppose that the distribution function* $F_L(x)$ *of L satisfies that the sum in* (2.5) *converges. Then*

$$d_L(u,v) \xrightarrow{d} Y^{(1)}+Y^{(2)}, \tag{2.9}$$

where $Y^{(1)}, Y^{(2)}$ *are i.i.d. copies of some a.s. finite random variable.*

Theorems 2.4 and 2.9 together describe typical distances in the configuration model with power law degrees, with exponent $\tau \in (2,3)$ for all edge-weight distributions L. Next we discuss some related literature and pose some open problems.

2.2 Discussion and Open Problems

Relation to age-dependent branching processes The configuration model has a tree-like local structure. Since most cycles are long, the local neighborhood of a uniformly chosen vertex exploration around a vertex can be coupled to a branching process (BP). When the edge-weights are incorporated in the model and in the coupling, this BP becomes age-dependent. In an age-dependent BP, individuals have an i.i.d. lifetime and give birth to their i.i.d. number of offspring upon death. Let us denote such a BP with offspring distribution X and life-time distribution σ by $\mathrm{BP}(X,\sigma)$. Let us write $\mathrm{BP}(D,X,\sigma)$ for D i.i.d. copies of $\mathrm{BP}(X,\sigma)$. Then, the local neighborhood of a vertex in $\mathrm{CM}_n(\boldsymbol{d},\boldsymbol{L})$ can be approximated by $\mathrm{BP}(D,B,L)$. Explosion of a BP means that the BP produces infinitely many individuals in finite time, with positive probability. In 2013 Amini et al [4] gave a necessary and sufficient condition for the explosion of $\mathrm{BP}(X,L)$, for offspring distributions X that satisfy $\mathbb{P}(X\geq x)\geq x^{-1-\varepsilon}$ for some $\varepsilon>0$. In an unpublished note [34], under the stronger assumption that X satisfies $x^{-1-\varepsilon}\leq \mathbb{P}(X\geq x)\leq x^{-\varepsilon}$ for some $\varepsilon>0$, the second author simplified this criterion to the sum in (2.5) being finite. The criterion (2.5) comes from the following observation: the L-length of any path in a BP leading to infinity can be lower bounded by the sum of the minimum edge-lengths in each generation. In generation k, the number of individuals is doubly exponential in k. The minimum of this many i.i.d. random variables of distribution L is approximately the kth term in the sum in (2.5). If this sum is infinite, the BP cannot explode, thus the summability of the minimums in each generation is necessary for the BP to explode. In [4], the authors showed that this notion of minimum-summability is sufficient as well by constructing an algorithm that finds an infinite path with finite total length.

To show distributional convergence of weighted distances in $\mathrm{CM}_n(\boldsymbol{d},\boldsymbol{L})$, as in Theorem 2.9, when the underlying BP explodes was the content of [8]. It remained open to characterise the growth of weighted distances when explosion does not happen. It follows from [4] that in the nonexplosive case, for offspring distribution X satisfying (2.3), the time to reach the first individual in generation ℓ grows as

$$\sum_{k=1}^{\ell} F_L^{-1}(\exp\{-(\tau-2)^{-k}\}).$$

This, combined with the fact that the graph distance of a typical vertex to a maximal degree vertex is $\log\log n/|\log(\tau-2)|$, gives a strong intuitive explanation for the formula for $d_L(u,v)$ in Theorem 2.4. Unfortunately, the BP approximation for $\mathrm{CM}_n(\boldsymbol{d})$ fails much earlier than reaching the maximal degree vertex, and the BP techniques do not reveal enough information on the structure of the optimal path leading generation k of the BP, in particular, they do not provide good enough lower bounds on the degrees along the path. These are the reasons why we need to use a different technique, degree-dependent percolation, to show the upper bound on $d_L(u,v)$. Unfortunately, this technique is not fine enough to show distributional convergence of the fluctuations of $d_L(u,v)$ around its typical value.

Problem 2.10 (*Tightness, distributional covergence*). Consider $\mathrm{CM}_n(\boldsymbol{d},\boldsymbol{L})$ satisfying Assumptions 2.1–2.2 and let u and v be uniformly chosen vertices from $[n]$. Suppose the distribution function $F_L(x)$ satisfies (2.5). Determine the conditions under which

$$d_L(u,v) - 2\sum_{i=1}^{\left\lfloor \frac{\log\log n}{|\log(\tau-2)|}\right\rfloor} F_L^{(-1)}\left(\mathrm{e}^{-\left(\frac{1}{\tau-2}\right)^i}\right), \qquad d_H(u,v) - 2\frac{\log\log n}{|\log(\tau-2)|} \tag{2.10}$$

are tight sequences of random variables. Do these sequences converge in distribution?

Infinite mean degrees, i.e., when $\tau \in (1,2)$, is investigated in [10,21], where the authors show that the graph distance is whp 2 or 3, the weighted distance converges to the sum of two random variables. Finite variance degrees, $\tau > 3$, is studied in [11,13,22,27]. In this case typical graph distances are of order $\log n$, weighted distances scale as a constant times $\log n$ with converging fluctuations around this value, while the hopcount, centered around another constant times $\log n$, satisfies a central limit theorem.

It still remains open to characterise weighted distances for the boundary exponents, i.e., when $\tau \in \{2,3\}$. For the $\tau = 3$ case, even the explosion of the underlying age dependent BP is an open question. For $\tau = 2$, local neighborhoods grow faster than double-exponential and the precise growth depends sensitively on the slowly varying function involved, thus the techniques used here do not apply directly.

Problem 2.11 ($\tau = 2$ or 3). Characterise weighted distances for the case when the degree distribution follows a power law (with a slowly varying function correction term) when $\tau = 2$ and when $\tau = 3$.

We further expect that similar results hold for a large class of power-law graph models, specially in the $\tau \in (2,3)$ regime, including inhomogeneous random graphs (e.g. the Chung-Lu or Norros-Reitu models), spatial models such as the geometric inhomogeneous random graphs and scale-free-percolation.

2.3 Overview of the Proof

Next we give an overview of the proof of Theorem 2.4. The proof consists of two parts, a lower and an upper bound that use slightly different techniques.

2.3.1 Lower Bound

Let us denote the graph distance ball of radius k around a vertex q in $\mathrm{CM}_n(\boldsymbol{d})$ by $B_k^G(q)$, and the set of vertices precisely at graph distance k away from q by $\Delta B_k^G(q)$. For the

proof of the lower bound we show that $B_k^G(u)$, $B_k^G(v)$ can be coupled to two independent branching processes (BPs), with the first generation having distribution function F_D and all further generations having distribution function F_B from Assumption 2.2. We show that the coupling can be maintained until two random indices $\kappa_n(u)$, $\kappa_n(v)$ such that polynomially many vertices in n are found around both vertices u, v, and that $B^G_{\kappa_n(u)}(u)$, $B^G_{\kappa_n(v)}(v)$ are whp disjoint. Since any path connecting u, v must intersect the boundaries of these sets, we obtain the lower bound

$$\begin{aligned} d_L(u,v) &\geq d_L\left(u, \Delta B^G_{\kappa_n(u)}(u)\right) + d_L\left(v, \Delta B^G_{\kappa_n(v)}(v)\right) \\ &\geq \sum_{q=u,v} \sum_{i=0}^{\kappa_n(q)-1} \min_{x\in\Delta B_i^G(q), y\in\Delta B_{i+1}^G(q)} \{L_{(x,y)}\} \end{aligned} \tag{2.11}$$

where we obtained the second line by lower bounding $d_L(q, \Delta B^G_{\kappa_n(q)}(q))$ by the sum of the minimal edge lengths connecting $\Delta B_i^G(q)$ to $\Delta B_{i+1}^G(q)$ over i. We show that this sum of minima is larger than $(1-\varepsilon)$ times the denominator of the lhs of (2.6) whp.

2.3.2 Upper Bound

The upper bound also couples the neighborhoods $B_k(u)$, $B_k(v)$ to two disjoint BPs, but we exploit the coupling only until we reach vertices u_{K_n}, v_{K_n} of degree at least $\widetilde{K}_n$, for some carefully chosen $\widetilde{K}_n$ that tends to infinity with n, but $d_L(u, u_{K_n})$ and $d_L(v, v_{K_n})$ are still of negligible length compared to the denominator of the lhs of (2.6). Then we connect the vertices u_{K_n} and v_{K_n} using degree-dependent percolation that we describe now.

The idea for degree-dependent percolation originates from [7], and it is an extension of a construction by Janson [31]. In the percolated graph, we keep edges independently of each other, with probabilities that depend on the degrees of the end vertices of the edge. We use the i.i.d. edge lengths to realise the percolation, i.e., an edge $e = (x, y)$ is kept if and only if its edge length satisfies

$$L_{x,y} \leq \xi(d_x, d_y), \tag{2.12}$$

for some appropriately chosen threshold function $\xi(\cdot,\cdot)$. We use a result from [7,31], that states that the percolated graph can be looked at as a subgraph G^r of a configuration model with a new degree sequence $\mathbf{d^r}$. We choose ξ in such a way that the new degree sequence still satisfies the power-law condition in (2.1), for the same τ but possibly different C, γ. Note that G^r is a subgraph of the original $\mathrm{CM}_n(\boldsymbol{d}, \boldsymbol{L})$, and as a result any path present in G^r was necessarily also present in $\mathrm{CM}_n(\boldsymbol{d}, \boldsymbol{L})$. We show that u_{K_n}, v_{K_n} has percolated degree at least K_n.

Then, we construct two paths, emanating from u_{K_n} and v_{K_n}, and reaching vertices $u^\star$, $v^\star$ of percolated degree at least $n^{\alpha(\tau-2)}$, respectively, in this percolated graph. We control the (growing) degrees of vertices along these paths and as a result (2.12) gives an upper bound on the edge-lengths along these paths. More precisely, analogous to [7], we define a sequence $y_i(K_n)$ with $y_0 = K_n$ and layers in the graphs $\Gamma_i := \{v \in [n] : d_v \geq y_i(K_n)\}$, for $0 \leq i \leq i_{\max}$ with $i_{\max}$ the number of layers. We show that a vertex in Γ_i is connected to a vertex in Γ_{i+1} whp, moreover the total error probability over all the layers tends to zero as $K_n \to \infty$. Thus whp there exist paths from u_{K_n}, v_{K_n} where the ith vertex along the path has degree at least $y_i(K_n)$. Finally, we connect the vertices $u^\star$, $v^\star$ in G^r via a path of length at most three edges using vertices with degree at least $n^{1/2+\delta}$ for some small $\delta > 0$. The length of the constructed path is at most

$$d_L(u,v) \le d_L(u,u_{K_n}) + d_L(v,v_{K_n}) + 2\sum_{i=0}^{i_{\max}} \xi(y_i(K_n), y_{i+1}(K_n)) + 3\xi(n^{\alpha(\tau-2)}, n^{1/2}). \tag{2.13}$$

The first two terms on the rhs, coming from the branching processes, are negligible due to the choice of K_n, and the last term also since it tends to zero with n. With the proper choice of $\xi(\cdot,\cdot)$, the middle term becomes at most $(1+\varepsilon/2)$ times the denominator of (2.6).

3 Exploration Around Two Vertices

The goal of this section is to couple the neighborhoods $B_k(u)$, $B_k(v)$ of the two uniformly chosen vertices u, v to two independent BPs. We first show that, for $q \in \{u, v\}$ the coupling can be maintained until $k = \kappa_n(q) = \log\log n/|\log(\tau-2)|+$ a tight random variable. Then, using the growth of the BPs, we make (2.11) quantitative by giving a whp lower bound on the minimum of edge-lengths connecting consecutive generations in the BPs.

As a preparation for the upper bound, as in (2.13), we determine M_n, the number of generations needed to reach a vertex with degree at least K_n, that we denote by q_{K_n}, for $q \in \{u, v\}$. Finally, we give an upper bound on $d_L(q, q_{K_n})$ for $q \in \{u, v\}$.

3.1 Coupling of the Exploration to a Branching Process

First we explain the coupling of the neighborhoods of the vertices u and v to branching processes. The coupling uses an exploration, where we reveal the pairs of half-edges and thus the neighbors of vertices together with their degrees one-by-one, in a breadth-first-search manner. By $\mathcal{U}_t, \mathcal{V}_t$ we denote the subgraphs consisting of vertices at graph distance of at most t from u and v, respectively. The forward degree of a vertex v in the exploration denotes then the number of new (not previously discovered) neighbors of a vertex upon exploration. We slightly adjust [29, Lemma 2.2]) to our setting, since Assumptions 2.1–2.2 are a special case of the assumptions of [29, Lemma 2.2]). An alternative formulation and proof can be found in [11, Proposition 4.7].

Lemma 3.1 (Coupling error of the exploration process, [29]). *Consider* $\mathrm{CM}_n(\boldsymbol{d})$ *satisfying Assumptions 2.1–2.2. Then, in the exploration process started from two uniformly chosen vertices u and v, the forward degrees* $(X_k^{(n)})_{k\le s_n}$ *of the first* s_n *newly discovered vertices can be coupled to an i.i.d. sequence* B_k *from distribution B as in Assumption 2.2. So, there is a coupling* $(X_k^{(n)}, B_k)_{k\le s_n}$ *with the following error bound*

$$\begin{aligned}\mathbb{P}(\exists k \le s_n, X_k^{(n)} \ne B_k) &\le C s_n^{(2\tau-2-2\varepsilon)/(\tau-\varepsilon)} n^{(2-\tau+\varepsilon)/(\tau-\varepsilon)} \\ &\quad + C s_n^2 n^{\frac{(1+\varepsilon)}{\tau-1}-1} + s_n n^{-\kappa}.\end{aligned} \tag{3.1}$$

An immediate corollary is the following:

Corollary 3.2 (Whp coupling of the exploration to BPs, [29]). *In the configuration model satisfying Assumptions 2.1 and 2.2, let t be such that*

$$\begin{aligned}|\mathcal{U}_t \cup \mathcal{V}_t| &\le 2\min\{n^{(1-(1+\varepsilon)/(\tau-1)-\delta)/2}, n^{-(\tau-2-2\varepsilon)/2(\tau-1-\varepsilon)}, n^{(2-\tau+\varepsilon)/(\tau-\varepsilon)}, n^{\kappa-\delta}\} \\ &=: 2n^{\theta(\delta)}\end{aligned} \tag{3.2}$$

for some $\delta > 0$. Then $(\mathcal{U}_t, \mathcal{V}_t)$ can be whp coupled to two i.i.d. BPs with generation sizes $(Z_k^{(u)}, Z_k^{(v)})_{k>0}$ with distribution F_B for the offspring in the second and further generations, and with distribution F for the offspring in the first generation.

The proofs of Lemma 3.1 and Corollary 3.2 can be found in [29, Sect. 2].

Davies in [19] shows that for a BP with offspring distribution satisfying the tail behavior in (2.3), the sequence of random variables $Y_k := (\tau-2)^k \log(Z_k)$ converges almost surely. It is elementary to extend his result to a BP where the root has a different offspring distribution (see [8] for details). Having this result in mind for the two BPs coupled to the neighborhoods of u, v, we rewrite the generation sizes as

$$Z_k^{(q)} =: \exp\left\{ \left(\frac{1}{\tau-2} \right)^k Y_k^{(q)} \right\}. \tag{3.3}$$

Fixing a small $\delta > 0$, we define for $q \in \{u, v\}$,

$$\kappa_n(q) := \max\left\{ k : Z_k^{(q)} \le n^{\theta(\delta)} \right\}, \tag{3.4}$$

and then Corollary 3.2 implies that $\Delta B_{\kappa_n(q)}(q)$ has size $Z_{\kappa_n(q)}^{(q)}$ since the coupling can still be maintained. Combining (3.3) and (3.4) we obtain that $Y_{\kappa_n(q)}^{(q)}$ converges *in distribution* to two independent copies of the same random variable Y. The convergence now is only distributional, since there is no coupling between the BPs for different values of n. Using (3.3) and (3.4), we can provide an *implicit* description of $\kappa_n(q)$, $q \in \{u, v\}$.

Claim 3.3 (Last generation of the exploration). *Consider* $\mathrm{CM}_n(\boldsymbol{d})$ *satisfying Assumptions 2.1–2.2. Let u and v be two uniformly chosen vertices. Then we can couple the BFS-exploration around u and v to two BPs until generation that has the implicit representation*

$$\kappa_n(q) = \frac{\log\log n + \log\left(\theta(\delta) f_n(q) / Y_{\kappa_n(q)}^{(q)} \right)}{|\log(\tau-2)|} \text{ for } q \in \{u, v\}, \tag{3.5}$$

where $\kappa_n(q)$ is an integer; $Y_{\kappa_n(q)}^{(q)}$, for $q \in \{u, v\}$ are independent and converge in distribution, and $f_n(q) \in (\tau-2, 1]$ describes the exponent $\theta(n) f_n(q)$ that satisfies $Z_{\kappa_n(q)}^{(q)} = n^{\theta(n) f_n(q)}$.

The message of this claim is that the coupling can be maintained until $\log\log n / |\log(\tau-2)|$ + a tight random variable many generations.

Proof Fix $\delta > 0$ small enough and set $\theta(\delta)$ as in Corollary 3.2. Corollary 3.2 then implies that the coupling error converges to zero as long as $|B_{k_1}(u) \cup B_{k_2}(v)| \le 2n^{\theta(\delta)}$. The definition of $\kappa(q)$ in (3.5) implies that this is indeed satisfied for $k_1 := \kappa_n(u), k_2 := \kappa_n(v)$. The value of $\kappa_n(q)$ in (3.5) is obtained by an elementary rearrangement of the formula (3.3) when k is replaced by $\kappa_n(q)$, and we took the integer part of the obtained expression. Note that $\kappa_n(q)$ is well-defined this way since the generation sizes are increasing doubly exponentially for all large enough k due to (3.3) and the fact that $Y_k^{(q)}$ would converge if we would let k tend to infinity, and as a result the total size of $B_k(q)$ is $1 + o(1)$ times the last generation size. □

Next we make (2.11) quantitative by giving a lower bound on the length of the path from q to generation $\kappa_n(q)$. Recall that $\Delta B_k(w)$ is the set of vertices at distance k from a vertex w in $\mathrm{CM}_n(\boldsymbol{d})$.

Lemma 3.4 (Lower bound on shortest path length). *Consider* $\mathrm{CM}_n(\boldsymbol{d})$ *satisfying Assumptions 2.1–2.2 with i.i.d. edge lengths from distribution* L *with distribution function* F_L *satisfying* (2.5). *Let* u, v *be two uniformly chosen vertices. Then, for* $q \in \{u, v\}$, *with* $\kappa_n(q)$ *as in* (3.5),

$$\lim_{n\to\infty} \mathbb{P}\left(d_L(u, \Delta B_{\kappa_n(q)}(q)) > (1-\varepsilon) \sum_{i=1}^{\left\lfloor \frac{\log\log n}{|\log(\tau-2)|} \right\rfloor} F_L^{(-1)}\left(\mathrm{e}^{-\left(\frac{1}{\tau-2}\right)^i} \right) \right) = 1. \tag{3.6}$$

Proof Under the coupling between the neighborhoods of u, v to the BPs established in Corollary 3.2 and Claim 3.3, for all $i \le \kappa_n(q)$, $|\Delta B_i(q)| = Z_i^{(q)}$, the size of generation i in the BP coupled to the neighborhood of u. Using the idea in (2.11), $d_L(u, \Delta B_{\kappa_n(q)}(q))$ is longer than the sum of the minimum edge-lengths between consecutive generations. That is,

$$d_L(u, \Delta B_{\kappa_n(q)}(u)) \ge \sum_{i=0}^{\kappa_n(q)-1} \min\left\{ L_{i,1}^{(q)}, \dots L_{i,Z_i^{(u)}}^{(q)} \right\}, \tag{3.7}$$

where $L_{i,j}^{(u)}$ are i.i.d. for all i, j and $q \in \{u, v\}$. We let

$$C_n := \max\left\{ \sup_{1\le k\le \kappa_n(q)} Y_k^{(q)}, h(n) \right\} \tag{3.8}$$

with $h(n)$ a function defined later on. By (3.3), $Z_i^{(q)} = \exp\{(\tau-2)^{-i} Y_i^{(q)}\}$, thus, using, (3.8),

$$Z_i^{(q)} \le \exp\{(\tau-2)^{-i} C_n\}. \tag{3.9}$$

For i.i.d. L_j, the following tail bound holds for any $N \in \mathbb{N}$, $z(N) > 0$:

$$\mathbb{P}\left(\min_{j\le N} L_j > z(N) \right) = (1 - F_L(z(N)))^N \ge 1 - N F_L(z(N)) \tag{3.10}$$

which is at most $1/N^\xi$ when we set, for some $\xi > 0$, $z(N) := F_L^{(-1)}(1/N^{1+\xi})$. Using that the minimum in (3.7) is non-increasing when increasing the number of variables involved, by (3.9), we can set N to be $\exp\{(\tau-2)^{-i} C_n\}$ to estimate the ith term in (3.7) from below using (3.10). Conditioning on the value C_n, $\kappa_n(q)$, combined with a union bound, yields that the inequality

$$\sum_{i=0}^{\kappa_n(q)-1} \min\left\{ L_{i,1}, \dots L_{i,Z_i^{(u)}} \right\} \ge \sum_{i=0}^{\kappa_n(q)-1} F_L^{(-1)}\left(e^{-(\tau-2)^{-i} C_n (1+\xi)} \right), \tag{3.11}$$

holds with error probability (conditioned on C_n, $\kappa_n(q)$) at most

$$E(C_n) := \sum_{i=0}^{\infty} e^{-\left(\frac{1}{\tau-2}\right)^i C_n \xi} \le C_1 e^{-C_n \xi} \tag{3.12}$$

for some constant $C_1 > 0$. Combining (3.7) and (3.11) yields that under the coupling, with error probability given in (3.12),

$$d_L(q, \Delta B_{\kappa_n(q)}(q)) \ge \sum_{i=0}^{\kappa_n(q)-1} F_L^{(-1)}\left(e^{-\left(\frac{1}{\tau-2}\right)^i C_n (1+\xi)} \right). \tag{3.13}$$

Next we transform the rhs to match the format in (3.6). With $\lceil a \rceil = \min\{m \in \mathbb{Z}, m \geq a\}$, $\lfloor b \rfloor = \max\{m \in \mathbb{Z}, m \leq a\}$, we use the following inequalities valid for monotone non-increasing functions g with $g(0) < \infty$:

$$\sum_{k=\lceil a \rceil+1}^{\lfloor b \rfloor} g(k) \overset{(\star)}{\leq} \int_a^b g(x)\mathrm{d}x, \qquad \int_0^{z+1} g(x)\mathrm{d}x \overset{(\triangle)}{\leq} \sum_{k=0}^{z} g(k). \tag{3.14}$$

We use $(\triangle)$ to bound the rhs of (3.13) from below, then we carry out the variable transformation $1/(\tau-2)^x C_n(1+\xi) = 1/(\tau-2)^y$, and transform the integral back to a sum using $(\star)$. The variable transformation shifts the summation boundaries by $\widetilde{C}_n := \log(C_n(1+\xi))/|\log(\tau-2)|$, and we obtain that

$$\sum_{i=0}^{\kappa_n(q)-1} F_L^{(-1)}\left(\mathrm{e}^{-\left(\frac{1}{\tau-2}\right)^i C_n(1+\xi)}\right) \geq \sum_{i=\lceil \widetilde{C}_n+1 \rceil}^{\kappa_n(q)+\lfloor \widetilde{C}_n \rfloor} F_L^{(-1)}\left(\mathrm{e}^{-\left(\frac{1}{\tau-2}\right)^i}\right). \tag{3.15}$$

We bound the upper summation boundary on the rhs from below. Recall C_n from (3.8), then

$$\lfloor \widetilde{C}_n \rfloor = \left\lfloor \frac{\log((1+\xi)\max\{\sup_{1\leq k\leq \kappa_n} Y_k^{(q)}, h(n)\})}{|\log(\tau-2)|} \right\rfloor \geq \frac{\log Y_{\kappa_n(q)}}{|\log(\tau-2)|} - 1 \tag{3.16}$$

Using now the formula for $\kappa_n(q)$ from (3.5),

$$\kappa_n(q) + \lfloor \widetilde{C}_n \rfloor \geq \frac{\log\log n + \log(\theta(n) f_n(q)(\tau-2))}{|\log(\tau-2)|}. \tag{3.17}$$

Next, the lower summation boundary on the rhs of (3.15) is not 1, and, if $\widetilde{C}_n \to \infty$, then this might cause too much difference from the desired sum in (3.6). Thus, for any fixed $\varepsilon > 0$ we define

$$R_n(\varepsilon) := \max_z \left\{ \sum_{k=1}^{z-1} F_L^{(-1)}\left(e^{-\left(\frac{1}{\tau-2}\right)^i}\right) \leq \frac{\varepsilon}{2} \sum_{i=1}^{\lfloor \log\log n/|\log(\tau-2)| \rfloor} F_L^{(-1)}\left(e^{-\left(\frac{1}{\tau-2}\right)^i}\right) \right\}. \tag{3.18}$$

Since the sum on the rhs between the brackets tends to infinity with n, so will $R_n(\varepsilon)$. Setting $\widetilde{C}_n = R_n(\varepsilon)$, combined with (3.17) and the fact that the summands tend to zero then implies that

$$\sum_{i=\lceil \tilde{C}_n+2 \rceil}^{\kappa_n(q)+\lfloor \widetilde{C}_n \rfloor} F_L^{(-1)}\left(e^{-\left(\frac{1}{\tau-2}\right)^i}\right) > (1-\varepsilon) \sum_{i=1}^{\left\lfloor \frac{\log\log n}{|\log(\tau-2)|} \right\rfloor} F_L^{(-1)}\left(\mathrm{e}^{-\left(\frac{1}{\tau-2}\right)^i}\right), \tag{3.19}$$

as desired. The choice $\widetilde{C}_n = \log(C_n(1+\xi))/|\log(\tau-2)| = R_n(\varepsilon)$, establishing the choice $h(n) = 1/(1+\xi)\,(\tau-2)^{-R_n(\varepsilon)}$ in (3.8). Since $R_n(\varepsilon)$ tends to infinity, so will C_n, ensuring that the error probability in (3.12) tends to zero as well. This finishes the proof of the lower bound. □

Next we do some preparations for the proof of the upper bound. First we investigate the number of generations we need to explore to reach a vertex of degree at least $\widetilde{K}_n$.

Lemma 3.5 (Generations needed to reach degree $\widetilde{K}_n$). *Consider* $\mathrm{CM}_n(\boldsymbol{d})$ *satisfying Assumptions 2.1–2.2. Let* u, v *be two uniformly chosen vertices and* $\widetilde{K}_n = O(\log n)$ *a sequence that tends to infinity with* n. *Then, for* $q \in \{u, v\}$, *for any* M *with* $M|\log(\tau-2)| > 1$,

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{w\in \Delta B_{M\log\log \widetilde{K}_n}(q)} d_w < \widetilde{K}_n \right) = 0. \tag{3.20}$$

Proof For brevity we write $M_n := M \log\log \widetilde{K}_n$. By Corollary (3.2), and Claim 3.3, for $q \in \{u, v\}$ $(B_k(q))_{k\le \kappa_n(q)}$ can be coupled to a BP, where, in each generation the degrees are i.i.d. from distribution B as in Assumption 2.2. Recall (3.3), write $1+\delta = M|\log(\tau-2)|$ and condition on whether $Y^{(q)}_{M_n}$ is less than $1/(\log \widetilde{K}_n)^{\delta/2}$ or not. Then

$$\begin{aligned}\mathbb{P}\left(\max_{w\in \Delta B_{M_n}(q)} d_w \le \widetilde{K}_n \right) &\le \mathbb{P}\left(\max_{w\in Z^{(q)}_{M_n}} B_w \le \widetilde{K}_n \mid Y^{(q)}_{M_n} > 1/(\log \widetilde{K}_n)^{\delta/2} \right) \\ &\quad + \mathbb{P}\left(Y^{(q)}_{M_n} < 1/(\log \widetilde{K}_n)^{\delta/2} \right). \end{aligned} \tag{3.21}$$

By Davies [19], the limiting variable $\lim_{n\to\infty} Y^{(q)}_k$ is almost surely positive on survival of the BP. By Assumption 2.1, $\mathbb{P}(B \ge 1) = 1$ and thus the BP cannot go extinct and therefore $\mathbb{P}(Y^{(q)}_{M_n} = 0) = 0$. Thus, the second term on the rhs in (3.21) converges to zero. For the first term,

$$\mathbb{P}\Big(\max_{w\in Z^{(q)}_{M_n}} B_w \le \widetilde{K}_n \mid Z^{(q)}_{M_n} \Big) = (F_B(\widetilde{K}_n))^{Z^{(q)}_{M_n}}.$$

Using the lower bound on F_B from (2.3), with $L(x) := \exp\{-C^\star(\log x)^\gamma\}$ we obtain

$$(F_B(\widetilde{K}_n))^{Z^{(q)}_{M_n}} \le \Big(1 - \frac{L(\widetilde{K}_n)}{\widetilde{K}_n{}^{\tau-2}}\Big)^{Z^{(q)}_{M_n}} \le \exp\Big\{ -L(\widetilde{K}_n) Z^{(q)}_{M_n} / \widetilde{K}_n{}^{\tau-2} \Big\}. \tag{3.22}$$

Using that $Z^{(q)}_{M_n} = \exp\{(\tau-2)^{-M_n} Y^{(q)}_{M_n}\}$ by (3.3) in (3.22), we obtain that

$$\begin{aligned} &\mathbb{P}\Big(\max_{w\in Z^{(q)}_{M_n}} B_w \le \widetilde{K}_n \mid Y^{(q)}_{M_n} > \frac{1}{(\log \widetilde{K}_n)^{\delta/2}} \Big) \\ &\le \exp\Big\{ - \frac{\exp\{-C^\star(\log \widetilde{K}_n)^\gamma\} \exp\{(\log \widetilde{K}_n)^{1+\delta/2}\}}{\widetilde{K}_n{}^{\tau-2}} \Big\} \\ &\le \exp\Big\{ - \exp\big\{ (\log \widetilde{K}_n)^{1+\delta/2} (1 - C^\star (\log \widetilde{K}_n)^{\gamma-1-\delta/2}) - (\tau-2)(\log \widetilde{K}_n)^{-\delta/2} \big\} \Big\} \\ &\le \exp\Big\{ - \exp\big\{ 1/2 (\log \widetilde{K}_n)^{1+\delta/2} \big\} \Big\} \xrightarrow{n\to\infty} 0. \end{aligned} \tag{3.23}$$

where we used that $M_n = M \log\log K_n$, with $1+\delta = M|\log(\tau-2)|$, the last inequality holds for $\widetilde{K}_n$ large. Thus both probabilities on the rhs in (3.21) tend to zero as $\widetilde{K}_n$ tends to infinity, which completes the proof. □

In the proof of the upper bound of the main theorem we run the exploration algorithm until we reach a vertex of degree $\widetilde{K}_n$. The path to this vertex is the sum of i.i.d. copies of edge weights. We show that whp this sum is less than some $\varepsilon_1 > 0$ times the denominator of the lhs of (2.6).

Lemma 3.6 (Upper bound on the length of the path in the exploration process). *Let $(L_i)_{i\geq 1}$ be i.i.d. from distribution F_L satisfying (2.5). Then for all $\varepsilon_1 > 0$ there exists a choice of M_n such that M_n tends to infinity with n and*

$$\lim_{n\to\infty} \mathbb{P}\left(\sum_{i=0}^{M_n} L_i < \varepsilon_1 \sum_{i=1}^{\left\lfloor \frac{\log\log n}{|\log(\tau-2)|} \right\rfloor} F_L^{(-1)}\left(\mathrm{e}^{-\left(\frac{1}{\tau-2}\right)^i} \right) \right) = 1. \tag{3.24}$$

The lemma immediately follows from the following, more general result.

Claim 3.7 *Let $(L_i)_{i\geq 1}$ be i.i.d. random variables from distribution F_L, and let $(a_m)_{m\geq 1}$ be an arbitrary sequence that tends to infinity as $m \to \infty$. Then, there exists a deterministic sequence $(z_L(m))_{m\geq 1}$ with $\lim_{m\to\infty} z_L(m) = \infty$ such that*

$$\lim_{m\to\infty} \mathbb{P}\Big(\sum_{i=1}^{z_L(m)} L_i \leq a_m \Big) = 1.$$

Proof We distinguish two cases, based on the tail behavior of L. When F_L does not satisfy any of these cases, L can be stochastically dominated by a random variable that does satisfy (at least) one of these cases and then the result follows by a simple stochastic domination argument.

Case (1): $\mathbb{E}[L] < \infty$. In this case, for some $\delta \in (0, 1)$, let

$$z_L(m) := \max_{z\in\mathbb{N}} \left\{ z \leq a_m^{\delta} \right\} \tag{3.25}$$

Clearly $z_L(m)$ tends to infinity with m when a_m does. Markov's inequality implies that

$$\mathbb{P}\left(\sum_{i=0}^{z_L(m)} L_i \geq a_m \right) \leq \frac{z_L(m)\mathbb{E}[L]}{a_m} \to 0. \tag{3.26}$$

Case (2): $\mathbb{E}[L] = \infty$ and additionally $\mathbb{P}(L_i > x) \leq 1/g(x)$ for some non-decreasing function $g(x)$. For some small $\delta, \varepsilon_2 > 0$, we define $z_L(m)$ implicitly by

$$z_L(m) g^{(-1)}((z_L(m))^{1+\varepsilon_2}) = a_m^{1-\delta}, \tag{3.27}$$

where $g^{(-1)}(x) = \inf\{y \in \mathbb{R} : g(y) \geq x\}$. Since $g(x)$ is non-decreasing and a_m tends to infinity, $z_L(m)$ tends to infinity as well. Note that when $g(x) \geq x^a$ for some $a \in (0, 1)$, capturing regularly varying cases, a lower bound on (3.27) can be explicitly calculated:

$$z_L(m) \geq a_m^{(1+(1+\varepsilon_2)/a)^{-1}(1-\delta)}$$

To estimate the lhs of (3.26) in this case, we use a truncation argument. We condition (3.24) on the maximum of the L_i being larger than $T_m := g^{(-1)}\left((z_L(m))^{1+\varepsilon_2}\right)$ or not, which gives us the following upper bound

$$\mathbb{P}\left(\sum_{i=0}^{z_L(m)} L_i \geq a_m \right) \leq \mathbb{P}\left(\exists i \leq z_L(m) : L_i > T_m\right) + \mathbb{P}\left(\sum_{i=0}^{z_L(m)} L_i \mathbb{1}_{\{L_i \leq T_m\}} \geq a_m \right). \tag{3.28}$$

First we focus on the first term in (3.28). Using that $\mathbb{P}(L_i > x) = 1/g(x)$ and the value T_m,

$$\mathbb{P}\left(\exists i \leq z_L(m) : L_i > T_m\right) \leq z_L(m)\mathbb{P}\left(L_i \geq T_m\right) \leq \frac{z_L(m)}{(z_L(m))^{1+\varepsilon_2}} = (z_L(m))^{-\varepsilon_2}, \tag{3.29}$$

that tends to zero as m tends to infinity. Next we investigate the second term in (3.28) which we bound with Markov's inequality,

$$\mathbb{P}\left(\sum_{i=0}^{z_L(m)} L_i \mathbb{1}_{\{L_i \leq T_m\}} \geq a_m\right) \leq \frac{z_L(m)\mathbb{E}[L_i \mathbb{1}_{\{L_i \leq T_m\}}]}{a_m} \tag{3.30}$$

Now we observe that $\mathbb{E}[L_i \mathbb{1}_{\{L_i \leq T_m\}}] \leq T_m$, and use this bound on the rhs of (3.30), and (3.27),

$$\frac{z_L(m)\mathbb{E}[L_i \mathbb{1}_{\{L_i \leq T_m\}}]}{a_m} \leq \frac{z_L(m)T_m}{a_m} \leq \frac{a_m^{1-\delta}}{a_m} \leq a_m^{-\delta} \tag{3.31}$$

Combining (3.29) and (3.31) implies that (3.28) tends to zero as m tends to infinity. This finishes the proof. □

4 Degree-Dependent Percolation on the Configuration Model

In this section we make the degree dependent-percolation precise, that we have described in Sect. 2.3.2. Percolation for the configuration model was studied in [31] and later adjusted for the degree-dependent version in [7].

The induced subgraph of G on vertex set S is the largest subgraph of G with edges that have both endpoints in S. We denote the induced graph of a graph G restricted to the vertices in a set S by $G_{|S}$. Let $p(d) : \mathbb{N} \to [0, 1]$ be a monotone decreasing function of d. For a half-edge s we write the percolation probability shortly as $p_s := p(d_{v(s)})$ with $v(s)$ the vertex that s is attached to and $d_{v(s)}$ the degree of vertex $v(s)$. Now we define two different ways to percolate the configuration model. After that we show equality in distribution for the two different percolated graphs.

Definition 4.1 (*Edge percolation*). Consider a configuration model $\mathrm{CM}_n(\boldsymbol{d})$ with half-edges already paired into edges. Delete any edge between vertices with degrees d, d' in the graph independently of all other edges with probability $p(d)p(d')$. We denote the resulting graph by $\widetilde{\mathrm{CM}}_n^{p(d)}(\boldsymbol{d})$.

As described in Sect. 2.3.2, we can realize the egde percolation on $\mathrm{CM}_n(\boldsymbol{d}, \boldsymbol{L})$ by using the i.i.d. edge-lengths in $(L_e)_e$ as auxiliary variables to determine which edges to keep. Then, the threshold function $\xi(d, d')$ as in (2.12) must satisfy $\mathbb{P}(L \leq \xi(d, d')) = p(d)p(d')$ for all $d, d' \in \mathbb{N}$.

Definition 4.2 (*Half-edge percolation*). Given a degree sequence $\boldsymbol{d} = (d_1, \ldots, d_n)$ and a half-edge s, we keep a half-edge with probability p_s independently. If we do not keep it, then we create a new vertex with one half-edge corresponding to the deleted half-edge. We call the newly created vertex and half-edge artificial. We denote the total number of artificial vertices by A. After this procedure is carried out for all half-edges, we pair all the half-edges uniformly at random, (including the artificial ones as well). In the end we take the induced subgraph on the n original vertices. We denote the resulting graph by $\mathrm{CM}_n^{p(d)}(\boldsymbol{d})$.

By denoting the number of half-edges that are kept at vertex i by d_i^r, and $\mathbf{1}(A)$ a sequence with A repetitions of the value 1, $\mathrm{CM}_n^{p(d)}(\boldsymbol{d})$ is nothing but $CM_{n+A}((\boldsymbol{d}^r, \mathbf{1}(A)))_{|[n]}$, i.e., the induced subgraph of the first n vertices of a configuration model with $n + A$ vertices, and degree sequence $(\boldsymbol{d}^r, \mathbf{1}(A))$ that is d_i^r for $i \leq n$ and 1 for $i \geq n$. A result in [7] is the following:

Proposition 4.3 (Equality in distribution of two percolated graphs). *Consider a function $p(d)$, the degree-dependent percolation $\widetilde{\mathrm{CM}}_n^{p(d)}(\boldsymbol{d})$ as in Definition 4.1. Then $\mathrm{CM}_n^{p(d)}(\boldsymbol{d}) \stackrel{d}{=} \widetilde{\mathrm{CM}}_n^{p(d)}(\boldsymbol{d})$, where $\mathrm{CM}_n^{p(d)}(\boldsymbol{d})$ is the half-edge percolation as described in Definition 4.2.*

The message of Proposition 4.3 is that we can understand the (connectivity) properties of the graph after the degree-dependent edge percolation $\widetilde{\mathrm{CM}}_n^{p(d)}(\boldsymbol{d})$ by studying a configuration model $\mathrm{CM}_{n+A}(d^r, 1(A))$ restricted to the first n vertices. In some sense this corollary enables to change the order of percolation and pairing. So, now on we focus on studying the properties of $\mathrm{CM}_n^{p(d)}(\boldsymbol{d}) = \mathrm{CM}_{n+A}(d^r, 1(A))|_{[n]}$. Importantly, we need to control the new degree sequence in $\mathrm{CM}_n^{p(d)}(\boldsymbol{d})$. Recall that the vector $\boldsymbol{d^r} := \{d_1^r, \ldots, d_n^r\}$ denotes the number of kept half edges attached to vertices in $[n]$ in Defintion 4.2. Let us write $F_n^r(x) := \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{d_i^r \leq x\}}$. The goal is to find conditions on $p(d)$ such that $F_n^r(x)$ still satisfies the conditions of (2.1), when F_n did so.

Lemma 4.4 (Empirical degree distribution after half-edge percolation). *Consider $\mathrm{CM}_n(\boldsymbol{d})$ with degree sequence satisfying Assumption 2.1. Perform half-edge percolation as described in Definition 4.2 on $\mathrm{CM}_n(\boldsymbol{d})$ with percolation function $p(d)$ satisfying*

$$p(d) > b \exp\{-c(\log(d))^\eta\} \tag{4.1}$$

for some constants $b, c > 0$ and $\eta \in (0, 1)$. Then there exists $\theta > 0$ such that for all $x \in [\theta, n^\alpha]$ the empirical degree distribution $F_n^r(x)$ of the degrees after percolation still satisfies Assumption 2.1, except the condition on the minimal degree, with the same τ, α, but possibly different $\gamma \in (0, 1)$.

Proof By Definition 4.2, half-edges are kept independently, and thus, given d_i, $d_i^r \stackrel{d}{=} \mathrm{Bin}(d_i, p(d_i))$, where $\mathrm{Bin}(n, p)$ is a binomial random variable with parameters n and p. As a result the random variables $(d_i^r)_{i \leq n}$ are independent given the initial degrees $(d_1, \ldots, d_n)$. The upper bound in Assumption 2.1 for F_n^r is elementary since

$$1 - F_n^r(x) = \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{\mathrm{Bin}(d_i, p(d_i)) > x\}} \leq \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{d_i > x\}} = 1 - F_n(x). \tag{4.2}$$

Next we show the lower bound. First we define for all $x < n^\alpha$

$$S(x) := \{v : d_v \geq s(x)\} \tag{4.3}$$

where $s(x) > x$ is a function of x that is defined later. Clearly,

$$1 - F_n^r(x) \geq \frac{1}{n}\sum_{i \in S(x)} \mathbb{1}_{\{\mathrm{Bin}(d_i, p(d_i)) > x\}} \tag{4.4}$$

We choose the value of $s(x)$ such that the probability that the indicators within the sum are 1 with high enough probability, for all $i \in S(x)$. Namely, if we choose $s(x)$ such that the expectation of the binomial, $d_i\, p(d_i)$, is higher than $2x$ for all vertices in $S(x)$, then we can

use the concentration of binomial random variables [25, Theorem 2.21] to get an upper bound on the probability that the indicator functions are 1. Let

$$s(x) = \frac{2x}{b} \mathrm{e}^{2c(\log(2x/b))^\eta} \tag{4.5}$$

then we find that

$$\begin{aligned} s(x)p(s(x)) &\geq 2x\mathrm{e}^{2c(\log(2x/b))^\eta} \mathrm{e}^{-c(\log(2x/b)+2c(\log(2x/b))^\eta))^\eta} \\ &= 2x\mathrm{e}^{2c(\log(2x/b))^\eta} \mathrm{e}^{-c\left(\log(2x/b)(1+2c(\log(2x/b))^{\eta-1})\right)^\eta}. \end{aligned}$$

Since $\eta < 1$, the factor $1 + 2c(\log(2x/b))^{\eta-1}$ in the exponent of the last factor is at most $3/2$ whenever $x \geq \frac{b}{2} \exp\left\{(4c)^{1-\eta}\right\} := \frac{b}{2}\widehat{\theta}$. Using this fact and the monotonicity of $dp(d)$, we find that for all $d_i > s(x)$,

$$d_i\, p(d_i) \geq s(x)p(s(x)) \geq 2xe^{c/2(\log 2x/b)^\eta} \geq 2x. \tag{4.6}$$

Then, for all $d_i > s(x)$, by [25, Theorem 2.21],

$$\begin{aligned} \mathbb{P}\left(\mathrm{Bin}(d_i,\, p(d_i)) > x\right) &\leq \mathbb{P}\left(\mathrm{Bin}(d_i,\, p(d_i) > \frac{s(x)p(s(x))}{2}\right) \\ &\leq e^{-s(x)p(s(x))/8} \leq e^{-x/4} < \frac{1}{8}, \end{aligned} \tag{4.7}$$

whenever $x > 4\log 8$. Using this we get for all $x \geq \max\left\{\frac{b}{2}\widehat{\theta}, 4\log 8\right\} := \theta$

$$\begin{aligned} &\mathbb{P}\left(n(1 - F_n^r(x)) \leq \frac{S(x)}{4}\right) \leq \mathbb{P}\left(\sum_{i\in S(x)} \mathbb{1}_{\{\mathrm{Bin}(d_i,\, p(d_i))>x\}} \leq \frac{S(x)}{4}\right) \\ &\leq \mathbb{P}\left(\mathrm{Bin}(|S(x)|, 7/8) \leq \frac{|S(x)|}{4}\right) \leq \mathrm{e}^{-|S(x)|/8}. \end{aligned}$$

Combining this estimate with a union bound,

$$\mathbb{P}\left(\exists x \in [\theta, n^\alpha] : n(1 - F_n^r(x)) \leq \frac{|S(x)|}{4}\right) \leq \sum_{x=\theta}^{n^\alpha} e^{-|S(x)|/8} \leq n^\alpha e^{-|S(n^\alpha)|/8}, \tag{4.8}$$

since $|S(x)|$ decreases as x increases. Using (2.1) $|S(x)|$ can be bounded from below as follows

$$|S(x)| = n(1 - F_n(s(x))) \geq n\frac{1}{s(x)^{\tau-1}} \mathrm{e}^{-c(\log s(x))^\eta}.$$

It is elementary to calculate that, with $s(x)$ as in (4.5), the rhs satisfies satisfies the lower bound in Assumption 2.1, with the same τ, α, while the new value of γ is $\max\{\gamma^{\mathrm{old}}, \eta\}$. Using this bound for $|S(x)|$ within the probability sign in (4.8) and for $|S(n^\alpha)|$ on the rhs of (4.8) we arrive at:

$$\mathbb{P}\left(\exists x \in [\theta, n^\alpha] : 1 - F_n^r(x) \leq \frac{1}{s(x)^{\tau-1}} \mathrm{e}^{-c(\log s(x))^\eta}\right) \leq \mathrm{e}^{\alpha \log n} \mathrm{e}^{-n^\varepsilon \mathrm{e}^{-c(\log n)^\eta}} \xrightarrow{n\to\infty} 0. \tag{4.9}$$

□

Next we prepare more for the proof of the upper bound of Theorem 2.4, by comparing the degree of a fixed vertex before and after the half-edge percolation. This will be used to ensure that u_{K_n}, v_{K_n} in Sect. 2.3.2 still has sufficiently high degree in the percolated subgraph.

Lemma 4.5 (Degree after percolation vs original degree). *Apply half-edge percolation as described in Definition 4.2 with percolation function $p(d)$ satisfying (4.1) on $\mathrm{CM}_n(\boldsymbol{d})$. Let $\widetilde{K}_n = O(\log n)$ an arbitrary sequence that tends to infinity with n. We define*

$$K_n := \sup\left\{m : 2m \le \widetilde{K}_n b \mathrm{e}^{-c(\log \widetilde{K}_n)^\eta}\right\}. \tag{4.10}$$

Then a vertex w with $d_w \ge \widetilde{K}_n$ in $\mathrm{CM}_n(\boldsymbol{d})$ has $d_w^r \ge K_n$ in $\mathrm{CM}_n^{p(d)}(\boldsymbol{d})$ whp.

Proof First we investigate the expected degree of a vertex after percolation. Consider a vertex w with degree d_w, as before d_w^r denotes the degree after the half-edge percolation. Recall that $d_w^r \stackrel{d}{=} \mathrm{Bin}(d_w, p(d_w))$. Thus

$$\mathbb{E}[d_w^r] \ge \mathbb{E}[\mathrm{Bin}(d_w, b\mathrm{e}^{-c(\log d_w)^\eta})] = d_w b \mathrm{e}^{-c(\log d_w)^\eta} = b\mathrm{e}^{\log d_w (1 - c(\log d_w)^{\eta-1})}.$$

The rhs is monotone increasing in d_w, and tends to infinity as $d_w \to \infty$. Therefore, by setting K_n as in (4.10), K_n tends to infinity when $\widetilde{K}_n$ does. By (4.10), the expectation of a $\mathrm{Bin}(d_w, p(d_w))$, for any $d_w \ge \widetilde{K}_n$, is larger than $2K_n$. Knowing that, we can use the concentration of binomial random variables [26, Theorem 2.21] to obtain a bound on the probability that the binomial is smaller than K_n, i.e.

$$\mathbb{P}\left(\mathrm{Bin}(d_w, p(d_w)) < K_n \mid d_w \ge \widetilde{K}_n\right) \le \exp\{-K_n/4\},$$

since K_n tends to infinity this finishes the proof. □

5 Upper and Lower Bound on Weighted Distances

In this section we give the proofs of Theorems 2.4 and 2.5. We start with the main result as stated in Theorem 2.4, after that we give the proof of Theorem 2.5. We start with the lower bound as stated in the following lemma:

Lemma 5.1 (Lower bound on the weighted graph distance). *Consider $\mathrm{CM}_n(\boldsymbol{d})$ satisfying Assumptions 2.1–2.2 and let u and v be uniformly chosen from $[n]$. Suppose the edge lengths are i.i.d. with distribution function $F_L(x)$ that satisfies (2.5). Then for all $\varepsilon > 0$*

$$\lim_{n\to\infty} \mathbb{P}\left(d_L(u,v) > (1-\varepsilon)2 \sum_{i=1}^{\left\lfloor \frac{\log\log n}{|\log(\tau-2)|}\right\rfloor} F_L^{(-1)}\left(\mathrm{e}^{-\left(\frac{1}{\tau-2}\right)^i}\right)\right) = 1, \tag{5.1}$$

and for the hopcount

$$\mathbb{P}\left(\mathrm{d}_H(u,v) > (1-\varepsilon)2\log\log n/|\log(\tau-2)|\right) = 1. \tag{5.2}$$

Proof We consider two uniformly chosen vertices u and v. We do a BFS-exploration on both sides and by Lemma 3.3, we can couple these explorations whp to two independent BPs until generations $\kappa_n(u), \kappa_n(v)$ respectively. We write $\Delta B_{\kappa_n(q)}(q)$ for the set of vertices at distance $\kappa_n(q)$ from vertex $q \in \{u, v\}$, respectively. By the coupling, these explorations are disjoint

whp. Since any path connecting u, v must intersect $\Delta B_{\kappa_n(u)}(u)$, $\Delta B_{\kappa_n(v)}(v)$, we have the following lower bounds on the weighted distance and the hopcount between u, v:

$$\begin{aligned} d_L(u,v) &\geq d_L(u, \Delta B_{\kappa_n(u)}(u)) + d_L(v, \Delta B_{\kappa_n(v)}(v)), \\ \mathrm{d}_H(u,v) &\geq \kappa_n(u) + \kappa_n(v). \end{aligned} \tag{5.3}$$

Then, (5.1) directly follows from the first inequality and Lemma 3.4. By (5.3), the result of the lemma follows by a union bound. For the hopcount, the second inequality combined with Lemma 3.3 yields (5.2), since $Y^{(q)}_{\kappa_n(q)}$ converges in distribution. □

For the proof of the upper bound we use a proposition, similar to [7, Proposition 2.1], which gives an upper bound on the path length between two vertices of a fixed degree of at least K. In our setting the vertices have degree at least K_n with K_n tending to infinity with n. We provide the adjusted proof since the adjustments are non-trivial.

Proposition 5.2 *Consider* $\mathrm{CM}_n(\boldsymbol{d})$ *satisfying* (2.1) *for all* $x \in [\theta, n^\alpha]$ *for some given* $\theta \in \mathbb{R}$ *and some* $\alpha > 1/2$. *Let* w *be a vertex with degree at least* K_n. *Then, whp, there exists a path* $(\pi_0 = w, \pi_1 \ldots, \pi_{i_{\max}} = w^\star)$ *from* w *to a vertex* $w^\star$ *with degree at least* $n^{(\tau-2)\alpha}$ *such that the degree* d_{π_i} *of the* i*th vertex on the path satisfies*

$$d_{\pi_i} \geq \left(K_n^{1-\delta_n}\right)^{\left(\frac{1}{\tau-2}\right)^i} =: \widetilde{y}_i(K_n), \tag{5.4}$$

with $\delta_n \to 0$ *as* $K_n \to \infty$. *Whp,* $i_{\max}$, *the length of this path is at most*

$$i_{\max} \leq \frac{\log\log n}{|\log(\tau-2)|} - \frac{\log\log K_n}{|\log(\tau-2)|}. \tag{5.5}$$

Proof We shall denote the number of edges on the path from w to $w^\star$ by $i_{\max}$ and we define the following sets of vertices

$$\Gamma_i := \{v \in [n] : d_v \geq y_i(K_n)\} \tag{5.6}$$

for some increasing sequence $y_i(K_n) =: y_i$ to be determined shortly. $(\Gamma_i)_{i \leq i_{\max}}$ can be seen as layers of the graph, where $i_{\max}$ is the maximal i such that Γ_i is non-empty. Our goal is to prove that there exists a sequence $y_i(K_n)$, defining the layers Γ_i, such that the following holds:

$$\lim_{n\to\infty} \sum_{i=0}^{i_{\max}} \mathbb{P}\left(\pi_i \in \Gamma_i, \pi_i \nrightarrow \Gamma_{i+1} \mid d_w \geq K_n\right) = 0, \tag{5.7}$$

where π_i is a vertex chosen from Γ_i according to the size-biased distribution, equivalently, vertex π_i is the vertex that a uniformly chosen half-edge from Γ_i is attached to. Conditioning on the total number of half edges $\mathcal{H}_n$ in $\mathrm{CM}_n(\boldsymbol{d})$, and $\mathcal{H}_{y_i}$, the number of half-edges attached to vertices in the set Γ_i, by pairing the half-edges of a vertex $z \in \Gamma_i$, we can pair at least $y_i/2$ half-edges before all the half-edges of z are paired, and each of these half-edges is paired to a half-edge attached to a vertex in Γ_{i+1} with probability at least $1 - \mathcal{H}_{y_{i+1}}/\mathcal{H}_n$. Thus,

$$\mathbb{P}\left(z \in \Gamma_i, z \nrightarrow \Gamma_{i+1} \mid \mathcal{H}_{y_{i+1}}, \mathcal{H}_n\right) \leq \left(1 - \frac{\mathcal{H}_{y_{i+1}}}{\mathcal{H}_n}\right)^{y_i/2}. \tag{5.8}$$

Note in particular that this bound holds when the vertex z is chosen randomly from Γ_i in a way that does not take into account its connections, in particular it holds when z is chosen according to the size-biased distribution from Γ_i. Since any vertex in Γ_i has degree larger

than y_i and $|\Gamma_i| = n(1 - F_n(y_i))$, $\mathcal{H}_{y_i} \geq y_i n(1 - F_n(y_i))$. Under Assumptions 2.1, 2.2, $\mathcal{H}_n \leq \varphi n$ for some $\varphi \in \mathbb{R}$, thus, we have by (5.8)

$$\mathbb{P}(\pi_i \in \Gamma_i, \pi_i \nrightarrow \Gamma_{i+1}) \leq \exp\left\{-\frac{y_i y_{i+1}(1 - F_n(y_{i+1}))}{2\varphi}\right\}. \tag{5.9}$$

For now we focus on the term in the exponent. Using (2.1), we lower bound

$$\frac{y_i y_{i+1}(1 - F_n(y_{i+1}))}{2\varphi} \geq \tilde{c} y_i y_{i+1}^{2-\tau} \mathrm{e}^{-C(\log y_{i+1})^{\gamma}} = \tilde{c} y_i y_{i+1}^{2-\tau-C(\log y_{i+1})^{\gamma-1}}, \tag{5.10}$$

with C defined in (2.1) and $\tilde{c}$ some positive constant. Now we would like to choose the sequence $y_i = y_i(K_n)$ such that (5.9) converges to zero in particular that (5.7) holds. We claim that this holds when y_i is given by the following recursion

$$y_0 = K_n, \qquad y_{i+1} = y_i^{(\tau-2+A(\log y_i)^{\gamma-1})^{-1}} \tag{5.11}$$

with $A > 0$ defined later. Note that for sufficiently large K_n, since $\gamma < 1$,

$$\tau - 2 + A(\log y_0)^{\gamma-1} < 1 \tag{5.12}$$

Now let $\varkappa_n = A(\log K_n)^{\gamma-1}$, then

$$y_{i+1} \geq y_i^{(2-\tau+\varkappa_n)^{-1}} \geq \cdots \geq K_n^{(\tau-2+\varkappa_n)^{-i}}, \tag{5.13}$$

in particular, y_i is increasing doubly exponentially. We use the recursion relation of (5.11) in (5.10)

$$\tilde{c} y_i y_{i+1}^{2-\tau-C(\log y_{i+1})^{\gamma-1}} = \tilde{c} y_i^{\frac{2-\tau-C(\log y_{i+1})^{\gamma-1}}{\tau-2+A(\log y_i)^{\gamma-1}}+1} \geq \tilde{c} y_i^{\frac{A(\log y_i)^{\gamma-1}-C(\log y_{i+1})^{\gamma-1}}{\tau-2+\varkappa_n}} \tag{5.14}$$

Choose $A \geq 2C$ and use that the sequence y_i is increasing and the lower bound in (5.13), then

$$\tilde{c} y_i^{\frac{A(\log y_i)^{\gamma-1}-C(\log y_{i+1})^{\gamma-1}}{\tau-2+\varkappa_n}} \geq \tilde{c} \exp\left\{\frac{C(\log y_i)^{\gamma}}{\tau - 2 - \varkappa_n}\right\} \geq \tilde{c} \exp\left\{\frac{\widetilde{C}(\log K_n)^{\gamma}}{(\tau - 2 + \varkappa_n)^{i\gamma}}\right\}, \tag{5.15}$$

with $\widetilde{C} = C/(\tau - 2 - \varkappa_n)$. Combining everything from (5.10), we can use this lower bound in the exponent on the rhs of (5.9), and, since $\tau - 2 + \varkappa_n < 1$, the rhs of (5.9) is summable in i. Summing the lhs of (5.9) over i and then use the above bound we obtain

$$\sum_{i=0}^{\infty} \mathbb{P}(\pi_i \in \Gamma_i, \pi_i \nrightarrow \Gamma_{i+1} \mid d_w \geq K_n) \leq \hat{C} \exp\Big\{-\tilde{c} \exp\Big\{\frac{C(\log K_n)^{\gamma}}{(\tau - 2 + \varkappa_n)^{\gamma}}\Big\}\Big\}, \tag{5.16}$$

which tends to zero with n as K_n tends to infinity with n. This result yields the statement of (5.7). Using the result of [7, Lemma 2.6], the lower bound in (5.13) can be improved to

$$y_i \geq \left(K_n^{1-\delta_n}\right)^{(\tau-2)^{-i}} = \widetilde{y}_i(K_n), \tag{5.17}$$

with $\delta_n \to 0$ as $K_n \to \infty$.[1]

The path in the statement of the lemma is then constructed as follows, starting from the first vertex $\pi_0 := w$. By the first term in the sum in (5.7), π_0 is whp connected to at least one vertex in Γ_1. By the fact that the pairs of the half edges of π_0 are chosen uniformly, π_1 is a vertex attached to a uniformly chosen half-edge in Γ_1. As a result, π_1 is chosen

[1] From the proof of [7, Lemma 2.4], it is immediate that $\delta_n \leq \widetilde{D}(\log K_n)^{\gamma-1}$ for some constant $\widetilde{D} > 0$.

according to the size-biased distribution within Γ_1. Then we iterate this procedure to obtain $\pi_2, \pi_3, \ldots$ in $\Gamma_2, \Gamma_3, \ldots$ until the layers become empty by the constraint that the maximal degree guaranteed to exist is of order n^α. We denote the last vertex on the path by $w^\star$. The constructed path might jump across a layer, so the number of layers is an upper bound on the length of the path from w to $w^\star$. The last layer that is nonempty is then $\Gamma_{i_{\max}}$ with $i_{\max}$ is the largest integer with

$$K_n^{(1-\delta_n)(\tau-2)^{-i_{\max}}} \leq n^\alpha. \tag{5.18}$$

Since $\delta_n \to 0$, $\alpha < 1$, the following upper bound then holds

$$i_{\max} \leq \frac{\log\log n - \log\log K_n + \log\alpha - \log(1-\delta_n)}{|\log(\tau-2)|} \leq \frac{\log\log n - \log\log K_n}{|\log(\tau-2)|}. \tag{5.19}$$

We yet have to show that $y_{i_{\max}} \geq n^{\alpha(\tau-2)}$. For this, elementary rearrangement yields that the lhs of (5.18) equals $n^{\alpha(\tau-2)^\beta}$, with $\beta \in [0, 1)$ the fractional part of the middle term in (5.19). This finishes the proof. □

Lemma 5.3 (Upper bound on the weighted graph distance). *Consider* $\mathrm{CM}_n(\boldsymbol{d}, \boldsymbol{L})$ *satisfying Assumptions 2.1–2.2 and* u, v *two uniformly chosen vertices. Suppose the edge weights are i.i.d. from* $F_L(x)$ *that satisfies (2.5). Then for all* $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}\left(d_L(u,v) < (1+\varepsilon)2 \sum_{i=1}^{\left\lfloor \frac{\log\log n}{|\log(\tau-2)|} \right\rfloor} F_L^{(-1)}\left(\mathrm{e}^{-\left(\frac{1}{\tau-2}\right)^i} \right) \right) = 1. \tag{5.20}$$

Further, there exists a path between u, v *with at most* $(1+\varepsilon)2\log\log n/(\tau-2)$ *edges and having total length at most* $(1+\varepsilon)2\sum_{i=1}^{\left\lfloor \frac{\log\log n}{|\log(\tau-2)|} \right\rfloor} F_L^{(-1)}\left(\mathrm{e}^{-\left(\frac{1}{\tau-2}\right)^i} \right)$.

Proof For brevity let $a_n := \sum_{i=1}^{\left\lfloor \frac{\log\log n}{|\log(\tau-2)|} \right\rfloor} F_L^{(-1)}\left(\mathrm{e}^{-\left(\frac{1}{\tau-2}\right)^i} \right)$. First we construct the initial segments of the connecting path from both ends from u, v, as described heuristically in Sect. 2.3.2. Let M_n be as in Lemma 3.6, with $\varepsilon_1 := \varepsilon/3$. Then, consider any vertex w of graph distance M_n away in $\mathrm{CM}_n(\boldsymbol{d})$ from $q \in \{u, v\}$, chosen *independently* of $(L_e)_e$. Then, since the edge-lengths in $\mathrm{CM}_n(\boldsymbol{d}, \boldsymbol{L})$ are i.i.d. on the edges of the path from q to w, by Lemma 3.6, $\mathrm{d}_L(q, w) \leq a_n\varepsilon/3$ in $\mathrm{CM}_n(\boldsymbol{d}, \boldsymbol{L})$ whp. As a result of Lemma 3.5, for any M with $M|\log(\tau-2)| > 1$, at graph distance $M\log\log\widetilde{K}_n$ away from $q \in \{u, v\}$, there is at least one vertex with degree $\widetilde{K}_n$ in $\mathrm{CM}_n(\boldsymbol{d})$ whp. Thus, by defining $\widetilde{K}_n$ via $M_n = M\log\log\widetilde{K}_n$, (equivalently, $\widetilde{K}_n := \exp\left\{\exp\{M_n/M\}\right\}$), we find vertices with degree at least $\widetilde{K}_n$ at graph distance M_n away from $q \in \{u, v\}$, whp. Then, pick q_{K_n} for $q \in \{u, v\}$ in an arbitrary way among these vertices that is independent of $(L_e)_e$. Then, the previous argument applies and whp,

$$\mathrm{d}_L(q, q_{K_n}) \leq a_n\varepsilon/3 \tag{5.21}$$

in $\mathrm{CM}_n(\boldsymbol{d}, \boldsymbol{L})$ for $q \in \{u, v\}$.

Next we connect u_{K_n}, v_{K_n} using degree-dependent percolation. When applying edge-dependent percolation (as in Definition 4.1) on $\mathrm{CM}_n(\boldsymbol{d}, \boldsymbol{L})$, we can use the edge-lengths $(L_e)_e$ as auxiliary variables to decide which edge to keep. Namely, we keep edge e iff $L_e \leq \xi(d, d')$, with $\xi(d, d')$ satisfying $\mathbb{P}(L \leq \xi(d, d')) = p(d)p(d')$. By Proposition 4.3,

we can consider the percolated (sub)graph as an instance of a configuration model where the new degree sequence is $(\boldsymbol{d}^r, \mathbf{1}(A))$. We yet have to specify the percolation function that we use. For some $c > 0$, $\eta \in (0, 1)$ to be determined later, let

$$p(d) = \exp\{-c(\log d)^\eta\}. \tag{5.22}$$

The conditions of Lemma 4.5 apply, thus, with K_n as in (4.10), $d^r_{q_{K_n}} \geq K_n$ whp for $q \in \{u, v\}$. Further, the conditions of Lemma 4.4 are also satisfied, thus the $(\boldsymbol{d}^r, \mathbf{1}(A))$ sequence obtained after percolation still satisfies Assumption 2.1 (except the condition on the minimal degree being at least 2). Hence, following Proposition 5.2, we construct a path connecting u_{K_n}, v_{K_n} in the *percolated graph*, with good control on the (percolated) degrees along the path. Note that this path only uses vertices with percolated degree strictly larger than one, i.e., artificial vertices are not used, and as a result the constructed path is part of $\mathrm{CM}_n^{p(d)}(\boldsymbol{d})$ by definition.

For $q \in \{u, v\}$, we use the constructed path as described in Proposition 5.2 starting from q_{K_n} to reach a vertex $q^\star$ with $d^r_{q^\star} \geq n^{\alpha(\tau-2)}$. A lower bound on the degree of the ith vertex on this path is given by $\widetilde{y}_i(K_n) =: \widetilde{y}_i$ in (5.4). Since $p(d)$ is monotone decreasing, $\xi(d, d')$ is non-increasing in both variables. Thus, the edge-lengths on the constructed path are at most $\xi(\widetilde{y}_i, \widetilde{y}_{i+1})$ for $i = 0, 1, \ldots, i_{\max} - 1$. Hence, for $q \in \{u, v\}$

$$d_L(q_{K_n}, q^\star) \leq \sum_{i=0}^{i_{\max}-1} \xi(\widetilde{y}_i, \widetilde{y}_{i+1}). \tag{5.23}$$

Next we connect the two high-degree vertices $u^\star$ and $v^\star$. Recall that $\mathcal{H}_n$ stands for the total number of half-edges, and is at least some constant φ times n under Assumption 2.1 even without the minimal degree assumption. Fix $\delta \in (0, \alpha - 1/2)$ and write $\Lambda_{1/2+\delta} := \{w : d^r_w \geq n^{1/2+\delta}\}$, as well as $\mathcal{H}_{1/2+\delta} := \sum_{w \in \Lambda_{1/2+\delta}} d^r_w$. Then, following (5.8)–(5.10),

$$\begin{aligned}\mathbb{P}\left(q^\star \nleftrightarrow \Lambda_{1/2+\delta} \mid \mathcal{H}_n, \mathcal{H}_{1/2+\delta}\right) &\leq \left(1 - \frac{\mathcal{H}_{1/2+\delta}}{\mathcal{H}_n}\right)^{n^{\alpha(\tau-2)}/2} \\ &\leq \exp\left\{-cn^{\alpha(\tau-2)+(2-\tau)(1/2+\delta)-o(1)}\right\},\end{aligned} \tag{5.24}$$

which tends to zero as $n \to \infty$ since $1/2 + \delta < \alpha$. Thus, we can find vertices $u^{\star\star}, v^{\star\star} \in \Lambda_{1/2+\delta}$ such that $(q^\star, q^{\star\star})$ are kept edges in the percolated graph whp. Finally, we show that the edge $(u^{\star\star}, v^{\star\star})$ is also present whp in the percolated graph.

$$\mathbb{P}\left(u^{\star\star} \nleftrightarrow v^{\star\star} \mid \mathcal{H}_n\right) \leq \left(1 - \frac{n^{1/2+\delta}}{\mathcal{H}_n}\right)^{n^{1/2+\delta}/2} \leq \exp\left\{-cn^{(1+\delta)-1}\right\}, \tag{5.25}$$

which tends to zero as $n \to \infty$. By the monotonicity of ξ, whp, the vertices $u^\star$ and $v^\star$ are connected via at most 3 edges with length at most

$$d_L(u^\star, v^\star) \leq 3\xi(n^{\alpha(\tau-2)}, n^{1/2}). \tag{5.26}$$

Combining (5.21), (5.23) and (5.26), we arrive at (2.13). In what follows we show that the rhs of (5.23) is at most $(1 + \varepsilon/3)a_n$. By the definition of edge-percolation in Definition 4.1, we keep an edge connecting vertices with degrees y_i, y_{i+1} with probability at most $p(y_i)p(y_{i+1})$. Using the form $p(d)$ from (5.22) and its monotonicity, and the lower bound on y_i from Prop. 5.2,

$$\begin{aligned}p(y_i)p(y_{i+1}) &\leq \exp\{-c(\log y_i)^\eta - (\log y_{i+1})^\eta\} \\ &\leq \exp\left\{-\widetilde{c}(\tau-2)^{-\eta(i+1)}(\log K_n^{1-\delta_n})^\eta\right\},\end{aligned} \tag{5.27}$$

with $\widetilde{c} = c(1 + (\tau - 2)^{\eta})$. Recall that we keep an edge e between vertices with degrees d, d' iff its edge-length is at most $\xi(d, d')$ in the edge-percolation. This gives us $\xi(d, d') = F_L^{(-1)}(p(d)p(d'))$, and, by monotonicity again, from (5.27) it follows that

$$\xi(y_i, y_{i+1}) \leq F_L^{(-1)}\Big(\exp\big\{-c(\log K_n^{1-\delta_n})^{\eta}(\tau-2)^{-\eta(i+1)}\big\}\Big). \tag{5.28}$$

Combining (5.27) with the bound on $i_{\max}$ from Proposition 5.2, (5.23) can be bounded above as

$$d_L(q_{K_n}, q^{\star}) \leq \sum_{i=1}^{\left\lfloor \frac{\log(\log n/\log K_n)}{|\log(\tau-2)|}\right\rfloor} F_L^{(-1)}\Big(\exp\big\{-c(\log K_n^{1-\delta_n})^{\eta}(\tau-2)^{-\eta i}\big\}\Big). \tag{5.29}$$

Similar to the proof of Lemma 3.4, we need to transform the rhs to the desired form in (5.20). Using similar bounds as in (3.14), we rewrite the sum to an integral, change variables as $(\tau-2)^{-\eta x}(\log K_n^{1-\delta_n})^{\eta} =: (\tau-2)^{-y}$, and change the integral back to a sum. This operation shifts the summation boundaries by $\eta \log\log K_n/|\log(\tau-2)|$ and multiplies the whole sum by η. We obtain

$$d_L(q_{K_n}, q^{\star}) \leq \frac{1}{\eta} \sum_{i=\left\lfloor \eta \frac{\log\log K_n}{|\log(\tau-2)|}\right\rfloor}^{\left\lceil \eta \frac{\log\log n}{|\log(\tau-2)|}\right\rceil} F_L^{(-1)}(\mathrm{e}^{-1/(\tau-2)^i}). \tag{5.30}$$

By choosing $\eta \in (0, 1)$ in (5.22) such that $1/\eta < 1 + \varepsilon/3$ so we obtain that

$$d_L(q_{K_n}, q^{\star}) \leq (1+\varepsilon/3) \sum_{i=1}^{\left\lfloor \frac{\log\log n}{|\log(\tau-2)|}\right\rfloor} F_L^{(-1)}\left(\mathrm{e}^{-\left(\frac{1}{\tau-2}\right)^i}\right). \tag{5.31}$$

Finally, it is not hard to see that $3\xi(n^{\alpha}(\tau-2), n^{1/2}) \leq a_n\varepsilon/3$ holds as well for all large enough n. Combining everything, we arrive at

$$\begin{aligned} \mathrm{d}_L(u, v) &\leq \sum_{q\in\{u,v\}} \big(\mathrm{d}_L(q, q_{K_n}) + \mathrm{d}_L(q_{K_n}, q^{\star})\big) + \mathrm{d}_L(u^{\star}, v^{\star}) \\ &\leq 2a_n(\varepsilon/3 + (1+\varepsilon/3)) + a_n\varepsilon/3 \leq 2a_n(1+\varepsilon). \end{aligned} \tag{5.32}$$

This finishes the proof of (5.20). For the second statement, recall that for some $M \geq 1/|\log(\tau-2)|$, $M_n = M\log\log(\widetilde{K}_n)$ and note that the number of edges on the constructed path is at most

$$2M\log\log\widetilde{K}_n + 2\frac{\log\log n - \log\log K_n}{|\log(\tau-2)|} + 3, \tag{5.33}$$

where the relation between $\widetilde{K}_n$ and K_n is described in Lemma 4.5 in (4.10). From (4.10) it is elementary to check that for all n large enough

$$\log\log K_n = \log\log\widetilde{K}_n + \log(1 - c(\log\widetilde{K}_n)^{\eta-1}) = \log\log\widetilde{K}_n + o(1), \tag{5.34}$$

thus, writing $M := (1+z)/|\log(\tau-2)|$ for some $z > 0$, the number of edges in the constructed path is at most

$$\frac{2\log\log n + 2z\log\log \widetilde{K}_n + o(1)}{|\log(\tau-2)|} + 3 \le (1+\varepsilon)\frac{2\log\log n}{|\log(\tau-2)|}, \tag{5.35}$$

as desired. □

Proof of Theorem 2.4 Lemma 5.1 states the proof of the lower bound and Lemma 5.3 the proof of the upper bound. These combined prove the statement of the theorem. □

5.1 Erased Configuration Model

In this section we prove Theorem 2.5.

Proof of Theorem 2.5, lower bound The strategy of the proof is the following: first we show that the lower bound is also valid in the erased model. Then, we show that the constructed paths in the proof of the upper bound between vertices q, q_{K_n} and $q_{K_n}, q^\star$ are whp simple for $q \in \{u, v\}$, and as a result they survive the erasing procedure whp. Finally, we connect $u^\star, v^\star$ in the erased model in some other way than that in the original model.

First we start with the lower bound. The proof of Lemma 5.1 consists of a BFS exploration around the two vertices u and v. These explorations can whp be coupled to two BP trees and therefore all edges within these trees are whp simple. So this lemma remains valid after erasure and thus the lower bound follows both for the weighted distance as well as for the hopcount. □

In the proof of the upper bound we again use a coupling to BP trees to find u_{K_n}, v_{K_n}. Thus, the path between q, q_{K_n} is again whp simple and thus it survives erasure. Next we investigate the constructed path between $q_{K_n}, q^\star$. This path is constructed in the percolated graph. The erasure happens *before* the degree-dependent percolation, so edges of the path constructed in Proposition 5.2 could in principle be deleted earlier in the erasure procedure. We show that the edges on the constructed path were not part of a multiple edge whp, meaning that they were whp not erased before. For this, we state a lemma that gives a bound on the original degree of a vertex, given its percolated degree d^r. This lemma is the 'reverse' of Lemma 4.5.

Claim 5.4 (Degree after percolation vs original degree). *Apply half-edge percolation as described in Definition 4.2 with percolation function $p(d)$ satisfying (4.1) on $\mathrm{CM}_n(\boldsymbol{d})$. Let $\omega(n)$ be an arbitrary sequence that tends to infinity with n. Let $s(x)$ be defined as in (4.5). Then, for a vertex $w \in \mathrm{CM}_n(\boldsymbol{d})$,*

$$\mathbb{P}(d_w \ge s(x) \mid d_w^r \le x) \le c\exp\{-x/4\} \tag{5.36}$$

for some $c > 0$.

Proof The proof directly follows from Bayes' theorem applied to the lhs of (5.36), and following the calculations between (4.5) and (4.7). □

Lemma 5.5 (No multiple edges on the path $q_{K_n}, q^\star$). *Let u_i and u_{i+1} be two consecutive vertices on the constructed path between $u_{K_n}, u^\star$ in Proposition 5.2 and $i = 0, \dots i^{\max} - 1$, then*

$$\lim_{n\to\infty} \mathbb{P}(\ge 2 \text{ edges connecting } u_i \leftrightarrow u_{i+1} | \ge 1 \text{ edge connecting } u_i \leftrightarrow u_{i+1}) = 0. \tag{5.37}$$

Proof Note that the path in Prop. 5.2 is later, in the proof of Theorem 2.4 is constructed in the *percolated graph*. Thus u_i, the ith vertex on this path has percolated degree at least as in (5.4). Without loss of generality we can assume that

$$d^r_{u_i} \le \left(K_n^{1-\delta_n}\right)^{\left(\frac{1}{\tau-2}\right)^{i+1}} =: y_{i+1}, \tag{5.38}$$

since otherwise the path has 'jumped' a layer and one can consider the path to be shorter by an edge. Recall that $d_{u_{i_{\max}}} \ge n^{\alpha(\tau-2)}$ holds as well. Applying Claim 5.4 on $(u_i)_{i\le i_{\max}-1}$, using the upper bound in (5.38),

$$\mathbb{P}\left(\exists i \le i_{\max}-1, d_{u_i} \ge s(y_{i+1}) \mid d^r_{u_i} \le y_{i+1} \forall i \le i_{\max}-1\right) \le \sum_{i=0}^{i_{\max}-1} c\exp\{-y_{i+1}/4\} \tag{5.39}$$

which tends to zero with n since it is a constant times the first term.

We can rewrite the probability in (5.37) as

$$\frac{1-\mathbb{P}(1 \text{ edge } u_i \leftrightarrow u_{i+1}) - \mathbb{P}(u_i \not\leftrightarrow u_{i+1})}{1-\mathbb{P}(u_i \not\leftrightarrow u_{i+1})}. \tag{5.40}$$

We investigate the probabilities in (5.40) separately starting with the probability that there is exactly one edge between those two vertices. We lower bound the probability that precisely the jth half-edge of u_i connects to u_{i+1}, and the others do not. Note that for the kth half-edge the probability of not connecting to u_{i+1} is at least $(\mathcal{H}_n - d_{u_{i+1}} - 2(k-1))/(\mathcal{H}_n - 2(k-1)-1) \ge 1 - d_{v_{i+1}}/\mathcal{H}_n$. Thus

$$\begin{aligned}\mathbb{P}(1 \text{ edge } u_i \leftrightarrow u_{i+1}) &\ge \sum_{j=1}^{d_{u_i}} \frac{d_{u_{i+1}}}{\mathcal{H}_n - 2d_{u_i}} \prod_{k=1}^{d_{u_i}-1}\left(1-\frac{d_{u_{i+1}}}{\mathcal{H}_n}\right) \\ &\ge \frac{d_{u_{i+1}}d_{u_i}}{\mathcal{H}_n - 2d_{u_i}}\left(1-\frac{d_{u_{i+1}}}{\mathcal{H}_n}\right)^{d_{u_i}}. \end{aligned} \tag{5.41}$$

Next we bound the probability that there is no edge between two consecutive vertices, both from above and below.

$$\mathbb{P}(u_i \not\leftrightarrow u_{i+1}) \le \prod_{k=1}^{\lfloor d_{u_i}/2\rfloor}\left(1-\frac{d_{u_{i+1}}}{\mathcal{H}_n - 2d_{u_i}}\right) = \left(1-\frac{d_{u_{i+1}}}{\mathcal{H}_n - 2d_{u_i}}\right)^{d_{u_i}/2}, \tag{5.42}$$

$$\mathbb{P}(u_i \not\leftrightarrow u_{i+1}) \ge \prod_{k=1}^{d_{u_i}}\left(1-\frac{d_{u_{i+1}}}{\mathcal{H}_n}\right) = \left(1-\frac{d_{u_{i+1}}}{\mathcal{H}_n}\right)^{d_{u_i}}. \tag{5.43}$$

Using series expansion for Eqs. (5.41)–(5.43), we obtain an upper bound on (5.40):

$$\frac{1-\mathbb{P}(1 \text{ edge } u_i \leftrightarrow u_{i+1}) - \mathbb{P}(u_i \not\leftrightarrow u_{i+1})}{1-\mathbb{P}(u_i \not\leftrightarrow u_{i+1})} \le \frac{2d_{u_i}d_{u_{i+1}}}{\mathcal{H}_n} \Big/ \left(1-\frac{d_{u_i}d_{u_{i+1}}}{4(\mathcal{H}_n - d_{u_i})}\right). \tag{5.44}$$

By (5.39), whp, $d_{u_{i+1}} \le s(y_{i+1})$, and further, by the definition of $i_{\max}$ in (5.18), and $s(\cdot)$ in (4.5), $s(y_{i_{\max}-k}) \le n^{\alpha(\tau-2)^k(1+o(1))}$ for $k \in \{1,2\}$. Thus, for all $i \le i_{\max}-2$, whp

$$\frac{d_{u_i}d_{u_{i+1}}}{\mathcal{H}_n} \le c\frac{s(y_{i_{\max}-2})s(y_{i_{\max}-1})}{n} \le n^{\alpha(\tau-2)(\tau-1)(1+o(1))-1}.$$

The rhs converges to zero as n tends to infinity as long as $\alpha < ((\tau-2)(\tau-1))^{-1}$, which we have assumed in Assumption 2.1. □

Proof of Theorem 2.5, upper bound As mentioned before, we construct a path in $\mathrm{ECM}_n(\boldsymbol{d}, \boldsymbol{L})$ to connect u, v. For this it is enough to construct a path with all its edges begin simple edges in $\mathrm{CM}_n(\boldsymbol{d}, \boldsymbol{L})$. This path has a huge overlap with the path in the upper bound of Theorem 2.4. Namely, the segments between u, u_{K_n} and v, v_{K_n} are whp using simple edges by the coupling to BP trees. The segments between $u_{K_n}, u_{i_{\max}-1}$ and $v_{K_n}, v_{i_{\max}-1}$ are whp using simple edges again so they survives erasure. Next we connect $u_{i_{\max}-1}$ to $v_{i_{\max}-1}$. Note that the constructed path in $\mathrm{CM}_n(\boldsymbol{d}, \boldsymbol{L})$ might use multiple edges so we need a different connecting path. However, $q_{i_{\max}-1}$ for $q \in \{u, v\}$ are vertices with degree at least $n^{\alpha(\tau-2)(1+o(1))}$. In the proof of Theorem 2.4, we created a 3-hop connection between $u_{i_{\max}} = u^\star$ and $v_{i_{\max}} = v^\star$ in the *percolated* graph, see (5.25)–(5.26). When we erase a multiple edge, we keep one edge independently of its edge-length. Thus, from every multiple edge at least one edge remains. Hence, an analogous construction as in (5.25)–(5.26) can be repeated, not for the percolated graph but for the original graph, developing a 5-hop connection between $u_{i_{\max}-1}, v_{i_{\max}-1}$. The edge-lengths on this path are simply i.i.d. copies of L. Thus,

$$d_L^e(u,v) \le d(u,u_{K_n}) + d(v,v_{K_n}) + 2\sum_{i=0}^{i_{\max}-2} \xi(y_i, y_{i+1}) + \sum_{i=1}^{5} L_i. \tag{5.45}$$

For all $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}\left(\sum_{i=1}^{5} L_i \le \varepsilon/3 \sum_{i=1}^{\left\lfloor \frac{\log\log n}{|\log(\tau-2)|} \right\rfloor} F_L^{(-1)}\left(\mathrm{e}^{-\left(\frac{1}{\tau-2}\right)^i} \right) \right) = 0.$$

Then we treat the terms in (5.45) similarly as we did in the proof of Theorem 2.4 (see (5.23) and (5.27)–(5.31)) finishes the proof. □

Acknowledgements This work is supported by the Netherlands Organisation for Scientific Research (NWO) through VENI grant 639.031.447 (JK).

References

1. Achard, S., Salvador, R., Whitcher, B., Suckling, J., Bullmore, E.: A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. J. Neurosci. **26**(1), 63–72 (2006)
2. Aiello, W., Bonato, A., Cooper, C., Janssen, J., Prałat, P.: A spatial web graph model with local influence regions. Internet Math. **5**(1–2), 175–196 (2008)
3. Albert, R., Barabási, A.-L.: Statistical mechanics of complex networks. Rev. Mod. Phys. **74**(1), 47 (2002)
4. Amini, O., Devroye, L., Griffiths, S., Olver, N.: On explosions in heavy-tailed branching random walks. Ann. Probab. **41**(3B), 1864–1899 (2013)
5. Barabási, A.-L., Albert, R.: Emergence of scaling in random networks. Science **286**(5439), 509–512 (1999)
6. Barabási, A.-L., Albert, R., Jeong, H.: Scale-free characteristics of random networks: the topology of the world-wide web. Phys. A **281**(1), 69–77 (2000)

7. Baroni, E., Hofstad, R.v.d., Komjáthy, J.: Tight fluctuations of weight-distances in random graphs with infinite-variance degrees. arXiv:1609.07269 (2016)
8. Baroni, E., Hofstad, Rvd, Komjáthy, J.: Nonuniversality of weighted random graphs with infinite variance degree. J. Appl. Probab. **54**(1), 146–164 (2017)
9. Bender, E.A., Canfield, E.R.: The asymptotic number of labeled graphs with given degree sequences. J. Comb. Theory Ser. A **24**(3), 296–307 (1978)
10. Bhamidi, S., Hofstad, Rvd, Hooghiemstra, G.: Extreme value theory, poisson-dirichlet distributions, and first passage percolation on random networks. Adv. Appl. Probab. **42**(3), 706–738 (2010a)
11. Bhamidi, S., Hofstad, Rvd, Hooghiemstra, G.: First passage percolation on random graphs with finite mean degrees. Ann. Appl. Probab. **20**(5), 1907–1965 (2010b)
12. Bhamidi, S., Hofstad, Rvd, Hooghiemstra, G.: First passage percolation on the Erdős-Rényi random graph. Comb. Probab. Comput. **20**, 683–707 (2011)
13. Bhamidi, S., Hofstad, R.v.d., Hooghiemstra, G.: Universality for first passage percolation on sparse random graphs. Ann. Probab. **45**(4), 2568–2630, 07 (2017)
14. Boguná, M., Papadopoulos, F., Krioukov, D.: Sustaining the Internet with hyperbolic mapping. Nat. Commun. **1**(62) (2010)
15. Bollobás, B.: A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. Eur. J. Comb. **1**(4), 311–316 (1980)
16. Bringmann, K., Keusch, R., Lengler, J.: Geometric inhomogeneous random graphs. arXiv:1511.00576 (2015)
17. Bullmore, E., Sporns, O.: Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. **10**(3), 186 (2009)
18. Chung, F., Lu, L.: The diameter of sparse random graphs. Adv. Appl. Math. **26**(4), 257–279 (2001)
19. Davies, P.L.: The simple branching process: a note on convergence when the mean is infinite. J. Appl. Probab. **15**(3), 466–480 (1978)
20. Deijfen, M., Hofstad, R.v.d., Hooghiemstra, G.: Scale-free percolation. Annales de l'Institut Henri Poincaré, Probabilités et Statistiques, **49**(3), 817–838 (2013)
21. Esker, H.v.d., Hofstad, R.v.d., Hooghiemstra, G., Znamenski, D.: Distances in random graphs with infinite mean degrees. Extremes **8**(3), 111–141 (2005)
22. Esker, H.v.d., Hofstad, R.v.d., Hooghiemstra, G.: Universality for the distance in finite variance random graphs. J. Stat. Phys. **133**(1), 169–202 (2008)
23. Faloutsos, M., Faloutsos, P., Faloutsos, C.: On power-law relationships of the internet topology. ACM SIGCOMM Comput. Commun. Rev. **29**(4), 251–262 (1999)
24. Hammersley, J.M., Welsh, D.: First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory. In: Bernoulli 1713 Bayes 1763 Laplace 1813, pp. 61–110. Springer (1965)
25. Hofstad, R.v.d.: Random Graphs and Complex Networks, Vol. I. Cambridge University Press (2016)
26. Hofstad, R.v.d.: Random Graphs and Complex Networks, Vol. II. Cambridge University Press (2016). to appear
27. Hofstad, R.v.d., Hooghiemstra, G., Van Mieghem, P.: Distances in random graphs with finite variance degrees. Random Struct. Algorithms **27**(1), 76–123 (2005)
28. Hofstad, R.v.d., Hooghiemstra, G., Znamenski, D.: Distances in random graphs with finite mean and infinite variance degrees. Electron. J. Probab. **12**(25), 703–766 (2007)
29. Hofstad, R.v.d., Komjáthy, J.: When is a scale-free graph ultra-small? J. Stat. Phys. **169**(2), 223–264 (2017)
30. Jacob, E., Mörters, P.: A spatial preferential attachment model with local clustering. In: International Workshop on Algorithms and Models for the Web-Graph, pp. 14–25. Springer (2013)
31. Janson, S.: On percolation in random graphs with given vertex degrees. Electron. J. Probab. **14**, 86–118 (2009)
32. Janson, S., Luczak, M.J.: A new approach to the giant component problem. Random Struct. Algorithms **34**(2), 197–216 (2009)
33. Kolossváry, I., Komjáthy, J.: First passage percolation on inhomogeneous random graphs. Adv. Appl. Probab. **47**(2), 589–610 (2015)
34. Komjáthy, J.: Explosive crump-mode-jagers branching processes. arXiv preprint arXiv:1602.01657 (2016)
35. Milgram, S.: The small world problem. Psychol. Today May, 60–67 (1967)
36. Montoya, J.M., Solé, R.V.: Small world patterns in food webs. J. Theor. Biol. **214**(3), 405–412 (2002)
37. Newman, M.E.J.: The structure of scientific collaboration networks. Proc. Natl. Acad. Sci. **98**(2), 404–409 (2001)
38. Redner, S.: How popular is your paper? an empirical study of the citation distribution. Eur. Phys. J. B-Condens. Matter Complex Syst. **4**(2), 131–134 (1998)

39. Reittu, H., Norros, I.: On the power-law random graph model of massive data networks. Perform. Eval. **55**(1–2), 3–23 (2004)
40. Watts, D.J., Strogatz, S.H.: Collective dynamics of 'small-world' networks. Nature **393**, 440–442 (1998)