CrossMark

When is a Scale-Free Graph Ultra-Small?

Remco van der Hofstad[1] · **Júlia Komjáthy**[1]

Abstract In this paper we study typical distances in the configuration model, when the degrees have asymptotically infinite variance. We assume that the empirical degree distribution follows a power law with exponent $\tau \in (2, 3)$, up to value n^{β_n} for some $\beta_n \gg (\log n)^{-\gamma}$ and $\gamma \in (0, 1)$. This assumption is satisfied for power law i.i.d. degrees, and also includes *truncated power-law* empirical degree distributions where the (possibly exponential) truncation happens at n^{β_n}. These examples are commonly observed in many real-life networks. We show that the graph distance between two uniformly chosen vertices centers around $2 \log \log(n^{\beta_n})/|\log(\tau - 2)| + 1/(\beta_n(3 - \tau))$, with tight fluctuations. Thus, the graph is an *ultrasmall world* whenever $1/\beta_n = o(\log \log n)$. We determine the distribution of the fluctuations around this value, in particular we prove these form a sequence of tight random variables with distributions that show log log-periodicity, and as a result it is non-converging. We describe the *topology and number of shortest paths*: We show that the number of shortest paths is of order $n^{f_n \beta_n}$, where $f_n \in (0, 1)$ is a random variable that oscillates with n. We decompose shortest paths into three segments, two 'end-segments' starting at each of the two uniformly chosen vertices, and a middle segment. The two end-segments of any shortest path have length $\log \log(n^{\beta_n})/|\log(\tau - 2)|$+tight, and the total degree is increasing towards the middle of the path on these segments. The connecting middle segment has length $1/(\beta_n(3 - \tau))$+tight, and it contains only vertices with degree at least of order $n^{(1-f_n)\beta_n}$, thus all the degrees on this segment are comparable to the maximal degree. Our theorems also apply when instead of truncating the degrees, we start with a configuration model and we *remove* every vertex with degree at least n^{β_n}, and the edges attached to these vertices. This sheds light on the *attack vulnerability* of the configuration model with infinite variance degrees.

✉ Júlia Komjáthy
  j.komjathy@tue.nl

  Remco van der Hofstad
  rhofstad@win.tue.nl

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

1 Introduction and results

Many real-world networks are claimed to be *small worlds*, meaning that their graph distances
are quite small. In social networks, such small distances go under the name of the 'six-degrees-
of-separation' paradigm and have attracted attention due to the interesting experiments by
Milgram [41,59]. See also Pool and Kochen [55] as well as [17], where a related experiment is
described on the basis on email messages. After this start in social sciences, the small-world
nature of many other networks was first described by Strogatz and Watts [62]. A popular
account of small-world aspects of networks can be found in the book by Watts [61]. See also
the surveys by Newman [49] and Albert and Barabási [2] on real-world networks.

The reported small-world nature of many real-world networks has incited a deep and
thorough study of typical distances in random graphs. See the highly influential paper by
Newman, Strogatz and Watts [51], who pioneered this line of research. There is a deep
relation between the small-world nature of networks and their other often reported common
feature, the *scale-free paradigm*, which states that the proportion of vertices of degree k in
many networks scales as an inverse power of k for k large. This scale-free nature implies
that there are many vertices with very high degrees, and these *hubs* drastically shrink graph
distances. The common picture in many random graph models is that graph distances are
asymptotically *logarithmic* in the graph size when the degrees have finite-variance degrees
[23,30,51], while they are *doubly logarithmic* or *ultrasmall* when the graphs have infinite-
variance degrees [11,12,15,18,31,52]. See also [29] for a discussion of many of the available
results. The conclusion is that typical distances in random graphs are closely related to their
degree structure, and larger degrees significantly shrink graph distances.

In many real-world networks, not only power-law degree sequences are observed, but also
power laws with a so-called *exponential truncation*. This means that, even though the degree
distribution for small values is close to a power law, for large values the tail distribution
becomes exponentially small. This occurs e.g. in sexual networks [39], the Internet Movie
Data base [4], and in scientific collaboration networks [47,48]. Of course, it is not easy
to guess whether a distribution has a power law, or rather a power law with exponential
truncation. Newman and collaborators give sensible suggestions on how to approach these
issues in real-world networks [13,50].

The value above which the exponential decay starts to set in is the *truncation parameter*.
Naturally, when the degrees already had finite variance before truncation, they will remain
to have so after truncation, so nothing much happens and graph distances ought to behave
as in the finite-variance setting in [30]. Further, any power-law distribution with exponential
truncation with a *bounded* truncation parameter has all moments, and thus distances ought
to become *logarithmic* as in the finite-variance case in e.g., [30], accounting to 'strictly'
small-world networks [62]. The situation changes dramatically when dealing with infinite-
variance degrees, that is, when the power-law exponent is below 3. Indeed, in this setting,
it is well known that typical paths realizing the graph distance pass through the vertices
of highest possible degrees (i.e., the *hubs*) and typical distances grow much slower, *doubly-
logarithmically* with the size of the graph, accounting to *ultrasmall worlds* [31,53]. However,
truncating the degrees could possibly have a dramatic effect, and could possibly increase

the graph distances rather substantially. In fact, even though the asymptotic variance of the degrees remains infinite after truncation, distances might grow substantially with the truncation and the graph might fail to be an ultra-small world.

The main aim of this paper is to quantify the effect of truncation of the degrees in random graphs models, in particular, in the configuration model that we define below. For simplicity, we state our result as a theorem without bothering to explain the somewhat tedious details and conditions. Below, we give several more accurate and detailed versions of this theorem.

Theorem *Let us consider the configuration model on n vertices with empirical degree distribution that follows a power law with exponent $\tau \in (2, 3)$, satisfying some appropriate regularity assumptions and truncated at degrees n^{β_n}, where $\beta_n (\log n)^\gamma \to \infty$ for some $\gamma \in (0, 1)$. Let $d_G(v_r, v_b)$ denote the graph distance between two uniformly chosen vertices v_r, v_b. Then*

$$d_G(v_r, v_b) - 2 \frac{\log \log(n^{\beta_n})}{|\log(\tau - 2)|} - \frac{1}{(3 - \tau)\beta_n} \tag{1.1}$$

is a tight sequence of random variables.

For the precise result see Theorem 1.8 below. In that theorem, we also identify the distributional limit of the tight sequence of random variables along subsequences. This was first done in [31] using different methods for the special case $\beta_n \equiv 1/(\tau - 1)$. We show in Section 6 that the seemingly different formulations, Theorem 1.8 and the result in [31] are actually the same.

Note that as soon as $\beta_n = o(1/\log \log n)$, the term containing $1/\beta_n$ becomes dominant, and the second order term is of order $\log \log n$. Thus, the random graph fails to be an ultrasmall world in a strict sense as soon as the truncation happens at degrees at most $n^{o(1/\log \log n)}$. However, when $\beta_n \log \log n \to \infty$, the dominant term is the first one, of order $\log \log n$. This theorem shows that distances in truncated power-law graphs with infinite asymptotic variance *interpolate* between ultrasmall worlds and small worlds.

This sheds light to a discussion in the physics literature [15,19,24,25,33,51] about the validity of the formula derived using generating function methods or n-dependent branching process approximations, stating that

$$d_G(v_r, v_b) = \frac{\log n}{\log \nu_n} + \text{tight} = \frac{1}{(3 - \tau)\beta_n} + \text{tight} \tag{1.2}$$

with ν_n standing for the empirical second moment[1] of the degrees and β_n is the truncation exponent. Compare this formula to the one in (1.1) and note that this formula yields the first order term if and only if $\beta_n = o(1/\log \log n)$, while even in this case it fails to capture the second order term, which is of order $\log \log n$.

While [51] questions the validity of this formula for $\tau \in (2, 3)$, [24,25,33] argue that a constant $\beta_n \equiv \beta$ for the truncation exponent yields bounded typical distances in this regime, in agreement with what (1.2) suggests. This contradicts the arguments in [15] where the authors show that the smallest achievable order for typical distances is $\log \log n$.

Closest to our result is the work of Dorogovtsev et al. [19], who, using (non-rigorous) generating function methods, study typical distances in the configuration model with truncated power-law distributions of the form $\mathbb{P}(D = k) \sim k^{-\tau} \zeta^k$, for constant $\zeta < 1$, and derive that

---

[1] We show in Claim 2.6 and (4.2) below that $\nu_n \sim n^{\beta_n(3-\tau)}$, thus, $\log n / \log \nu_n$ in (1.2) yields $1/(\beta_n(3-\tau))$, justifying the equality between the two expressions.

$$d_G(v_r, v_b) = \frac{\log n}{(3 - \tau)|\log(1 - \zeta)|} + \frac{2\log|\log(1 - \zeta)|}{|\log(\tau - 2)|} + \text{tight},$$

$$= \frac{1}{(3 - \tau)\beta_n} + \frac{2\log\log(n^{\beta_n})}{|\log(\tau - 2)|} + \text{tight}, \tag{1.3}$$

where we took the liberty to set $\zeta := \exp\{-1/n^{\beta_n}\}$ to obtain the second line. Observe that our (rigorous) result in (1.1) above is in perfect agreement with this. The region of validity of their approach [19, (106)] translates to the requirement that $\beta_n = o(1/\log\log n)$, can be interpreted that their method requires that the first term is the leading order term in (1.3). We explain below in Sect. 1.3 why the term $\log\log n$ is missing from (1.2) and where it comes from.

Let us make a comparison: two models with the same maximal degree, a truncated and an un-truncated one. More precisely, take a model with power-law exponent $\tau \in (2, 3)$ and truncation at $\beta_n = \beta < 1/(\tau - 1)$ fixed. In a model with power-law exponent $\tilde{\tau}$ and natural truncation (that is, $\tilde{\beta} = 1/(\tilde{\tau} - 1)$), the maximal degree is $n^{1/(\tilde{\tau}-1)}$. Setting the maximal degrees in the two models, $- n^\beta$ and $n^{1/(\tilde{\tau}-1)}$, respectively – to be equal yields the relation $\tilde{\tau} = 1 + 1/\beta$. Comparing distances in the two models, we see an interesting phenomenon. When $\beta_n < 1/2$, the un-truncated model has $\tilde{\tau} = 1 + 1/\beta > 3$, implying that $\lim_{n\to\infty} \mathbb{E}[(D'_n)^2] < \infty$, which, in turn, implies that distances jump up to *logarithmic* order. In the truncated model distances are of order $\log\log n$, as described by (1.1). When $\beta_n > 1/2$, the leading order of distances in the truncated model is $\log\log n/|\log(\tau - 2)|$+tight, while in the un-truncated model it is $\log\log n/|\log(\tilde{\tau}-2)|+tight = \log\log n/|\log(1/\beta-1)|+tight$. Since $\beta < 1/(\tau - 1)$ by assumption, $1/\beta - 1 > \tau - 2$ and hence the distances in the latter model are larger. This means that 're-parametrizing' the truncated model by another power-law (τ') that would more naturally reflect the maximal degree is not the same as truncating, even the leading term changes. This effect is extreme when $\beta < 1/2$, in which case the truncated and the un-truncated models do not even belong to the same universality class (ultrasmall versus small world).

Let us comment on the choice of model as well. We expect that the same result is true for the giant component of the Chung-Lu or Norros-Reitu model when the power law exponent is $\tau \in (2, 3)$. These models behave qualitatively similarly to the configuration model. In fact, the ultrasmall nature of these networks were pointed out in [52,53]. The independence of the edges conditioned on vertex-weights even makes path-counting methods easier, in fact, we expect that the same proof that we provide here could be applied for these models as well.

Attack vulnerability The removal of all vertices above a certain degree in a network is called a *targeted attack* or *deliberate attack*. An immediate corollary (see Corollary 1.17 below) of our work is that our results remain valid when instead of truncating the degrees, we *remove* all vertices with degree at least $n^{\tilde{\beta}_n}$ from a configuration model. In particular, distances after a targeted attack are described by (1.1), with β_n replaced by $\tilde{\beta}_n$. This theorem also sheds light to how distances gradually grow from ultrasmall to small world when vertices with smaller and smaller degree are gradually removed. A similar analysis have been carried out for a variant of the preferential attachment model in [20], for the special case when all degrees above order $\log n$ (equivalently, the oldest εn vertices for small small $\varepsilon > 0$) are removed. They show that distances in this case grow logarithmically, giving a strong base for our conjecture that the formula in (1.1) should also be valid for $\beta_n = \Theta(1/\log n)$, which is, at least with the methods of this paper, beyond our reach.

1.1 The Model and the Main Result

In this paper we work under the setting of the configuration model $CM_n(\boldsymbol{d})$. In this random graph model, there are n vertices, with prescribed degrees d_v, $v \in \{1, 2, \ldots, n\} := [n]$. To each vertex $v \in [n]$ we assign d_v half-edges and the half-edges are then paired uniformly at random to form edges. We assume that the total number of half-edges $\ell_n := \sum_{v \in [n]} d_v$ is even. We denote the outcome - a graph-valued random variable - by $CM_n(\boldsymbol{d})$.

1.1.1 Setting and Assumptions

We study the case when the empirical degree distribution follows a possibly truncated power law, with an exponent that gives rise to empirical variance tending to infinity with n and when the truncation happens at some polynomial of n. To make this precise, we impose the following three assumptions on the *empirical degree distribution*, $F_n(x) := \frac{1}{n} \sum_{v \in [n]} \mathbb{1}_{\{d_v \leq x\}}$:

Assumption 1.1 (*Power-law tail behavior*) There exists a $\beta_n \in (0, 1/(\tau - 1)]$ such that for all $\varepsilon > 0$, $F_n(x) = 1$ for $x \geq n^{\beta_n(1+\varepsilon)}$, while for all $x \leq n^{\beta_n(1-\varepsilon)}$,

$$1 - F_n(x) = \frac{L_n(x)}{x^{\tau-1}}, \tag{1.4}$$

with $\tau \in (2, 3)$, and a function $L_n(x)$ that satisfies for some constant $C_1 > 0$ and $\eta \in (0, 1)$ that

$$\exp\{-C_1 (\log x)^\eta\} \leq L_n(x) \leq \exp\{C_1 (\log x)^\eta\}. \tag{1.5}$$

Assumption 1.2 (*Minimal degree at least 2*) $\min_{v \in [n]} d_v \geq 2$.

See Remark 1.12 for an extension of our results in the case when $\min_{v \in [n]} d_v = 1$. We write D_n for a random variable having distribution F_n. Then D_n is the degree of a uniformly chosen vertex from $[n]$. We introduce $D_n^\star := (\text{the } size\text{-}biased \text{ } version \text{ of } D_n) - 1$ by

$$\mathbb{P}(D_n^\star = j) := \frac{j+1}{\ell_n} \sum_{v \in [n]} \mathbb{1}_{\{d_v = j+1\}} = \frac{(j+1)\mathbb{P}(D_n = j+1)}{\mathbb{E}[D_n]}, \quad j \geq 0. \tag{1.6}$$

We write $F_n^\star(x)$ for the distribution function of $D_n^\star$. Note that for all $x \leq n^{\beta_n(1-\varepsilon)}$,

$$1 - F_n^\star(x) = \frac{1}{\mathbb{E}[D_n]} \sum_{j \geq x} (j+1)\mathbb{P}(D_n = j+1) \geq \frac{1}{\mathbb{E}[D_n]} x(1 - F_n(x)),$$

and similarly, under Assumption 1.1,

$$1 - F_n^\star(x) \leq (x+1)(1 - F_n(x+1)) + \sum_{j \geq x}(1 - F_n(x))$$

$$\leq \frac{2}{\mathbb{E}[D_n]} \frac{L_n(x+1)}{x^{\tau-2}} + \frac{1}{\mathbb{E}[D_n]} \sum_{j \geq x} \frac{L_n(j)}{j^{\tau-1}}.$$

By a Karamata-type theorem (see [9, Proposition 1.5.10]) the latter sum on the rhs is at most $2L_n(x)/(x^{\tau-2}(\tau - 2))$ for all large enough x. Further, for $x \geq n^{\beta_n(1+\varepsilon)}$ it is obvious from the definition (1.6) that $1 - F_n^\star(x) = 0$. Thus, it follows that for all $\varepsilon > 0$ and $x \leq n^{\beta_n(1-\varepsilon)}$, $F_n^\star$ satisfies that

$$1 - F_n^\star(x) = \frac{L_n^\star(x)}{x^{\tau-2}}, \tag{1.7}$$

with a function $L_n^\star(x)$ satisfying (1.5) again (possibly with a different constant $C_1^\star$ instead of C_1 in the exponent in (1.5) for $L_n^\star$).

To be able to state convergence results, we will need an assumption that relates the behavior of F_n and $F_n^\star$ for different values of n to a limiting distribution function. We write $d_{\mathrm{TV}}(F, G) := \frac{1}{2} \sum_{x \in \mathbb{N}} |F(x+1) - F(x) - (G(x+1) - G(x))|$ for the total variation distance between two (discrete) probability measures. The weakest form of such assumption that we can pose is captured in the following assumption:

Assumption 1.3 (*Convergence to limiting distributions*) We assume that there exist distribution functions $F(x)$, $F^\star(x)$ such that $F_n(x) \to F(x)$ and $F_n^\star(x) \to F^\star(x)$ in all continuity points of $F(x)$, $F^\star(x)$. We assume that there exists a $\kappa > 0$,

$$\max\{d_{\mathrm{TV}}(F_n, F), d_{\mathrm{TV}}(F_n^\star, F^\star)\} \leq n^{-\beta_n \kappa}. \tag{1.8}$$

Let us write D, $D^\star$ for random variables following the limiting distributions F, $F^\star$ in Assumption 1.3, respectively. Since total variation convergence equals weak convergence for discrete random variables, $D_n \xrightarrow{d} D$ and $D_n^\star \xrightarrow{d} D^\star$. Using (1.6), we further obtain that

$$\mathbb{P}(D^\star = j) = \frac{(j+1)\mathbb{P}(D = j)}{\mathbb{E}[D]}, \tag{1.9}$$

thus $D^\star$ is the (size-biased version-1) of D. Further note that F can be written in the form (1.4), and $F^\star$ as in (1.7) with limiting L, $L^\star$ that satisfies (1.5). Note that the limit variables are not truncated. Further, the bound $n^{-\beta_n \kappa}$ in Assumption 1.3 is best possible since $d_{\mathrm{TV}}(F_n, F) \geq \mathbb{P}(D > n^{\beta_n}) \geq n^{-\beta_n(\tau-1)}\ell(n^{-\beta_n(\tau-1)}) \geq n^{-\beta_n(\tau-1-\delta)}$. It is also reasonable, e.g. it can be shown that it is satisfied in Examples 1.18-1.21 below.

Under Assumptions 1.1 and 1.2, the graph almost surely has a unique connected component of size $n(1 - o_{\mathbb{P}}(1))$ see e.g. [28, Vol II., Theorem 4.1] or [42,43], or the recent paper [22].

We provide three examples in Sect. 1.2 below (i.i.d. degrees, exponential and hard truncation) that satisfy Assumption 1.1, see Examples 1.18, 1.20, 1.21 below, as well as collect some references to networks following such empirical degree distributions.

In this paper we study typical distances, that is, the graph distance d_G between *two uniformly chosen vertices* in the graph. For the sake of the proof, we denote these vertices by v_r and v_b and think of them as being red and blue, respectively.

Definition 1.4 (*With high probability*) We say that a sequence of events $\mathcal{E}_n$ happens with high probability under the measure $\mathbb{Q}$ (and abbreviate this as $\mathbb{Q}$-whp), if $\mathbb{Q}(\mathcal{E}_n) \to 1$ as $n \to \infty$. We write simply 'whp' when the measure is the annealed measure of the configuration model and the two uniformly chosen vertices v_r, v_b.

We emphasize that in the setting of this paper, whp statements should be read as follows: for asymptotically almost every realizations of the random graph $\mathrm{CM}_n(\boldsymbol{d})$ and almost all pairs of vertices (v_r, v_b), the statement is true.

Definition 1.5 ($\sim$ *notation*) We use the shorthand notation $X_n \sim a_n$ if there exists a constant $\theta \in (0, 1)$ such that

$$X_n \sim a_n \iff \mathbb{P}\left(X_n \in \left[a_n \mathrm{e}^{-(\log a_n)^\theta}, a_n \mathrm{e}^{(\log a_n)^\theta}\right]\right) \to 1. \tag{1.10}$$

We call vertices with degree at least $\sim n^{(\tau-2)\beta_n}$ *hubs*.

Note that $X_n \sim n^a$ is somewhat stronger than stating that $X_n = n^{a(1+o_{\mathbb{P}}(1))}$.

The statement of the main theorem uses some knowledge about infinite-mean branching processes, as well as their coupling to the local neighborhood of the two vertices v_r, v_b. So, before stating the result, we have to do a small excursion into defining these objects. In particular, Lemma 2.2 and Corollary 2.3 below, based on [7, Proposition 4.7], states that under Assumption 1.3, whp, the number of vertices and their forward degrees in an exploration of the neighborhood of v_r, v_b can be coupled to two independent branching processes (that are embedded in the graph disjointly, and have offspring distribution $F^\star$ for the second and further generations, and with offspring distribution given by F for the first generation), as long as the total number of vertices of the explored clusters does not exceed n^ϱ for some small $\varrho > 0$. Let us do the exploration in a breadth-first-search manner, and then the exploration at time t contains all the vertices with at most distance t away from $\mathcal{C}_0^{(r)} := \{v_r\}$ and $\mathcal{C}_0^{(b)} := \{v_b\}$. We shall denote these clusters by $\mathcal{C}_t^{(r)}$, $\mathcal{C}_t^{(b)}$ (i.e., the vertices and their graph structure), and think of them as being colored red and blue, respectively. Similarly, we denote the number of vertices in the kth generation of the coupled branching processes by $(Z_k^{(r)}, Z_k^{(b)})_{k>0}$. The next definition, describing the *double-exponential growth rates* of these neighborhoods, uses this coupling:

Definition 1.6 (*Double-exponential growth rates of local neighborhoods*) Let $(Z_k^{(r)}, Z_k^{(b)})$ denote the number of individuals in the kth generation of the two independent copies of a Galton-Watson process, coupled to the breadth-first-search exploration process of the neighborhoods of v_r and v_b in the configuration model. In these branching processes, the size of the first generation has distribution F, and all further generations have offspring distribution $F^\star$ from Assumption 1.3. Then, for some $\varrho_n' < (\tau - 2) \min \left\{ \beta_n\kappa, (1 - \beta_n(1 + \varepsilon))/2, (\tau - 2 - 2\varepsilon)/(2(\tau - 1)) \right\}$, let us define

$$Y_r^{(n)} := (\tau - 2)^{t(n^{\varrho_n'})} \log \left(Z_{t(n^{\varrho_n'})}^{(r)} \right), \quad Y_b^{(n)} := (\tau - 2)^{t(n^{\varrho_n'})} \log \left(Z_{t(n^{\varrho_n'})}^{(b)} \right), \quad (1.11)$$

where $t(n^{\varrho_n'}) = \inf_k \{\max\{Z_k^{(r)}, Z_k^{(b)}\} \geq n^{\varrho_n'}\}$. Let us further introduce

$$Y_r := \lim_{k \to \infty} (\tau - 2)^k \log \left(Z_k^{(r)} \right), \quad Y_b := \lim_{k \to \infty} (\tau - 2)^k \log \left(Z_k^{(b)} \right). \quad (1.12)$$

Note that the limit variables in (1.12) are independent of ρ_n'. Further, note that $(Y_r^{(n)}, Y_b^{(n)})$ is a subsequence of the convergent sequence $\left((\tau - 2)^k \log(Z_k^{(r)}), (\tau - 2)^k \log(Z_k^{(b)}) \right)$, taken at the subsequence $k_n := t(n^{\rho_n'})$. Since for any $\rho' > 0$, $t(n^{\rho_n'}) \to \infty$ as $n \to \infty$, under Assumption 1.3 we shall obtain that $(Y_r^{(n)}, Y_b^{(n)}) \xrightarrow{d} (Y_r, Y_b)$ as $n \to \infty$. When (and only when) $\beta_n = 1/(\tau - 1)$, we shall need one more assumption that concerns the limiting distribution of Y_r, Y_b:

Assumption 1.7 (*No pointmass of the measure of Y*) We assume that the limiting random variable $Y := \lim_{k \to \infty} (\tau - 2)^k \log(\max\{Z_k, 1\})$ of the branching process in Definition 1.6 has no point-mass on $(0, \infty)$.

The criteria on F in Assumption 1.3 required for this assumption to hold are not obvious. According to our knowledge, no necessary and sufficient condition for no point mass or absolute continuity can be found in the literature. A sufficient criterion for absolute continuity of Y is given in [56,57].

To be able to state the results shortly, let us define the σ-algebra generated by the induced subgraph on $\mathcal{C}^{(r)}_{t(n\varrho'_n)} \cup \mathcal{C}^{(b)}_{t(n\varrho'_n)}$:

$$\mathcal{F}_{\varrho'_n} := \sigma\left(\mathcal{C}^{(r)}_{t(n\varrho'_n)} \cup \mathcal{C}^{(b)}_{t(n\varrho'_n)}\right) \tag{1.13}$$

and introduce the shorthand notation[2]

$$\mathbb{P}_Y(\cdot) := \mathbb{P}(\cdot|\mathcal{F}_{\varrho'_n}), \quad \mathbb{E}_Y(\cdot) := \mathbb{E}(\cdot|\mathcal{F}_{\varrho'_n}) \tag{1.14}$$

Further define, for $q \in \{r, b\}$,

$$T_q(\beta_n) := \left\lfloor \frac{\log\log(n^{\beta_n}) - \log(Y_q^{(n)})}{|\log(\tau - 2)|} \right\rfloor - 1, \quad b_n^{(q)}(\beta_n) := \left\{ \frac{\log\log(n^{\beta_n}) - \log(Y_q^{(n)})}{|\log(\tau - 2)|} \right\}, \tag{1.15}$$

where $\lfloor x \rfloor$ denotes the largest integer that is at most x and $\{x\} = x - \lfloor x \rfloor$ denotes the fractional part of x. Further, let $\lceil x \rceil$ denote the smallest integer that is larger than x.

1.1.2 Typical Distances

Theorem 1.8 (Distances in truncated power-law configuration models) *Consider the configuration model with empirical degree distribution satisfying Assumptions 1.1-1.3 with some $\beta_n \in (0, 1/(\tau - 1)]$ such that $\beta_n(\log n)^\gamma \to \infty$ for some $\gamma \in (0, 1)$. When $\beta_n \to 1/(\tau - 1)$ then we require that Assumption 1.7 holds additionally.*

$$d_G(v_r, v_b) = T_r(\beta_n) + T_b(\beta_n) + \left\lceil \frac{1/\beta_n - (\tau - 2)^{b_n^{(r)}(\beta_n)} - (\tau - 2)^{b_n^{(b)}(\beta_n)}}{3 - \tau} \right\rceil + 1. \tag{1.16}$$

The sequence $Y_q^{(n)}$ in $T_q(\beta_n), b_n^{(q)}(\beta_n)$ converges in distribution as $n \to \infty$. It is straightforward to show that for a sequence of random variables X_n converging in distribution, the transforms $\lfloor \log\log(n^\beta) + X_n \rfloor$ and $\{\log\log n + X_n\}$ do not converge, since their distribution shifts along as $\log\log(n^\beta)$ moves from one integer to the next. We can rephrase the statement of Theorem 1.8 in terms of convergence in distribution by filtering out the parts that show loglog-periodicity.

Corollary 1.9 *The following distributional convergence holds:*

$$d_G(v_r, v_b) - \frac{2\log\log(n^{\beta_n})}{|\log(\tau - 2)|} - \left\lceil \frac{1/\beta_n - (\tau - 2)^{b_n^{(r)}(\beta_n)} - (\tau - 2)^{b_n^{(b)}(\beta_n)}}{3 - \tau} \right\rceil$$

$$+ b_n^{(r)}(\beta_n) + b_n^{(b)}(\beta_n) \xrightarrow{d} -1 + \frac{-\log(Y_r Y_b)}{|\log(\tau - 2)|}. \tag{1.17}$$

Alternatively, we obtain weak convergence along (double-exponentially growing) subsequences $(n_k)_{k \in \mathbb{N}}$ satisfying $\log\log(n_k^{\beta_n}) = k + c + o(1)$ for every $c \in [0, 1)$.

---

[2] The notation comes from the fact that $Y_r^{(n)}, Y_b^{(n)} \in \mathcal{F}_{\varrho'_n}$.

Remark 1.10 The criterion that $\beta_n (\log n)^\gamma \to \infty$ for some $\gamma \in (0, 1)$ is just slightly stronger than requiring that the empirical second moment of the degrees in the graph tends to infinity. Indeed, we give in (2.44) in Claim 2.6 the upper bound $n^{(3-\tau)\beta_n}$ on the empirical second moment of the degrees. A similar lower bound can be proved as well. This expression tends to infinity whenever $\beta_n \log n \to \infty$.

Remark 1.11 The message of Theorem 1.8 is that typical distances are centered around approximately $2 \log \log(n^{\beta_n})/|\log(\tau - 2)| + 1/(\beta_n(3 - \tau))$, when the truncation of the degrees happens at a value n^{β_n}, with tight fluctuations around this value. As pointed out before, the threshold for the dominance of the two terms is at $\beta_n = \Theta(1/\log \log n)$. Indeed, as soon as $\beta_n = o(1/\log \log n)$, the term containing $1/\beta_n$ in (1.17) becomes dominant, and the second order term[3] is of order $\log \log n$. On the other hand, when $\beta_n \log \log n \to \infty$, the dominant term is $\log \log n$.

Remark 1.12 (Dropping the condition on the minimal degree) With slightly more work it is possible to drop Assumption 1.2 from the assumptions in Theorems 1.8. Under Assumption 1.1 but without Assumption 1.2, the graph has a unique *giant component* of linear size, $\zeta n(1 - o_\mathbb{P}(1))$ for some $\zeta > 0$, see Janson and Luczak [34]. In this case, the statement of Theorem 1.8 remain valid *conditioned* on the event that both v_r, v_b are in the giant component of the graph. This conditioning can be done similarly as described in [31]. To keep our paper short, we omit to provide the proof here, since this is not the main focus of this paper.

1.1.3 Structure and Number of Shortest Paths

The next two theorems are by-products of the proof of Theorem 1.8. They reveal the structure of shortest paths and thus shed light on the topology of the graph in more detail. Let us denote a path connecting vertices u, v by $\mathcal{P}_{u,v}$, and let us denote any path that realizes the graph distance $d_G(u, v)$ by $\mathcal{P}_{u,v}^\star$. We write $w \in \mathcal{P}_{u,v}$ if w is a vertex $\neq u, v$ that is on the path $\mathcal{P}_{u,v}$. We write $d_G(u, v | \Lambda)$ for the length of the shortest path between two vertices u, v restricted to contain vertices in a set Λ. Finally, let us write

$$\Lambda_{\leq z} := \left\{ v \in [n] : d_v \leq n^z \right\}, \tag{1.18}$$

and for a triplet (z, x_1, x_2) of numbers let us define the 'upper' and 'lower' fractional-part of the following expression:

$$f^u(z, x_1, x_2) := \left\lceil \frac{1/z - x_1 - x_2}{3 - \tau} \right\rceil - \frac{1/z - x_1 - x_2}{3 - \tau}, \tag{1.19}$$

$$f^\ell(z, x_1, x_2) := \frac{1/z - x_1 - x_2}{3 - \tau} - \left\lfloor \frac{1/z - x_1 - x_2}{3 - \tau} \right\rfloor. \tag{1.20}$$

Note that either $f^u = f^\ell = 0$ or $f^u = 1 - f^\ell \in (0, 1)$.

Theorem 1.13 (Structure and number of shortest paths between hubs) *Under the same conditions as Theorem 1.8, let $\widetilde{\beta}_n \leq \beta_n$ be such that $\widetilde{\beta}_n (\log n)^\gamma \to \infty$. Let $x_1, x_2 > \tau - 2$ and v_1, v_2 be two vertices with degrees $d_{v_j} \sim n^{x_j \widetilde{\beta}_n}$ for $j = 1, 2$. Then, the distance between v_1, v_2 restricted to paths $\mathcal{P}_{v_1,v_2}$ that contain only vertices with degree at most $n^{\widetilde{\beta}_n}$ is whp*

$$d_G(v_1, v_2 \mid \Lambda_{\leq \widetilde{\beta}_n}) = \left\lceil \frac{1/(\widetilde{\beta}_n) - x_1 - x_2}{3 - \tau} \right\rceil + 1, \tag{1.21}$$

---

[3] Note that in (1.17), the terms containing $\log \log n$ and $-\log \beta_n$ both compete to be the second order term, but under the criterion on β_n, $-\log \beta_n = O(\log \log n)$.

while the number of shortest paths[4] *between* v_1, v_2 *within* $\Lambda_{\leq \widetilde{\beta}_n}$ *satisfies whp*

$$\# \left\{ \mathcal{P}^{\star}_{v_1, v_2} \mid \Lambda_{\leq \widetilde{\beta}_n} \right\} \sim n^{\widetilde{\beta}_n f^u(\widetilde{\beta}_n, x_1, x_2)}, \tag{1.22}$$

where $f^u(\widetilde{\beta}_n, x_1, x_2)$ *is defined in* (1.19). *Further, whp all vertices on* any *shortest path in* $\Lambda_{\leq \widetilde{\beta}_n}$ *connecting* v_1, v_2 *have degree at least* $n^{\widetilde{\beta}_n f^{\ell}(\widetilde{\beta}_n, x_1, x_2)}$. *I.e., for all* $\varepsilon > 0$,

$$\mathbb{P}\left(\exists \mathcal{P}^{\star}_{v_1, v_2} \in \Lambda_{\leq \widetilde{\beta}_n}, w \in \mathcal{P}^{\star}_{v_1, v_2} : d_w \leq n^{\widetilde{\beta}_n f^{\ell}(\widetilde{\beta}_n, x_1, x_2)(1-\varepsilon)} \right) \to 0. \tag{1.23}$$

Interpreting Theorem 1.13, first note that setting $\widetilde{\beta}_n < \beta_n$ in (1.21) reveals the distance between the two vertices when the path must avoid vertices with degree at least $n^{\widetilde{\beta}_n}$. While (1.21) for $\widetilde{\beta}_n \equiv \beta_n$ shows that the generating function approximation known from physics - the formula described in (1.2)—is valid when instead of typical distances, we consider distances between very high-degree vertices. The second statement, (1.22) shows that the number of shortest paths between hubs concentrate on a logarithmic scale, since the statement

$$\frac{\log \left(\# \{ \mathcal{P}_{v_1, v_2} \in \Lambda_{\leq \widetilde{\beta}_n} \} \right)}{f^u(\widetilde{\beta}_n, x_1, x_2) \widetilde{\beta}_n \log n} \xrightarrow{\mathbb{P}} 1. \tag{1.24}$$

is a direct consequence of (1.22). This implies that there are *many* shortest paths, $n^{\widetilde{\beta}_n f^u(\widetilde{\beta}_n, x_1, x_2)}$ many. Note however that as soon as any of $\widetilde{\beta}_n, x_1, x_2$ depends on n, the upper fractional part f^u starts to oscillate in the interval $[0, 1)$. Similarly, the statement in (1.23) shows that *all these shortest paths* use relatively high degree vertices - just a factor f^{ℓ} multiplies the exponent of the maximally allowed degree $n^{\widetilde{\beta}_n}$. Keep in mind that f^u and f^{ℓ} are does not necessarily are bounded away from 0.

Note again that as soon as any of β_n, x_1, x_2 depends on n, $f^{\ell}(\beta_n, x_1, x_2)$ oscillates together with this fractional part.

Comparing (1.22) and (1.23), one notes that $f^u + f^{\ell} = 1$ unless both of them are 0. One can avoid integer values by slightly changing β_n or x_1, x_2 by pushing some n-dependent terms into the $d_{v_j} \sim n^{x_j \beta_n}$ relation. Further, it is not hard to extend the proof of Theorem 1.13 to show that there is at least one shortest path that uses a vertex with degree $\sim n^{\widetilde{\beta}_n f^{\ell}(\widetilde{\beta}_n, x_1, x_2)(1+\varepsilon)}$ for arbitrary small $\varepsilon > 0$. Thus, we arrive to the following observation.

Observation 1.14 *Let* v_1, v_2 *be two hubs with* $d_{v_j} \sim n^{x_j \beta_n}$ *for* $x_j > \tau - 2$, $j \in \{1, 2\}$. *Then the number of shortest paths between* v_1, v_2 *times the lowest degree that these paths use gives approximately*[5] *the maximal degree in the graph, i.e.,* $\sim n^{\beta_n}$.

We provide a sketch proof of this observation in Section 5. Our next theorem analyses the number and structure of shortest paths between two uniformly chosen vertices v_r, v_b:

Theorem 1.15 (Structure and number of shortest paths) *Under the same conditions as in Theorem 1.8, there is a shortest path between* v_r, v_b *that has the following structure, whp:*

(1) (Degree-increasing phase) *For both* $q \in \{r, b\}$, *starting from* v_q, *a path segment of length* $T_q(\beta_n) = \log \log(n^{\beta_n}) / |\log(\tau - 2)| + \Theta_{\mathbb{P}}(1)$ *as in* (1.15) *ends with a vertex* $v_q^{\star}$ *with degree*

---

[4] Here, overlaps between paths are allowed. We consider two paths different if they have at least one different edge.

[5] up to error terms of order at most $\exp\{\pm(\log n^{\beta_n})^{\theta}\}$, for some $\theta < 1$.

$$d_{v_q^\star} \sim n^{\beta_n(\tau-2)^{b_n^{(q)}(\beta_n)}}, \tag{1.25}$$

where $b_n^{(q)}(\beta_n)$ is from (1.15). The vertex $v_q^\star$ can be chosen to be the maximal degree vertex among all vertices that are reachable from v_q on a path of length $T_q(\beta_n)$. For any $k < i_{\star(q)}(\beta_n)$, the degree of the $(T_q(\beta_n) - k)$th vertex on the path between $v_q, v_q^\star$ is $\sim n^{\beta_n(\tau-2)^{b_n^{(q)}(\beta_n)+k}}$, where $i_{\star(q)}(\beta_n)$ are tight random variables given below in (2.34) and (2.36).

(2) (Connection among high-degree vertices phase) *A path of length*

$$\left\lceil \frac{1/\beta_n - (\tau - 2)^{b_n^{(r)}(\beta_n)} - (\tau - 2)^{b_n^{(b)}(\beta_n)}}{3 - \tau} \right\rceil + 1 \tag{1.26}$$

connects $v_r^\star, v_b^\star$ *using only vertices with degree at least* $n^{\beta_n f_n^\ell}$, *where, in agreement with Theorem 1.13,* $f_n^\ell := f^\ell(\beta_n, (\tau - 2)^{b_n^{(r)}(\beta_n)}, (\tau - 2)^{b_n^{(b)}(\beta_n)})$. *Further, Phase (2) is valid for all shortest paths, whp. That is, whp, for any shortest path* $\mathcal{P}_{v_r, v_b}^\star$, *the segment between the* $T_r(\beta_n)$th *and the* $|\mathcal{P}_{v_r, v_b}^\star| - T_b(\beta_n)$th *vertex has length as in (1.26), and it only contains vertices with degree at least* $n^{\beta_n f_n^\ell}$.

Finally, as a consequence of Theorem 1.13, the number of shortest paths between v_r, v_b *satisfies* $\mathbb{P}_Y - whp$

$$\#\{\mathcal{P}_{v_r, v_b}^\star\} \sim n^{\beta_n f_n^u},$$

where $f_n^u := f^u(\beta_n, (\tau - 2)^{b_n^{(r)}(\beta_n)}, (\tau - 2)^{b_n^{(b)}(\beta_n)}) \in [0, 1)$ *fluctuates with* n.

Theorem 1.15 sheds light on the true structure of the shortest path between two uniformly chosen vertices: both ends of the path start with a segment where after a short initial randomness, degrees are essentially increasing in a deterministic fashion: each degree is asymptotically a power $1/(\tau - 2)$ of the previous degree on the path. This phase ends with a vertex $v_q^\star$ that has degree in the interval $[n^{(\tau-2)\beta_n}, n^{\beta_n}]$, for $q \in \{r, b\}$. The precise (random) prefactors of β_n for $q \in \{r, b\}$ in the exponent of n of the degree $d_{v_q^\star}$ determines the precise length of the next phase, that establishes a connecting path between $v_r^\star, v_b^\star$. For this path segment, Theorem 1.13 can be applied, which means that the length of this path is at most a constant $\lceil 2/(3-\tau) \rceil$ away from $1/(\beta_n(3 - \tau))$, and all vertices on this path have relatively high degree.

We can turn Observation 1.14 to hold for two uniformly chosen vertices as well: the non-integer condition holds whp under Assumption 1.7 with $(x_1, x_2) = ((\tau - 2)^{b_n^{(q)}(\beta_n)})_{q \in \{r, b\}}$, so we arrive to the following observation:

Observation 1.16 *Let* v_r, v_b *be two uniformly chosen vertices. Then whp under Assumption 1.7, the number of shortest paths between* v_r, v_b *times the lowest degree that these paths use in Phase (2) in Theorem 1.15 is approximately the maximal degree in the graph, i.e., it is* $\sim n^{\beta_n}$.

1.1.4 Attack Vulnerability

Let us mention that when we remove some set of vertices Λ from a configuration model and the edges attached to them (called an attack), the remaining graph is still a configuration model on $[n] \setminus \Lambda$, with a new empirical degree distribution that might become *random*, depending on the type of the attack. Observe that the shortest path between two vertices in

the remaining graph is the same as the shortest path in the original graph restricted to stay among vertices in $[n] \setminus \Lambda$. When the attack is so that it removes all vertices above a certain degree, we call it a *targeted attack* or *deliberate attack*. This is the meaning of setting $\widetilde{\beta}_n < \beta_n$ in Theorem 1.13 above.

An immediate corollary of the proof of Theorems 1.8 and 1.13 is that our results remain valid in the configuration model with targeted attack as well, that is, when instead of truncating the degrees, we *remove* all vertices with degree at least $n^{\widetilde{\beta}_n}$ from a configuration model. Equivalently, we can consider the length of the shortest path restricted to stay among vertices with degree at most $n^{\widetilde{\beta}_n}$.

Corollary 1.17 *Let us consider the configuration model under the same assumptions as the ones in Theorem 1.8. For a sequence $\widetilde{\beta}_n$ with $\widetilde{\beta}_n(\log n)^\gamma \to \infty$ for some $\gamma < 1$, let us remove all vertices with degree at least $n^{\widetilde{\beta}_n}$ and the edges attached to them from the graph on n vertices. Then, the typical distance between two vertices v_r, v_b chosen uniformly at random from the remaining set of vertices satisfies whp*

$$d_G(v_r, v_b | \Lambda_{\leq \widetilde{\beta}_n}) - \frac{2\log\log(n^{\widetilde{\beta}_n})}{|\log(\tau-2)|} - \left\lceil \frac{1/(\widetilde{\beta}_n) - (\tau-2)^{b_n^{(r)}(\widetilde{\beta}_n)} - (\tau-2)^{b_n^{(b)}(\widetilde{\beta}_n)}}{3-\tau} \right\rceil$$

$$+ b_n^{(r)}(\widetilde{\beta}_n) + b_n^{(b)}(\widetilde{\beta}_n) \xrightarrow{d} -1 + \frac{-\log(Y_r Y_b)}{|\log(\tau-2)|}. \tag{1.27}$$

Further, Theorem 1.15 also remains valid in this setting, with β_n replaced by $\widetilde{\beta}_n$ everywhere.

This corollary sheds light on the effect of a targeted attack - commonly known as the *attack vulnerability* of the network. In fact, Corollary 1.17 describes the way typical distances grow when we (gradually) remove the 'core' of the graph, meaning all vertices with degree at least $n^{\widetilde{\beta}_n}$. For example, starting with a configuration model with i.i.d. degrees, (corresponding to $\beta_n \equiv 1/(\tau-1)$), one has to go as far as to choose $\widetilde{\beta}_n = o(1/\log\log n)$ to change the order of magnitude of the length of shortest paths.

An alternative proof of this corollary could be the following: It can be shown that the number of vertices in $\Lambda_{\leq \widetilde{\beta}_n} = o(n)$, so, only $o(n)$ many vertices are removed. Then, one can show that – even though these are the highest degree vertices – the total number of half-edges attached to the removed vertices is still $o(n)$. When considering the degree of a remaining vertex in the remaining graph, the half-edges that were matched to removed vertices should also be removed. This results in a *thinning* of the degrees. This thinning is *not independent* for different half-edges and vertices. However, by a stochastic domination argument one can still show that the resulting new degree distribution still satisfies the conditions of Theorem 1.8, with now β_n replaced by $\widetilde{\beta}_n$. This makes clear why all expressions in (1.27) depend only on $\widetilde{\beta}_n$ and not on the original β_n.

1.2 Examples

Note that Assumption 1.1 is satisfied in the following cases that we keep in mind to study:

Example 1.18 The first example arises when the degrees are *independent and identically distributed* from a background power-law distribution F that satisfies (1.4) and (1.5) (for all $x \in \mathbb{R}$). In this case, it is not hard to see that the order of magnitude of the maximal degree in the graph is $n^{(1+o_{\mathbb{P}}(1))/(\tau-1)}$ whp. Further, using the concentration of binomial random variables (see [31] for the computations) shows that whp the empirical degree distribution

satisfies Assumption 1.1, with $\beta_n = 1/(\tau - 1)$, with a possibly larger constant C in the bound on $L_n(x)$ than the one in the background distribution F.

Pure power-law degrees were found for example in the internet backbone network [21], in metabolic reaction networks [35], in telephone call graphs [46], and most famously, in the world-wide-web [5,10,37].

Remark 1.19 (Fluctuations of typical distances in the i.i.d. degree case) In the special case of i.i.d. degrees as in Example 1.18, the value $\beta_n = 1/(\tau - 1)$. Under Assumption 1.7, the upper integer part in (1.16) simplifies to either 0 or 1, and typical distances in this case become

$$d_G(v_r, v_b) = T_r(\beta_n) + T_b(\beta_n) + 2 - \mathbb{1}\left\{\tau - 1 < (\tau - 2)^{b_n^{(r)}(\beta_n)} + (\tau - 2)^{b_n^{(b)}(\beta_n)}\right\}. \tag{1.28}$$

We emphasize that Theorem 1.8 implies that the typical distances in the graph are concentrated around $2\log\log n/|\log(\tau - 2)|$ with bounded fluctuations, a result that already appeared in [31] for the i.i.d. degree case. The statement of Corollary 1.9 'filters out' the bounded oscillations arising from fractional part issues that oscillate with n. We emphasize here that the statement of Theorem 1.8 applied to i.i.d. degrees and [31, Theorem 1.2] are essentially the same. However, they provide a different description of typical distances. The current proof here is much shorter than the one in [31] as well as it allows to treat the truncated degree case at the same time. We show in Section 6 that the two theorems are indeed the same.

Example 1.20 (Exponential truncation) The degrees are generated i.i.d. from an n-dependent *truncated power law* distribution $F^{(n)}$ that can be written in the form

$$1 - F^{(n)}(x) = \frac{L^{(n)}(x)}{x^{\tau-1}} \exp\{-cx/n^{\beta_n}\}, \tag{1.29}$$

with $L^{(n)}$ satisfying (1.5). In this case, the empirical distribution $F_n(x)$ satisfies Assumption 1.1 for all sufficiently large n, since for any $x \leq n^{\beta_n(1-\varepsilon)}$, the exponential term is at least $1/2$, say, while for any $x \geq n^{\beta_n(1+\varepsilon)}$, $F^{(n)}(x) = O(1/n^2)$, thus we shall not actually see vertices with such high degree in the graph for large enough n. A special case arises when $(d_i)_{i\in[n]}$ are i.i.d. with $d_i = \min\{X_i, G_i\}$, where G_i are i.i.d. geometric random variables with mean $\exp\{c/n^{\beta_n}\}$, and X_i i.i.d. as in Example 1.18.

Power-law degrees with exponential truncation were proposed e.g. in [45], and are observed for instance in the movie actor network [4], air transportation networks [27] and co-authorship networks [48,58], brain functional networks [1], ecological networks [44] such as coevolutionary networks of plant-animal interactions [36].

Example 1.21 (Hard truncation) The degrees are generated i.i.d. from an $n-dependent$ *truncated power-law* that can be written in the form

$$F^{(n)}(x) = \frac{L^{(n)}(x)}{x^{\tau-1}} \mathbb{1}_{x<n^{\beta_n}}, \tag{1.30}$$

with $L^{(n)}$ satisfying (1.5). A special case again arises when $(d_i)_{i\in[n]}$ are i.i.d. with $d_i = \min\{X_i, n^{\beta_n}\}$, where X_i are i.i.d. as in Example 1.18.

Probably the most important example for hard truncation is a *targeted attack*, since in this case every vertex above a certain degree in the network is removed. Scale-free graphs are often called attack vulnerable, see e.g. [3,14,32]. A theoretical example where the authors use a network model with hard truncation can be found in [26] or [40].

Perhaps surprisingly, the online social network of Facebook does not seem to follow a truncated power-law [60], even though the total number of friends of a person was limited to 5000 at the time of the measurement. Nevertheless, we would like to emphasize that our theorem allows for many possible truncations functions, among which the hard truncation is possibly the most strict.

1.3 Discussion and Open Questions

Heuristic explanation of the formula in Theorem 1.8.

In Theorem 1.8, we have determined that the distances are centered around

$$\frac{2 \log \log(n^{\beta_n})}{|\log(\tau - 2)|} + \frac{1}{\beta_n(3 - \tau)}$$

with tight fluctuations. Here we give a heuristic explanation of this formula. For graphs with locally tree-like structure, the usual BP approximation says that the number of vertices that are reachable on a path of length k from v_r, is approximately $Z_k^{(n)}$, the size of generation k of a BP with offspring D_n in the first and $D_n^\star$ in the consecutive generations. Generating function methods then yield that $\mathbb{E}[Z_k^{(n)}] = \mathbb{E}[D_n]\nu_n^{k-1}$, with $\nu_n = \mathbb{E}[D_n^\star] \sim n^{\beta_n(3-\tau)}$ as in (4.2) below, and then the approximation $Z_k^{(n)} \approx \mathbb{E}[Z_k^{(n)}]$ is often used. Unfortunately, for heavily skewed distributions like that of $D_n^\star$, it does not hold that $Z_k^{(n)} \approx \mathbb{E}[Z_k^{(n)}]$. This is so because $\mathbb{E}[D_n^\star]$ is characterised by the highest degree vertices, of degree n^{β_n}, while, on the other hand, low degree vertices are typically not connected to these hubs in the graph, and thus $Z_k^{(n)} \ll \mathbb{E}[Z_k^{(n)}]$.

It is true however that

$$Z_k^{(n)} \approx C^{1/(\tau-2)^k}, \tag{1.31}$$

for some random constant C [16]. Thus, as long as the BP approximation is valid, we see a 'degree-increasing phase' within the exploration clusters of the vertices v_r, v_b. From the approximation (1.31) it already follows that it takes $\log \log(n^{\beta_n})/|(\log(\tau-2))|+$tight number of steps to reach a hub $v_q^\star$ in the graph, for $q \in \{r, b\}$. Extreme value theory tells us that any hub will have some neighbors that are also hubs, and thus the approximation that $D_n^\star \approx \mathbb{E}[D_n^\star]$ and consequently $Z_k^{(n)} \approx \mathbb{E}[Z_k^{(n)}]$ suddenly becomes valid when considering the number of vertices of distance k away from $v_q^\star$. This means that it takes an *additional* $\log n / \log \nu_n = 1/\beta_n(3-\tau)+$tight number of steps to connect the two hubs $v_r^\star, v_b^\star$ to each other. This explains the formula in Theorem 1.8.

Comment About 'Structural Cutoff'

Often in physics literature, $\beta_n = 1/2$ is called 'structural cut-off' [54]. When $\beta_n > 1/2$, vertices with degree at least $n^{1/2+\varepsilon}$ form a complete subgraph of the graph, while for $\beta_n < 1/2$ this complete subgraph is not present. Further, when $\beta_n > 1/2$, a growing number of multiple edges appears, while for $\beta_n < 1/2$, the number of multiple edges in the graph stays bounded. Our theorems show that there is no significantly different behavior of typical distances when the truncation happens below versus above the structural cutoff $n^{1/2}$.

Open Questions

We believe the criterion $\beta_n (\log n)^\gamma \to \infty$ for some $\gamma \in (0, 1)$ can be relaxed to be $\beta_n \log n \to \infty$, at least when one imposes more strict bounds on the slowly varying function $L_n(x)$ in (1.4). The criterion $\beta_n \log n \to \infty$ is the weakest form that is necessary for the empirical second moment to tend to infinity. Provided one can generalize our results to hold whenever $\beta_n \log n \to \infty$, we obtain a perfect interpolation between doubly logarithmic and logarithmic distances. When $\beta_n \log n = \theta(1)$, the empirical second moment remains bounded and thus a finite mean BP approximation becomes available.

Notation

We write $[n]$ for the set of integers $\{1, 2, \ldots, n\}$. As usual, we write i.i.d. for independent and identically distributed, lhs and rhs for left-hand side and right-hand side. We use $\overset{d}{\longrightarrow}$, $\overset{\mathbb{P}}{\longrightarrow}$, $\overset{a.s.}{\longrightarrow}$ for convergence in distribution, in probability and almost surely, respectively. We use the Landau symbols $o(\cdot), O(\cdot), \Theta(\cdot)$ in the usual way. For sequences of random or deterministic variables X_n, Y_n we further write $X_n = o_{\mathbb{P}}(Y_n)$ and $X_n = O_{\mathbb{P}}(Y_n)$ if the sequence $X_n/Y_n \overset{\mathbb{P}}{\longrightarrow} 0$ and is tight, respectively.

Constants are typically denoted by c in lower and C in upper bounds (possible with indices to indicate which constant is coming from which bound), and their precise values might change from line to line. We introduce $(\mathcal{C}^{(r)}, \mathcal{C}^{(b)}) := (\mathcal{C}^{(r)}, \mathcal{C}^{(b)})$. At some time t along the matching or the exploration, for a set of vertices $\mathcal{A}_t$ we denote the set of unpaired half-edges at that moment attached to vertices in $\mathcal{A}_t$ as $\mathcal{H}(\mathcal{A}_t)$ and its size by $H(\mathcal{A}_t)$. Thus $H(\mathcal{A}_t) = \sum_{v \in \mathcal{A}_t} \sum_{s \text{ half-edge attached to } v} \mathbb{1}_{s \text{ not paired yet}}$. When the time is set to be 0, $H(\mathcal{A}) = \sum_{v \in \mathcal{A}} d_v$. We write d_v for the degree of vertex v.

1.4 Overview of the Proof

To determine the distance between v_r, v_b, we start growing two clusters that we call red and blue, respectively, in a breadth-first-search manner, and see how these two clusters reach the highest degree vertices. As long as the two clusters are disjoint, the growth is not necessarily simultaneous, i.e., we might stop the growth of one color earlier. To describe the growing clusters, we extensively use the fact that in the configuration model half-edges can be paired in an arbitrarily chosen order. This allows for a joint construction of the graph together with the growing of the two colored clusters. In Proposition 2.1 (Section 2) below, we show that the highest degree vertex $v_q^\star, q \in \{r, b\}$ that is reached by this path is of degree

$$d_{v_q^\star} \sim n^{\beta_n(\tau-2)^{b_n^{(q)}(\beta_n)}}, \tag{1.32}$$

for $q \in \{r, b\}$, respectively. The total length of this path from v_q is $T_q(\beta_n)$ as in (1.15) for $q \in \{r, b\}$. The proof of this proposition has the following ingredients: We couple the initial stages of the growth to two independent branching processes (BPs) (Section 2.1). The coupling fails when one of the colors (wlog we assume it is red) reaches size $n^{\varrho'_n}$ for some $\varrho'_n > 0$ sufficiently small. From the half-edges attached to the BP cluster of $v_q, q \in \{r, b\}$, we build a path through higher and higher degree vertices to a vertex with degree at least $\sim n^{(\tau-2)\beta_n}$. In Lemma 2.5 we give an upper bound on the degree of the maximal-degree vertex reached at any time $t(n^{\varrho'_n}) + i$ of the exploration, implying that in $T_q(\beta_n)$ hops no vertex of degree higher than $\sim n^{\beta_n(\tau-2)^{b_n^{(q)}(\beta_n)}}$ is reached from v_q. Thus, this lemma serves

also as a building block for the proof of Proposition 2.1, but beyond that, it also enables us to show two other important things, that form the content of Proposition 3.1 (Section 3):

(1) An early meeting is highly unlikely, i.e., the clusters $\mathcal{C}^{(q)}_{T_q(\beta_n)}$ are $\mathbb{P}_Y$-whp disjoint.
(2) The quantity in (1.32) bounds also the *total number of half-edges* attached to the explored cluster $\mathcal{C}^{(q)}_{T_q(\beta_n)}$.

Finally, in Sect. 4 we finish the proof of Theorem 1.8. For the lower bound, we count the number of z-length paths between the two disjoint clusters $\mathcal{C}^{(q)}_{T_q(\beta_n)}$. Here, we use a first moment method (i.e., we show that the expected number of paths is $o(1)$) when z is less than the expression in (1.26). For the upper bound, we establish the existence of a path of length as in (1.26) between $v_r^\star, v_b^\star$. We do this using a second moment method. This completes the proof of Theorem 1.8. We prove Theorems 1.13 and 1.15 in Sect. 5. In Sect. 6 we compare Theorem 1.8 in the special case $\beta_n \equiv 1/(\tau - 1)$ to the result in [31].

2 Distance from the Hubs

In this section we analyse the distance of v_r, v_b from the highest-degree vertices. The construction is similar to that of [6, Sect. 3], however, since Assumption 1.1 on F_n is weaker than the one in [6], we need to modify the proof. Recall that we call the vertices with degree at least $n^{\beta_n(\tau-2)}$ hubs and recall $T_q(\beta_n), b_n^{(q)}(\beta_n)$ from (1.15) More precisely, let us define

$$\text{hubs} := \{v \in [n] : d_v \geq n^{\beta_n(\tau-2)}\}, \tag{2.1}$$

and for a set of vertices $\mathcal{A} \subseteq [n]$ and a vertex $v \in [n]$,

$$d_G(v, \mathcal{A}) := \min_{a \in \mathcal{A}} d_G(v, a). \tag{2.2}$$

The next proposition determines the distance of the uniformly chosen vertices v_r, v_b from the hubs:

Proposition 2.1 (Distance from the hubs) *Let us consider the configuration model on n vertices with empirical degree distribution that satisfies Assumption 1.1, and let v_r, v_b be two uniformly chosen vertices. Then, for $q \in \{r, b\}$, $\mathbb{P}_Y$-whp,*

$$d_G(v_q, \text{hubs}) = T_q(\beta_n) = \frac{\log \log n + \log\left(\beta_n/Y_n^{(q)}\right)}{|\log(\tau - 2)|} - 1 - b_n^{(q)}(\beta_n).$$

More precisely, $\mathbb{P}_Y$-whp there is a vertex $v_q^\star \in \text{hubs}$ at distance $T_q(\beta_n)$ away from v_q with degree

$$d_{v_q^\star} \sim n^{\beta_n(\tau-2)^{b_n^{(q)}(\beta_n)}}, \tag{2.3}$$

while all vertices at distance at most $T_q(\beta_n) - 1$ from v_q are not hubs.

The main goal of this section is to prove this proposition. To show the upper bound, the proof has two main steps: the initial stage of the breadth-first-search exploration (BFS) is coupled to branching process trees (Sect. 2.1), while the later stage uses a decomposition of the vertices with degrees that are polynomial in n into shells (Sect. 2.2). To show the lower bound, we provide an upper bound on the degrees reached by the BFS in any shell at the time of first reaching that particular shell, see Lemma 2.5 (Sect. 2.3). This method is novel compared to the one in [31].

2.1 Coupling the Initial Stages of BFS to Branching Processes

In this section we investigate the initial stage of the spreading cluster of v_r, v_b. In the construction of the configuration model, at any time that we construct the matching of the half-edges, we are allowed to choose one of the not-yet-paired half-edges arbitrarily, and pair it to a uniformly chosen other not-yet-paired half-edge. Hence, we can do the pairing in an order that corresponds to the breadth-first-search (BFS) exploration started from v_r, v_b. That is, first we pair all the outgoing half-edges from the sources v_r, v_b (distance 1), then we pair the outgoing half-edges from the neighbors of the source vertices (distance 2), and so on, in a breadth-first-search manner. Whenever we finish pairing all the half-edges attached to vertices at a given graph distance from the source vertices, we increase the distance by 1. This process of joint construction of the BFS exploration and graph building is often called the *exploration process* in the literature.

Recall that $\mathcal{C}_t^{(r)}$, $\mathcal{C}_t^{(b)}$ denotes the subgraph that is at distance at most t from v_r, v_b. [7, Proposition 4.7] (see also [6, Lemma 2.2]) shows that that the number of vertices and their forward degrees[6] in the exploration process can be coupled to i.i.d. degrees having distribution function $F_n^\star$ from (1.6) as long as the total number of vertices of the colored clusters is not too large. There, a different assumption is posed on the maximal degree in the graph, so we shortly adjust the proof of [7, Proposition 4.7] to our setting below in Lemma 2.2. The distribution $F_n^\star$ arises from the fact that as long as the set of explored vertices is relatively small, a forward degree j is generated when the uniformly chosen half-edge of a pairing belongs to a vertex with degree $j + 1$. The probability of choosing a half-edge that belongs to a vertex with degree $j + 1$ is approximately equal to $(j + 1)\mathbb{P}(D_n = j + 1)/\mathbb{E}[D_n]$, and thus $F_n^\star$ and thus $F^\star$ are the natural candidates for the forward degrees in the exploration process.

Lemma 2.2 (Coupling error of the forward degrees) *Consider the configuration model with degree sequence that satisfies Assumption 1.1. Then, in the exploration process started from two uniformly chosen vertices v_r, v_b, 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 $D_{n,k}^\star$ from distribution $D_n^\star$ with the following error bound*

$$\mathbb{P}(\exists k \le s_n, D_{n,k}^\star \ne X_k^{(n)}) \le C s_n^2 n^{\beta_n(1+\varepsilon)-1} + C n^{-(\tau-2-2\varepsilon)/(\tau+\varepsilon)} s_n^{2(\tau-1+\varepsilon/2)/(\tau+\varepsilon)} \quad (2.4)$$

If further Assumption 1.3 holds, then there is a coupling of $(X_k^{(n)}, D_{n,k}^\star, D_k^\star)_{k \le s_n}$ with

$$\mathbb{P}(\exists k \le s_n, D_{n,k}^\star \ne X_k^{(n)} \text{ or } D_{n,k}^\star \ne D_k^\star) \le s_n n^{-\beta_n \kappa} + C s_n^2 n^{\beta_n(1+\varepsilon)-1}$$
$$+ C n^{-(\tau-2-2\varepsilon)/(\tau+\varepsilon)} s_n^{2(\tau-1+\varepsilon/2)/(\tau+\varepsilon)}. \quad (2.5)$$

By choosing s_n in Lemma 2.2 so that the rhs of the bound in (2.5) still tends to zero, we obtain the following corollary:

Corollary 2.3 (Whp coupling of the exploration to two BPs) *In the configuration model satisfying Assumptions 1.1 and 1.3, let t be such that*

$$|\mathcal{C}_t^{(r)} \cup \mathcal{C}_t^{(b)}| \le \min\{n^{\beta_n(\kappa-\delta)}, n^{(1-\beta_n(1+\varepsilon)-\delta)/2}, n^{-(\tau-2-\varepsilon-\delta)/(2(\tau-1+\varepsilon/2))}\} \quad (2.6)$$

---

[6] Forward degree means the number of newly available half-edges when a new vertex is discovered, that will be paired to new vertices later on.

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

Proof of Lemma 2.2 We would like to couple the forward degrees $(X_k^{(n)})_{k \leq s_n}$ in the exploration to an i.i.d. sample of size s_n from distribution $D_n^\star$ as in (1.7) as well as to $D^\star$ and estimate the coupling error. The idea of the coupling is to achieve size-biased sampling with and without replacement of the vertices at the same time: this is [8, Construction 4.2] that we informally recall here. Let us write $\mathcal{L}_n$ for the list of half-edges.

We use a sequence of uniform random variables $U_k \in [0, 1]$. Then, we sample a uniform half-edge h_k from $\mathcal{L}_n$, namely we set h_k to be the jth element of $\mathcal{L}_n$ if $U_k \in ((j-1)/\ell_n, j/\ell_n]$. We sample the i.i.d. $D_{n,k}^\star$ by setting $D_{n,k}^\star$ to be $d_{v(h_k)} - 1$, where $v(h)$ denotes the vertex that h is incident to.

At the same time, we keep a list of already sampled vertices $\mathcal{S}_k := \{v_r, v_b, v(h_1), \dots, v(h_k)\}$. As long as $v(h_k) \notin \mathcal{S}_{k-1}$, we can set $B_k^{(n)} := d_{v(h_k)} - 1$, this quantity describes the number of brother half-edges of a newly discovered vertex via a pairing. Note that in this sampling procedure there is no pairing of the half-edges yet. To ensure that the exploration cluster is a tree, at each step k we yet have to check if any of these $B_k^{(n)} = d_{v(h_k)} - 1$ half-edges create cycles when being paired, i.e., they shall be paired to vertices in $\mathcal{S}_{k-1}$. We write $X_k^{(n)}$ to be $B_k^{(n)}$ minus those half-edges that shall be paired to vertices in $\mathcal{S}_{k-1}$.

Thus, the coupling to a BP tree with offspring distribution $D_n^\star$ can fail in two ways: either $v(h_k) \in \mathcal{S}_{k-1}$ and the coupling between $B_k^{(n)}, D_{n,k}^\star$ fails (depletion-of-points effect), or $X_k^{(n)} < B_k^{(n)}$ and some of the $B_k^{(n)}$ half-edges create cycles (cycle-creation effect).

Introducing the σ-algebra $\mathcal{G}_k$ generated by v_r, v_b and the first k draws, [8, Lemma 4.3] bounds the coupling error between $B_k^{(n)}, D_{n,k}^\star$:

$$\mathbb{P}\big(B_k^{(n)} \neq D_{n,k}^\star \mid \mathcal{G}_{k-1}\big) \leq \frac{1}{\ell_n}\left(d_{v_r} + d_{v_b} + \sum_{s=1}^{k-1}(B_s^{(n)} + 1)\right), \tag{2.7}$$

while [8, Lemma 4.3] estimates the probability of creating a cycle at step k:

$$\mathbb{P}\left(X_k^{(n)} < B_k^{(n)} \mid \mathcal{G}_{k-1}\right)$$
$$\leq \frac{B_k^{(n)}}{\ell_n - 1 - \left(d_{v_r} + d_{v_b} + \sum_{s=1}^k\left(B_s^{(n)} + 1\right)\right)}\left(d_{v_r} + \sum_{s=1}^k(B_s^{(n)} + 1)\right). \tag{2.8}$$

Under Assumption 1.1 the maximal degree is at most $n^{\beta_n(1+\varepsilon)}$ in the graph, and $\ell_n = \mathbb{E}[D_n]n$ is of order n. Combining these observations, (2.7) is at most $Ckn^{\beta_n(1+\varepsilon)-1}$. Summing this bound over $k \leq s_n$, we obtain that

$$\mathbb{P}\big(\exists k \leq s_n, B_k^{(n)} \neq D_{n,k}^\star\big) \leq Cs_n^2 n^{\beta_n(1+\varepsilon)-1}. \tag{2.9}$$

Then, on the event $\{\forall k \leq s_n : B_k^{(n)} = D_{n,k}^\star\}$, $B_s^{(n)}$ in (2.8) can be replaced by the i.i.d. $D_{n,s}^\star$. Taking expectations of the rhs (2.8) does not work, since $\mathbb{E}[D_n^\star]$ is infinite. Thus we apply a truncation argument. Take $\varepsilon > 0$ so that $s_n^{(\tau-1+\varepsilon)/(\tau-2)} = o(n)$. Then the denominator on the rhs of (2.8) is at least cn for some $c > 0$. Further, for some truncation value K_n to be chosen later that satisfies that $s_n K_n = o(n)$,

$$\mathbb{P}\Big(\exists k \le s_n, X_k^{(n)} \ne B_k^{(n)} \mid \forall k \le s_n, B_k^{(n)} = D_{n,k}^\star\Big) \le \mathbb{P}\Big(\exists j \le s_n, D_{n,j}^\star > K_n\Big)$$

$$+ \frac{1}{cn} \sum_{k=1}^{s_n} \sum_{s=1}^{k} \mathbb{E}\Big[D_{n,k}^\star D_{n,s}^\star \mathbb{1}_{\{\forall j \le s_n : D_{n,j}^\star \le K_n\}}\Big]. \tag{2.10}$$

where we used that on the event that $\{\forall k \le s_n : B_k^{(n)} = D_{n,k}^\star, D_{n,k}^\star \le K_n, \}$ the denominator on the rhs of (2.8) is at least cn for some $c > 0$. Using (1.7) the first term on the rhs is at most $s_n L_n^\star(K_n) K_n^{-(\tau-2)} \le s_n K_n^{-(\tau-2)+\varepsilon}$ for arbitrarily small $\varepsilon > 0$ for sufficiently large n, while the second term is at most $Cn^{-1}K_n^2 s_n^2$. Making the order of the two error terms to be equal yields that the best choice of truncation value is at

$$K_n := (n/s_n)^{1/(\tau+\varepsilon)}. \tag{2.11}$$

Note that the initial criterium that $s_n K_n = o(n)$ is satisfied for all $s_n = o(n)$, while K_n is a polynomial of n with strictly positive exponent whenever $s_n = o(n^{(1-\varepsilon)/(\tau-1)})$. With K_n in (2.11) the sum of the error terms in (2.9) becomes

$$\mathbb{P}\Big(\exists k \le s_n, X_k^{(n)} \ne B_k^{(n)} \mid \forall k \le s_n, B_k^{(n)} = D_{n,k}^\star\Big) \le Cn^{-(\tau-2-2\varepsilon)/(\tau+\varepsilon)} s_n^{2(\tau-1+\varepsilon/2)/(\tau+\varepsilon)} \tag{2.12}$$

which tends to zero as long as $s_n = o\big(n^{(\tau-2-\varepsilon)/2(\tau-1+\varepsilon/2)}\big)$ for some arbitrarily small $\varepsilon > 0$.

Note that we have neglected coupling the forward degree of v_r, v_b to two i.i.d. copies of D_n. This coupling can be done in a very similar way, by choosing with and without replacement two uniform numbers from $[n]$, with a coupling error at most $1/n$, which is negligible compared to the rhs of (2.9) The probability that any of the d_{v_r} or d_{v_b} half-edges form cycles is at most of order $n^{\beta_n(1+\varepsilon)}/\ell_n$, which again can be merged into the rhs of (2.9) This finishes the proof of (2.4)

Next we extend the coupling between $\big(X_k^{(n)}, D_{n,k}^\star\big)$ to additionally couple $D_k^\star$ to them, using Assumption 1.3. On the event that $X_k^{(n)} = D_{n,k}^\star = \ell$, we use the optimal coupling that realizes the total variation distance between $D_n^\star$ and $D^\star$. Namely,

$$\mathbb{P}\Big(D_k^\star = \ell \mid X_k^{(n)} = D_{n,k}^\star = \ell\Big) := \frac{\min\{\mathbb{P}(D^\star = \ell), \mathbb{P}(D_n^\star = \ell)\}}{\mathbb{P}(D_n^\star = \ell)}$$

$$= \min\left\{1, \frac{\mathbb{P}(D^\star = \ell)}{\mathbb{P}(D_n^\star = \ell)}\right\}. \tag{2.13}$$

One can set the other possible values of $D_k^\star$ as described e.g. in [38, Chapter 1]. Nevertheless, the coupling error equals

$$\mathbb{P}\big(D_k^\star \ne D_{n,k}^\star\big) = \sum_{\ell \ge 1}\Big(1 - \min\{\mathbb{P}(D^\star = \ell), \mathbb{P}(D_n^\star = \ell)\}\Big) = d_{\mathrm{TV}}\big(F^\star, F_n^\star\big).$$

One can realize this coupling by using another independent uniform variable $\widetilde{U}_k$ to set the value of $D_k^\star$ once the value of $D_{n,k}^\star$ is determined. Thus, we obtain that

$$\mathbb{P}\big(\exists k \le s_n, D_{n,k}^\star \ne D_k^\star\big) \le d_{\mathrm{TV}}\big(F_n^\star, F^\star\big) \le s_n n^{-\kappa}. \tag{2.14}$$

Similarly, we can couple d_{v_r}, d_{v_b} to two i.i.d. copies of D with a coupling error at most $2d_{\mathrm{TV}}(F_n, F)$. This finishes the proof of (2.5) $\qquad\square$

A theorem by Davies [16] describes the growth rate of a branching process with a given offspring distribution G that satisfies (1.7) that we describe here informally. Let $\widetilde{Z}_k$ denote the k-th generation of a branching process with offspring distribution given by a distribution function G, that can be written in the form[7] as in (1.7) and (1.5) for some $\tau \in (2, 3)$ and some $x_0 > 0$ for all $x \geq x_0$. Then $(\tau - 2)^k \log (\widetilde{Z}_k \vee 1)$ converges almost surely to a random variable $\widetilde{Y}$. Further, the variable $\widetilde{Y}$ has exponential tails: if $J(x) := \mathbb{P}(\widetilde{Y} \leq x)$, then

$$\lim_{x \to \infty} \frac{-\log(1 - J(x))}{x} = 1. \tag{2.15}$$

We can apply Davies' theorem to each subtree of the two roots of the two i.i.d. BPs from Corollary 2.3, to obtain that the corresponding convergence in (1.12) in Definition 1.6. See [6, Lemma 2.4] for more details.

Recall from Corollary 2.3 that the coupling of the forward degrees in the BP and in the exploration fails when the total size of the BPs is too large, i.e., when (2.6)is not satisfied. Thus let us set[8]

$$\varrho'_n := (\tau - 2) \min \{\beta_n \kappa, (1 - \beta_n(1 + \varepsilon))/2, (\tau - 2 - \varepsilon)/(2(\tau - 1 + \varepsilon/2))\}. \tag{2.16}$$

Without loss of generality we assume that the cluster of v_r reaches size $n^{\varrho'_n}$ first (otherwise we switch the indices r, b). Thus, let us define

$$t\left(n^{\varrho'_n}\right) = \inf\left\{k : Z_k^{(r)} \geq n^{\varrho'_n}\right\}. \tag{2.17}$$

From the definition (1.11) an elementary rearrangement yields that (conditioned on $Y_r^{(n)}$),

$$t(n^{\varrho'_n}) = \frac{\log\left(\varrho'_n/Y_r^{(n)}\right) + \log\log n}{|\log(\tau - 2)|} + 1 - a_n^{(r)}, \tag{2.18}$$

where

$$a_n^{(r)} = \left\{\frac{\log\left(\varrho'_n/Y_r^{(n)}\right) + \log\log n}{|\log(\tau - 2)|}\right\}. \tag{2.19}$$

Note that $1 - a_n^{(r)}$ in (2.18) is there to make $t(n^{\varrho'_n})$ equal to the upper integer part of the fraction on the rhs of (2.18). Due to this effect, the last generation has a bit more vertices than $n^{\varrho'_n}$, namely

$$Z_{t(n^{\varrho'_n})}^{(r)} = n^{\varrho'_n(\tau-2)^{a_n^{(r)}-1}} =: m_r, \tag{2.20}$$

We obtain this expression by rearranging (1.11) and using the value $t(n^{\varrho'_n})$ from (2.18). The definition of ϱ'_n in (2.16) and $a_n^{(r)} \in [0, 1)$ implies that the exponent $\varrho'_n(\tau - 2)^{a_n^{(r)}-1}$ is still so small that the condition (2.6) in Corollary 2.3 is satisfied. Similarly, from (1.11) and (2.18), the blue cluster at this moment has size

$$Z_{t(n^{\varrho'})}^{(b)} = n^{\varrho'_n(\tau-2)^{a_n^{(r)}-1}Y_b^{(n)}/Y_r^{(n)}} =: m_b, \tag{2.21}$$

---

[7] Actually the theorem by Davies in [16] is somewhat more general, since it allows for a slightly larger class of slowly-varying functions than the criterion in (1.5) Nevertheless, the theorem applies for the case described in (1.7)and (1.5)

[8] The only reason to denote this quantity by ϱ'_n is to be consistent with the notation in [6].

where the assumption that red reaches size $n^{\varrho'_n}$ first is equivalent to the assumption that $Y_b^{(n)}/Y_r^{(n)} \leq 1$. This assumption together with (2.20) as well as the double-exponential growth apparent from (1.12) ensures that the total size of the two BPs is less than n^{ϱ_n} and thus Corolllary 2.3 still holds. Note that m_r, m_b are *random variables* that are measurable wrt $\mathcal{F}_{\varrho'_n}$.

2.2 Short Path to the Hubs Through Shells

To provide an upper bound on the distance of v_q and the hubs for $q \in \{r, b\}$, as well as to show (2.3) we build a path from $\mathcal{C}^{(q)}_{t(n^{\varrho'_n})}$ to the hubs. Let us set $C_2 := \max\{C_1, C_1^\star\}$, where $C_1, C_1^\star$ are the constants in the exponent in (1.5) for $L_n, L_n^\star$, respectively, and define the function

$$h(x) := \exp\left\{\frac{2C_2}{(\tau - 2)^\gamma}(\log x)^\gamma\right\}. \tag{2.22}$$

We shall repeatedly use that for any possible $L_n, L_n^\star$ satisfying (1.5) as $x \to \infty$,

$$\min\left\{L_n\left(x^{1/(\tau-2)}\right), L_n^\star\left(x^{1/(\tau-2)}\right)\right\} h(x) \to \infty, \tag{2.23}$$

$$\max\left\{L_n\left(x^{1/(\tau-2)}\right), L_n^\star\left(x^{1/(\tau-2)}\right)\right\} / h(x) \to 0. \tag{2.24}$$

Recall m_q from (2.20), (2.21) and that they are measurable wrt $\mathcal{F}_{\varrho'_n}$. Generally in the rest of the paper, random variables measurable wrt $\mathcal{F}_{\varrho'_n}$ are denoted by small letters, since they can be treated as constants under the measure $\mathbb{P}_Y$ in (1.14), and this is the meaure we mostly work with. In order to build the path, for both $q \in \{r, b\}$ we decompose the high-degree vertices in the graph into the following sets, that we call *shells*:

$$\Gamma_i^{(q)} := \left\{v : d_v > u_i^{(q)}\right\}, \tag{2.25}$$

where $u_i^{(q)}$ is defined recursively by

$$u_{i+1}^{(q)} = \left(\frac{u_i^{(q)}}{h(u_i^{(q)})}\right)^{1/(\tau-2)}, \qquad u_0^{(q)} := \left(\frac{m_q}{h(m_q)}\right)^{1/(\tau-2)}. \tag{2.26}$$

Setting $u_{-1}^{(q)} := m_q$, we obtain by iteration

$$u_i^{(q)} = m_q^{(\tau-2)^{-(i+1)}} / \prod_{k=1}^{i+1} h\left(u_{i-k}^{(q)}\right)^{(\tau-2)^{-k}}.$$

Clearly, $u_i^{(q)} \leq m_q^{(\tau-2)^{-(i+1)}}$. Using this upper bound to estimate the arguments of the function h in the denominator, as well as (2.22), the following lower bound holds with $K_\gamma = 1 - (\tau - 2)^{-(1-\gamma)}$:

$$u_i^{(q)} \geq m_q^{(\tau-2)^{-(i+1)}} \exp\left\{-\frac{2C_2}{K_\gamma(\tau-2)^\gamma}(\tau-2)^{-(i+1)}(\log m_q)^\gamma\right\}. \tag{2.27}$$

Since m_q tends to infinity with n (see (2.20), (2.21) and $\gamma < 1$, the second factor is of smaller order than the first factor for all sufficiently large n. This observation together with the upper bound yields that for any fixed i,

$$u_i^{(q)} \sim m_q^{(\tau-2)^{-(i+1)}}, \tag{2.28}$$

in the sense of (1.5). Note that $(\tau - 2)^{-1} > 1$, thus $u_i^{(q)}$ is growing and $\Gamma_i^{(q)} \supset \Gamma_{i+1}^{(q)}$.

To show that the initial stage (coupling to BPs) and the paths through shells has nonzero intersection, we will use the following claim:

Claim 2.4 *Let* X_i, $i = 1, \ldots, m$ *be i.i.d. random variables from distribution* $F_n^\star$ *or* $F^\star$. *Then*

$$\mathbb{P}\left(\max_{i \in [m]} X_i < \left(\frac{m}{h(m)} \right)^{1/(\tau-2)} \right) \le \exp \left\{ - \exp \left\{ \frac{C_1^\star}{(\tau-2)^\gamma} (\log m)^\gamma \right\} \right\} \to 0. \tag{2.29}$$

Proof We show it for $F_n^\star$. The proof for $F^\star$ is identical. Clearly

$$\mathbb{P}\left(\max_{i \in [m]} X_i < \left(\frac{m}{h(m)} \right)^{1/(\tau-2)} \right) = F_n^\star \left(\left(\frac{m}{h(m)} \right)^{1/(\tau-2)} \right)^m$$

$$\le \exp \left\{ m \left(1 - F_n^\star \left(\left(\frac{m}{h(m)} \right)^{1/(\tau-2)} \right) \right) \right\},$$

and the rest follows by using the function h from (2.22) as well as the form of $F_n^\star$ from (1.7), in particular the relation in (2.23). $\qquad \blacksquare$

Proof of Proposition 2.1, upper bound By the coupling established in Corollary 2.3, conditioned on the size of the last generation that we denote by m_q, the degrees in the last generation of the two BPs are an i.i.d. (either from $F_n^\star$ or from $F^\star$). Claim 2.4, applied conditionally on m_q, ensures that whp there are vertices with degree at least $u_0^{(q)} = (m_q / h(m_q))^{1/\tau-2}$ in the last generation of the two BPs, establishing that $\mathbb{P}_Y$-whp

$$\mathcal{C}_{t(n^{\varrho_n'})}^{(q)} \cap \Gamma_0^{(q)} \ne \varnothing. \tag{2.30}$$

The next step is to show that $\mathbb{P}_Y$-whp for all i such that $\Gamma_{i+1}^{(q)} \ne \varnothing$,

$$\Gamma_i^{(q)} \subset N(\Gamma_{i+1}^{(q)}), \tag{2.31}$$

where $N(S)$ stands for the set of vertices that are neighbors of S (Fig. 1).

This statement can be obtained by a modification of [6, Lemma 3.4] that we provide here for the reader's convenience. Recall that $H(\mathcal{A})$ denotes the number of half-edges attached to vertices in a set $\mathcal{A}$.

The algorithm to generate the configuration model makes it possible that when checking the connection of a vertex $v \in \Gamma_i^{(q)}$, we can start by pairing the half-edges of v one after another. Given $H(\Gamma_{i+1}^{(q)})$, the probability that a half-edge is not connected to any of the half-edges attached to vertices in $\Gamma_{i+1}^{(q)}$ is at most $1 - H(\Gamma_{i+1}^{(q)})/\ell_n$. Since we can pair at least $u_i^{(q)}/2$ half-edges before all the half-edges of v are paired[9], by a union bound for all $v \in H(\Gamma_i^{(q)})$,

$$\mathbb{P}_Y\left(\exists v \in \Gamma_i^{(q)} : v \nleftrightarrow \Gamma_{i+1}^{(q)} \mid H(\Gamma_{i+1}^{(q)}) \right) \le |\Gamma_i^{(q)}| \left(1 - \frac{H(\Gamma_{i+1}^{(q)})}{\ell_n} \right)^{u_i^{(q)}/2}$$

$$\le n \exp \left\{ -C u_i^{(q)} u_{i+1}^{(q)} (1 - F_n(u_{i+1}^{(q)})) \right\}. \tag{2.32}$$

By using the lower bound in (1.5) as well as the upper bound $u_{i+1}^{(q)} \le \left(u_i^{(q)} \right)^{1/(\tau-2)}$, (see also (2.23)),

$$u_i^{(q)} u_{i+1}^{(q)} \left(1 - F_n(u_{i+1}^{(q)}) \right) = h(u_i^{(q)}) L_n \left(u_{i+1}^{(q)} \right) \ge \exp \left\{ \widetilde{C} (\tau-2)^{-i} (\log m_q)^\gamma \right\},$$

---

[9] This is since in the worst case scenario the first $u_i^{(q)}/2$ half-edges are paired all back to half-edges of the same vertex v.

Fig. 1 An illustration of the layers and the mountain climbing phase at time $t(n^{\varrho'_n}) + 3$. Disclaimer: the degrees on the picture are only an illustration
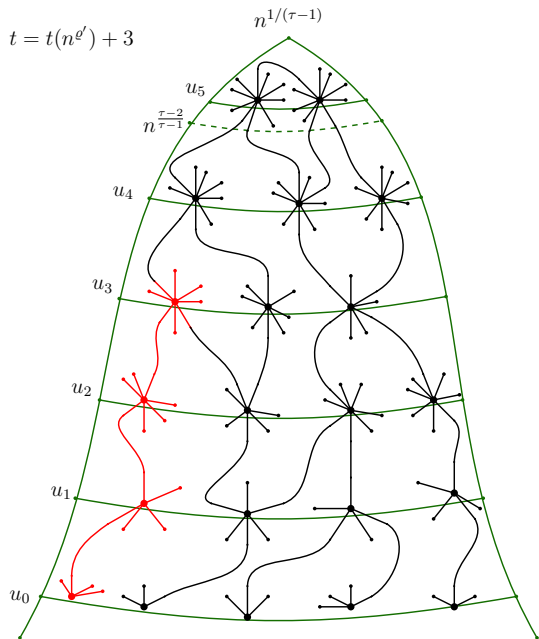
for some $\widetilde{C} > 0$. Since m_q ((2.20), (2.21)) tends to infinity with n, the bound in (2.32) tends to zero as $n \to \infty$ even when we sum over $i \geq 1$. This establishes (2.31).

The coupling to two BPs combined with (2.30) and (2.31) establishes the existence of a path to the hubs. This provides an *upper bound* on the distance between v_r, v_b and the hubs. It remains to calculate the length of these paths.

We write $i_{\star(q)}$ for the last index when $\Gamma_i^{(q)}$ is nonempty, i.e., by (1.4),

$$i_{\star(q)} := \sup \left\{ i : u_i^{(q)} \leq n^{\beta_n} \right\}. \tag{2.33}$$

Some calculation using the value of m_q from (2.20), combined with (2.28) and (2.27) shows that

$$i_{\star(r)} = -1 + \frac{\log \left(\beta_n / \left(\varrho'_n (\tau - 2)^{a_n^{(r)} - 1} \right) \right)}{|\log(\tau - 2)|} - b_n^{(r)}(\beta_n) + o_{\mathbb{P}_Y}(1), \tag{2.34}$$

where $b_n^{(r)}(\beta_n)$ is the fractional part of the previous term on the rhs. Using the value of $a_n^{(q)}$ from (2.19), plus the fact that $\{x - 1 + \{y\}\} = \{x + y\}$, we get that

$$b_n^{(r)}(\beta_n) = \left\{ \frac{\log(\beta_n / \varrho'_n)}{|\log(\tau - 2)|} + a_n^{(r)} - 1 \right\} = \left\{ \frac{\log(\beta_n / Y_r^{(n)}) + \log \log n}{|\log(\tau - 2)|} \right\}, \tag{2.35}$$

exactly as defined in (1.15). Similar calculation for $q = b$ yields

$$i_{\star(b)} = -1 + \frac{\log(\beta_n / \varrho'_n (\tau - 2)^{a_n^{(r)} - 1})) + \log(Y_r^{(n)} / Y_b^{(n)})}{|\log(\tau - 2)|} - b_n^{(b)}(\beta_n) + o_{\mathbb{P}_Y}(1), \tag{2.36}$$

where $b_n^{(b)}(\beta_n)$ is the fractional part of the previous term on the rhs. Again, some calculation yields that $b_n^{(b)}(\beta_n)$ is exactly as in (1.15). Note that the definition of ϱ'_n in (2.16) guarantees

that the ratio β_n/ϱ'_n is bounded, and hence $i_{\star(q)}$ is a tight random variable (measurable wrt $\mathcal{F}_{\varrho'_n}$). From (2.28) and (2.34) and (2.36) respectively, one can calculate that

$$u_{i_{\star(q)}}^{(q)} \sim n^{\beta_n(\tau-2)^{b_n^{(q)}(\beta_n)}}, \qquad (2.37)$$

and the error factor as in (2.27) is $o_{\mathbb{P}_Y}(1)$, since $i_{\star(q)}$ does not tend to infinity with n. Thus, the total length of the constructed path is, for $q \in \{r, b\}$,

$$T_q(\beta_n) = t(n^{\varrho'_n}) + i_{\star(q)} = \frac{\log\log(n^{\beta_n}) - \log Y_n^{(q)}}{|\log(\tau-2)|} - 1 - b_n^{(q)}(\beta_n), \qquad (2.38)$$

establishing the upper bound on $d_G(v_q, \text{hubs})$ in Proposition 2.1. Note that $T_q(\beta_n)$ only depends on the value ϱ'_n through the approximating variables $Y_n^{(q)}$, and $b_n^{(q)}(\beta_n)$ is exactly the fractional part of the expression on the rhs of $T_q(\beta_n)$. Since also $Y_n^{(q)} \xrightarrow{d} Y_q$ irrespective of the choice of ϱ'_n, this establishes that the choice of ϱ'_n is not relevant in the proof (at least not in the limit), but more a technical necessity. □

2.3 Upper Bound on the Degrees in the BFS

Now we turn towards providing a *matching lower bound* for the distance from the hubs. Similarly as in (2.26), let us define, for $q \in \{r, b\}$,

$$\begin{aligned}
\widehat{u}_0^{(q)} &:= (m_q h(m_q))^{1/(\tau-2)}, \\
\widehat{u}_{i+1}^{(q)} &:= \left(\widehat{u}_i^{(q)} h(\widehat{u}_i^{(q)})\right)^{1/(\tau-2)}, \\
\widehat{\Gamma}_i^{(q)} &:= \{v \in \text{CM}_n(\boldsymbol{d}) : d_v \geq \widehat{u}_i^{(q)}\},
\end{aligned} \qquad (2.39)$$

Note that $\widehat{u}_i^{(q)}$ grows faster than $u_i^{(q)}$ since here we multiply by h instead of dividing by it.

The next lemma handles the upper bound on the maximal-degree vertex reached at any time $t(n^{\varrho'_n}) + i$, but first some definitions. We say that a sequence of vertices and half-edges $\underline{\pi} := (\pi_0, s_0, t_1, \pi_1, s_1, t_2, \ldots, t_k, \pi_k)$ forms a *path* in $\text{CM}_n(\boldsymbol{d})$, if for all $0 < i \leq k$, the half-edges s_i, t_i are incident to the vertex π_i and (s_{i-1}, t_i) forms an edge between π_{i-1}, π_i.

For $q \in \{r, b\}$, we say that a path is *q-good* if $\deg(\pi_i) \leq \widehat{u}_i^{(q)}$ holds for every i. Otherwise we call it *q-bad*. We further decompose the set of q-bad paths in terms of where they 'turn' q-bad:

$$\begin{aligned}
Bad\mathcal{P}_k^{(q)} :=\Big\{& (\pi_0, s_0, t_1, \pi_1, s_1 \ldots, t_k, \pi_k) \text{ is a path,} \\
& \pi_0 \in \mathcal{C}_{t(n^{\varrho'_n})}^{(q)}, \ \deg(\pi_i) \leq \widehat{u}_i^{(q)} \ \forall i \leq k-1, \ \deg(\pi_k) > \widehat{u}_k^{(q)}\Big\}.
\end{aligned} \qquad (2.40)$$

The following lemma shows that q-bad paths $\mathbb{P}_Y$-whp do not occur:

Lemma 2.5 *For some constant $C > 0$, the following bound on the probability of having any bad paths holds for color $q \in \{r, b\}$:*

$$\mathbb{P}_Y\left(\exists k \in [0, i_{\star(q)}] : Bad\mathcal{P}_k^{(q)} \neq \varnothing\right) \leq C \exp\left\{-C(\log m_q)^\gamma\right\}. \qquad (2.41)$$

Before we prove this lemma, we need the following technical claim:

Claim 2.6 *Let $\left(D_{n,i}^\star\right)_{i \leq m}$ be i.i.d. from distribution $F_n^\star$ or $F^\star$. Then there exists a $0 < C < \infty$, so that*

$$\mathbb{P}_Y\left(\sum_{i=1}^m D_{n,i}^\star \geq (mh(m))^{1/(\tau-2)}\right) \leq \exp\left\{-C(\log m)^\gamma\right\}. \qquad (2.42)$$

Further, for any $y \in [0, n^{\beta_n}]$,

$$\sum_{d_\pi \geq y} \frac{d_\pi}{\ell_n} \leq L_n^{\star, up}(y) y^{2-\tau} \tag{2.43}$$

where we denote by $L_n^{\star, up}$ the upper bound in (1.5) on $L_n^{\star}$. Next, the empirical truncated second moment satisfies for all $y_n \to \infty$ and large enough n that

$$\sum_{\pi : d_\pi \leq y_n} \frac{d_\pi (d_\pi - 1)}{\ell_n} \leq \frac{2}{3 - \tau} (y_n)^{3-\tau} L_n^{\star, up}(y_n). \tag{2.44}$$

Proof The proof of (2.43) is the probably the easiest. Namely, by the definition of the empirical distribution as well as $F_n^{\star}$ the sum can be rewritten as follows:

$$\sum_{d_\pi \geq y} \frac{d_\pi}{\ell_n} = \sum_{j \geq y} \frac{\sum_{v \in [n]} j \mathbb{1}_{\{d_v = j\}}}{\ell_n} = 1 - F_n^{\star}(y - 1), \tag{2.45}$$

and an application of (1.7) establishes (2.43). Next, (2.44) can be rewritten similarly,

$$\sum_{\pi : d_\pi \leq y_n} \frac{d_\pi (d_\pi - 1)}{\ell_n} = \sum_{j \leq y_n} (j - 1) \frac{\sum_{v \in [n]} j \mathbb{1}_{\{d_v = j\}}}{\ell_n} = \sum_{j \leq y_n} (j - 1) \mathbb{P}(D_n^{\star} = j - 1)$$

$$\leq \sum_{j \leq y_n} \left(1 - F_n^{\star}(j)\right) \leq \sum_{j \leq y_n} \frac{L_n^{\star}(j)}{j^{\tau - 2}} \leq \sum_{j \leq y_n} \frac{L_n^{\star, up}(j)}{j^{\tau - 2}}. \tag{2.46}$$

To obtain the second line, we used the usual trick to relate the expectation to the tail of a distribution, namely,

$$\sum_{j \leq y_n} (j - 1) \mathbb{P}\left(D_n^{\star} = j - 1\right) = \sum_{j \leq y_n} \sum_{s \leq j - 1} \mathbb{P}\left(D_n^{\star} = j - 1\right) = \sum_{s \leq y_n - 1} \sum_{s < j \leq y_n} \mathbb{P}\left(D_n^{\star} = j - 1\right)$$

$$\leq \sum_{s \leq y_n - 1} \mathbb{P}\left(D_n^{\star} \geq s\right) = \sum_{s \leq y_n} \left(1 - F_n^{\star}(s)\right).$$

The condition that $y_n \to \infty$ as $n \to \infty$ enables us to apply the direct half of Karamata's theorem[10] (see [9, p. 26]), and obtain that for all large enough n, the following bound holds on the rhs of (2.46):

$$\sum_{j \leq y_n} \frac{L_n^{\star, up}(j)}{j^{\tau - 2}} \leq \frac{2}{3 - \tau} (y_n)^{3-\tau} L_n^{\star, up}(y_n), \tag{2.47}$$

finishing the proof of (2.44). The proof of (2.42) is the trickiest, we handle it with a truncation method. Let us shortly write $M_m := (mh(m))^{1/(\tau - 2)}$. First we use a union bound:

$$\mathbb{P}_Y \left(\sum_{i=1}^m D_{n,i}^{\star} \geq M_m\right) \leq \mathbb{P}\left(\exists i \leq m : D_{n,i}^{\star} \geq M_m\right)$$

$$+ \mathbb{P}\left(\sum_{i=1}^m D_{n,i}^{\star} \mathbb{1}_{\{D_{n,i}^{\star} < M_m\}} \geq M_m\right). \tag{2.48}$$

---

[10] Here we use that the *upper bound $L_n^{\star, up}$* on the function $L_n^{\star}$ is slowly varying.

Then, we estimate the first term on the rhs of (2.48) again by a union bound:

$$\mathbb{P}\left(\exists i \leq m : D_{n,i}^{\star} \geq M_m\right) \leq m\left(1 - F_n^{\star}(M_m)\right) = L_n^{\star}(M_m)/h(m). \tag{2.49}$$

We can use Markov's inequality on the second term on the rhs of (2.48):

$$\mathbb{P}_Y\left(\sum_{i=1}^{m} D_{n,i}^{\star} \mathbb{1}_{\left\{D_{n,i}^{\star} < M_m\right\}} \geq M_m\right) \leq \frac{m\mathbb{E}\left[D_{n,i}^{\star} \mathbb{1}_{\left\{D_{n,i}^{\star} < M_m\right\}}\right]}{M_m} \leq C\frac{mM_m^{3-\tau}L_n^{\star}(M_m)}{M_m}, \tag{2.50}$$

where we have used that the expectation in the numerator equals precisely the truncated empirical second moment as in (2.44) with $y_n := M_m$, thus this expectation can be handled in the same way as the rhs of (2.46). After elementary calculation, the sum of rhs of (2.49) and (2.50) equals

$$(C+1)\frac{L_n^{\star}(M_m)}{h(m)} \leq (C+1)\exp\left\{-\frac{C_2}{(\tau-2)^\gamma}(\log m)^\gamma\right\}. \tag{2.51}$$

This, together with (2.48) establishes (2.42).   □

Proof of Lemma 2.5 The statement can be proved by path-counting methods in the same way as [6, Lemma 5.2] in the Appendix of that paper. Some minor modifications are needed to that proof, for two reasons: First, Assumption 1.1 imposes an assumption on the *empirical degree distribution F_n* and the degrees are no longer i.i.d.. This makes certain estimates about truncated empirical moments easier (i.e., (2.43) and (2.44)). On the other hand, Assumption 1.1 is weaker than the assumption on the degrees there, compare to [6, (1.1)]. As a result, the recursion of $\widehat{u}_i^{(q)}$ uses the function h as a multiplier instead of the constant function $C \log n$. To make the argument easy to follow here, we recall the main steps in the proof and highlight differences only.

First of all, formulas [6, (A.3)–(A.5)] that count the expected number of bad paths apply word-for-word. That is, the expected number of paths in $\mathrm{CM}_n(\boldsymbol{d})$ through a fixed sequence of vertices $(\pi_0, \pi_1, \ldots, \pi_k)$ equals

$$\prod_{i=1}^{k} \frac{1}{\ell_n^\star - 2i + 1} d_{\pi_0}\left(\prod_{i=1}^{k-1} d_{\pi_i}(d_{\pi_i} - 1)\right) d_{\pi_k}, \tag{2.52}$$

where $\ell_n^\star$ denotes the number of unpaired half-edges at the moment of counting these paths. It is not hard to see (using the sizes m_q in (2.20) (2.21)) that $\ell_n^\star = \ell_n(1 - o_\mathbb{P}(1))$ when we apply this to a path emanating from $\mathcal{C}_{t(n^{\varrho_n'})}^{(q)}$.

This formula holds for any fixed sequence of vertices. We can count the expected number of q-bad paths in $Bad\mathcal{P}_k^{(q)}$ when we impose the same restrictions on π_i as in (2.40), and sum over all possible such options:

$$\mathbb{E}_Y\left[Bad\mathcal{P}_k^{(q)}|\right] \leq e^{Ck^2/\ell_n} \sum_{\pi_0 \in \mathcal{C}_{t(n^{\varrho_n'})}^{(q)}} d_{\pi_0} \prod_{i=1}^{k-1}\left(\sum_{\pi_i : d_{\pi_i} \leq \widehat{u}_i^{(b)}} \frac{d_{\pi_i}(d_{\pi_i} - 1)}{\ell_n^\star}\right)\left(\sum_{\substack{\pi_k \in [n] \\ d_{\pi_k} \geq \widehat{u}_k^{(b)}}} \frac{d_{\pi_k}}{\ell_n^\star}\right). \tag{2.53}$$

The next step is to estimate the different factors on the rhs of (2.53): here, the two proofs separate. In fact, we can use the bounds in Claim 2.6, and then we can spare all the arguments

between [6, (A.6)–(A.9)]. Let $\mathcal{E}_n := \left\{ \sum_{i=1}^{m_q} D_i^\star \leq \widehat{u}_0^{(q)} \right\}$, so that $\mathcal{E}_n$ holds $\mathbb{P}_Y$-whp by Claim 2.6, with $m := m_q$. Using the estimates (2.43) and (2.44) in Claim 2.6 with $y_n = \widehat{u}_i^{(q)}$, (2.53) turns into

$$\mathbb{E}_Y \left[|Bad\mathcal{P}_k^{(q)}| \mid \mathcal{E}_n \right] \leq \widehat{u}_0^{(q)} \cdot \left(\widehat{u}_k^{(q)} \right)^{2-\tau} L_n^{\star,\mathrm{up}} \left(\widehat{u}_k^{(q)} \right) \cdot \prod_{i=1}^{k-1} \left(\widehat{u}_i^{(q)} \right)^{3-\tau} \frac{2}{3-\tau} L_n^{\star,\mathrm{up}} \left(\widehat{u}_i^{(q)} \right). \tag{2.54}$$

This formula replaces [6, (A.11)]. It is an elementary calculation using the defining recursion (2.39) that $\widehat{u}_j^{(q)} \cdot \left(\widehat{u}_{j+1}^{(q)} \right)^{3-\tau} h \left(\widehat{u}_j^{(q)} \right) = \widehat{u}_{j+1}^{(q)}$. Applying this equation to $j = 0, \ldots, k-2$ sequentially, we arrive at the identity

$$\widehat{u}_0^{(q)} \cdot \left(\widehat{u}_k^{(q)} \right)^{2-\tau} h \left(\widehat{u}_{k-1}^{(q)} \right) \cdot \prod_{i=1}^{k-1} \left(\widehat{u}_i^{(q)} \right)^{3-\tau} h \left(\widehat{u}_{i-1}^{(q)} \right) = 1. \tag{2.55}$$

Comparing (2.54) to (2.55), we see that[11]

$$\mathbb{E}_Y \left[|Bad\mathcal{P}_k^{(q)}| \mid \mathcal{E}_n \right] \leq \prod_{i=0}^{k-1} \frac{2}{3-\tau} \frac{L_n^{\star,\mathrm{up}} \left(\widehat{u}_{i+1}^{(q} \right)}{h \left(\widehat{u}_i^{(q)} \right)}. \tag{2.56}$$

Next we show that for all n large enough, the rhs of (2.56) tends to zero even when summed over all $k \geq 1$. For this, using the recursion on $\widehat{u}_i^{(q)}$ in (2.39) as well as the function h in (2.22),

$$\begin{aligned}
L_n^{\star,\mathrm{up}} \left(\widehat{u}_{i+1}^{(q)} \right) &= \exp \left\{ C_1^\star \left(\log \left(\widehat{u}_i^{(q)} \right)^{1/(\tau-2)} + 2C_2 \left(\log \left(\widehat{u}_i^{(q)} \right)^{1/(\tau-2)} \right)^\gamma \right)^\gamma \right\} \\
&= \exp \left\{ C_1^\star \left(\log \left(\widehat{u}_i^{(q)} \right)^{1/(\tau-2)} \right)^\gamma \left(1 + 2C_2 \left(\log \left(\widehat{u}_i^{(q)} \right)^{1/(\tau-2)} \right)^{\gamma-1} \right)^\gamma \right\} \\
&\leq \exp \left\{ \frac{3C_1^\star}{2} \left(\log \left(\widehat{u}_i^{(q)} \right)^{1/(\tau-2)} \right)^\gamma \right\},
\end{aligned} \tag{2.57}$$

where in the inequality we have used that since $\gamma < 1$ and $\widehat{u}_i^{(q)}$ tends to infinity with n, for all n large enough the last factor in the rhs of the second line is at most $3/2$. The denominator of the ith factor in (2.56) has the exact same form, except there the constant multiplier in the exponent is at least $2C_1^\star$. Thus, the rhs of (2.54) is at most

$$\begin{aligned}
\mathbb{E}_Y \left[|Bad\mathcal{P}_k^{(q)}| \mid \mathcal{E}_n \right] &\leq \prod_{i=0}^{k-1} \frac{2}{3-\tau} \exp \left\{ -\frac{C_1^\star}{2} \left(\log \left(\left(\widehat{u}_i^{(q)} \right)^{1/(\tau-2)} \right) \right)^\gamma \right\} \\
&\leq \frac{2^k}{(3-\tau)^k} \exp \left\{ -\frac{C_1^\star}{2} \sum_{i=0}^{k-1} (\tau-2)^{-(i+2)\gamma} \left(\log m_q \right)^\gamma \right\}, \tag{2.58}
\end{aligned}$$

where we used the lower bound $\widehat{u}_i^{(q)} \geq m_q^{(\tau-2)^{-(i+1)}}$ that follows from the recursion (2.39) (see also (2.60) below) to obtain the second line. Since $(\tau-2)^{-\gamma} > 1$, the sum in the exponent

---

[11] Let us compare the inequality (2.56) with the bound $O((\log n)^{-k})$ on $\mathbb{E}_Y[|\mathcal{C}^{(b)}ad\mathcal{P}_k^{(q)}|]$ in [6] that can be found in the second formula after [6, (A.11)]. In [6] the recursion in (2.39) was used with the special choice $h(x) \equiv C \log n$, and $L_n^\star$ was assumed to be a bounded function. Using these, the bound $O((\log n)^{-k})$ is a special case of the bound here in (2.56). Thus, (2.56) is a generalisation of the bound in [6] for general choice of h in the recursion (2.39). The generalisation was necessary since in this paper we allow a wide range of L_n in (1.4) while in [6] the more restrictive $0 < c \leq L_n \leq C < \infty$ assumption was set.

is of order $(\tau - 2)^{-k\gamma} (\log m_q)^\gamma$. So, by Markov's inequality we obtain for some constant $C > 0$ that

$$\mathbb{P}_Y\big(\exists k \geq 1, Bad\mathcal{P}_k^{(q)} \neq \varnothing\big) \leq \sum_{k=1}^\infty \mathbb{E}_Y\big[\big|Bad\mathcal{P}_k^{(q)}\big|\big] \leq C \exp\big\{-C\left(\log m_q\right)^\gamma\big\} \to 0 \tag{2.59}$$

as $n \to \infty$, since m_q is a positive power of n under $\mathbb{P}_Y$ (see (2.20), (2.21)). We yet have to add the case $k = 0$: note that $Bad\mathcal{P}_0^{(q)} \neq \varnothing$ means that there exists a vertex in the last generation of the BP with degree at least $\widehat{u}_0^{(q)}$. We have already estimated this probability in (2.49), and the error term obtained in (2.51) can be merged into the rhs of (2.59), establishing the statement of the lemma in (2.41). $\square$

Proof of Proposition 2.1, lower bound We argue that $T_q(\beta_n)$ is also whp a lower bound to reach the hubs, that is, there is whp no path to the hubs shorter than $T_q(\beta_n)$.

On the event $\big\{\forall k \in [0, i_{\star(q)}] : Bad\mathcal{P}_k^{(q)} = \varnothing, q \in \{r, b\}\big\}$, that occurs $\mathbb{P}_Y$-whp by Lemma 2.5, we can use the upper bound $\widehat{u}_i^{(q)}$ on the degrees at time $t(n^{\varrho'_n}) + i$ for all $i \leq i_{\star(q)}$. Hence we obtain that the time it takes to reach a vertex of degree at least $n^{(\tau-2)/(\tau-1)}$ is at least

$$\widehat{i}_{\star(q)} := \inf\left\{i : \widehat{u}_i^{(q)} \geq n^{\beta_n(\tau-2)}\right\},$$

which, considering the double exponential growth of $\widehat{u}_i^{(q)}$ by powers of $1/(\tau - 2)$, is similar to the definition of $i_{\star(q)}$ in (2.33). The lower bound follows once we show that $\widehat{i}_{\star(q)} = i_{\star(q)}$ holds $\mathbb{P}_Y$-whp. For this it is enough to show that (2.28) holds also for $\widehat{u}_i^{(q)}$. From the recursion (2.39),

$$\widehat{u}_i^{(q)} = m_q^{(\tau-2)^{-(i+1)}} \prod_{k=1}^{i+1} h\left(\widehat{u}_{i-k}^{(q)}\right)^{(\tau-2)^{-k}}. \tag{2.60}$$

After a somewhat lengthy calculation, using a similar argument as in the second and third line of (2.57) recursively, we obtain that the product on the rhs is at most

$$\exp\left\{2C_2 \left(\log\left(m_q^{1/(\tau-2)}\right)\right)^\gamma \frac{1}{(\tau-2)^{i+1}} \sum_{k=1}^i \frac{1}{(\tau-2)^{k(\gamma-1)}}\right\}$$

$$\leq \exp\left\{\widetilde{C} \left(\log\left(m_q^{1/(\tau-2)}\right)\right)^\gamma \frac{1}{(\tau-2)^{i+1}}\right\},$$

since $\gamma < 1$. Recall that m_q tends to infinity with n and comparing this to (2.60) as well as to the definition of $\sim$ in (1.5). Nevertheless, the product in (2.60) is of much smaller order than the main term $m_q^{(\tau-2)^{-(i+1)}}$ and thus we obtain that (2.28) holds and also that $\widehat{i}_{\star(q)} = i_{\star(q)}$ $\mathbb{P}_Y$-whp. This finishes the proof of the lower bound. $\square$

3 Early Meeting is Unlikely

For the lower bound of Theorem 1.8, we crucially use the following proposition that shows that the two explorations are disjoint, i.e., the vertices at distance at most $T_b(\beta_n)$ away from v_b are all different from the vertices that are distance at most $T_r(\beta_n)$ away from v_r:

Proposition 3.1 *Let us consider the configuration model on n vertices with empirical degree distribution that satisfies Assumption* 1.1, *and let v_r, v_b be two uniformly chosen vertices. The event*

$$\mathcal{C}^{(r)}_{T_r(\beta_n)} \cap \mathcal{C}^{(b)}_{T_b(\beta_n)} = \varnothing \tag{3.1}$$

holds $\mathbb{P}_Y$-whp. Further, the total number of half-edges attached to vertices in $\mathcal{C}^{(r)}_{T_r(\beta_n)}, \mathcal{C}^{(b)}_{T_b(\beta_n)}$ is the same order of magnitude as the degree of $v^\star_q$ in Proposition 2.1 *up to smaller order correction terms. That is, $\mathbb{P}_Y$-whp,*

$$H\left(\mathcal{C}^{(r)}_{T_r(\beta_n)}\right) \sim n^{\beta_n(\tau-2)^{b^{(r)}_n(\beta_n)}} \quad and \quad H\left(\mathcal{C}^{(b)}_{T_b(\beta_n)}\right) \sim n^{\beta_n(\tau-2)^{b^{(b)}_n(\beta_n)}}. \tag{3.2}$$

Proof Recall that we write $\mathbb{P}_Y(\cdot), \mathbb{E}_Y[\cdot]$ for probabilities of events and expectations of random variables conditioned on $\mathcal{F}_{\varrho'_n}$. Recall from Lemma 2.5 that the event NoBad $:= \left\{BadP^{(q)}_k = \varnothing \ \forall k \le i_{\star(q)} \text{ for } q \in \{r, b\}\right\}$ holds $\mathbb{P}_Y$-whp. Since for any event A, $\mathbb{P}_Y(A) \ge \mathbb{P}_Y(A \mid \text{NoBad})\mathbb{P}_Y(\text{NoBad})$, it is enough to show that the event in (3.1) holds with probability tending to 1 when conditioned on NoBad.

To prove the proposition we first calculate the total number of free (unpaired) half-edges going out of the set $\mathcal{C}^{(r)}_{T_q(\beta_n)-\ell}$, (that we denote by $H\left(\mathcal{C}^{(r)}_{T_q(\beta_n)-\ell}\right)$), for any $\ell \in [0, i_{\star(q)}], q \in \{r, b\}$. We do this by counting the number of paths *with free ends*: we say that a sequence of vertices and half-edges $(\pi_0, s_0, t_1, \pi_1, s_1, t_2, \ldots, t_k, \pi_k, s_k)$ forms a *free-ended path* of length k in $\text{CM}_n(\boldsymbol{d})$, if for all $0 < i \le k$, the half-edges s_i, t_i are incident to the vertex π_i and (s_{i-1}, t_i) forms an edge between vertices π_{i-1}, π_i. Clearly, since the same vertex might be approached on several paths, the total number of free half-edges in $\mathcal{C}^{(q)}_{T_q(\beta_n)-\ell}$ can be bounded from above by the number of free-ended paths of length $T_q(\beta_n) - \ell$, starting from v_r. By the definition of $BadP^{(q)}_k$ in (2.40), on the event NoBad at time $t(n^{\varrho'_n}) + i$, $\widehat{u}^{(q)}_i$ defined in (2.39) is an upper bound on the degrees of color q vertices. We write $\mathcal{N}_k(\mathcal{A}, \text{free})$ for the set of, and $N_k(\mathcal{A}, \text{free})$ for the total number of, k-length free-ended paths starting from an unpaired half-edge that belongs to the set $\mathcal{A}$. Then, since $T_q(\beta_n) = t(n^{\varrho'_n}) + i_{\star(q)}$ (see (2.38)), for any $\ell \le i_{\star(q)}$,

$$H\left(\mathcal{C}^{(r)}_{T_q(\beta_n)-\ell}\right) \le N_{i_{\star(q)}-\ell}\left(\mathcal{C}^{(r)}_{t(n^{\varrho'_n})}, \text{free}\right), \tag{3.3}$$

and recall that $\mathcal{C}^{(r)}_{t(n^{\varrho'_n})}$ is coupled to the branching process described in Sect. 2.1. Hence, the degrees in the last generation of the BP phase are i.i.d. having distribution $D^\star_n$ satisfying (1.7). When counting free-ended paths through fixed vertices $(\pi_0, \ldots, \pi_k)$, (2.52) should be modified so that we have to choose two half-edges also from the end vertex π_k, thus there is an additional factor $d_{\pi_k} - 1$ that should multiply (2.52). The effect of this on (2.53) is that the factor containing π_k can be merged into the previous factor:

$$\mathbb{E}_Y\left[N_{i_{\star(q)}-\ell}\left(\mathcal{C}^{(r)}_{t(n^\varrho)}, \text{free}\right) \mid \text{NoBad}\right]$$

$$\le e^{Ci^2_{\star(q)}/\ell_n} \sum_{\pi_0 \in \mathcal{C}^{(q)}_{t(n^{\varrho'_n})}} d_{\pi_0} \cdot \prod_{i=1}^{i_{\star(q)}-\ell}\left(\sum_{\pi_i \in \Lambda_i} \frac{d_{\pi_i}(d_{\pi_i}-1)}{\ell_n}\right), \tag{3.4}$$

where we have applied the restriction that is valid under the event NoBad: $\pi_i \in \Lambda_i$, with $\Lambda_i = \left\{v \in [n] : D_v \le \widehat{u}^{(r)}_i\right\}$. Note that we could use again that $\ell^\star_n = \ell_n(1 - o_{\mathbb{P}}(1))$ by the

same argument that was used after formula (2.52). Using (2.42) and (2.44) from Claim 2.6, we obtain that

$$\mathbb{E}_Y\left[N_{i_{\star(r)}-\ell}\left(\mathcal{C}^{(r)}_{t(n^\varrho)}, \text{free}\right) \mid \text{NoBad}\right] \le \widehat{u}^{(r)}_0 \cdot \left(\prod_{i=1}^{i_{\star(r)}-\ell} \frac{2}{3-\tau}\left(\widehat{u}^{(q)}_i\right)^{3-\tau} L_n^{\star\text{up}}\left(\widehat{u}^{(q)}_i\right)\right) \mathrm{e}^{2i^2_{\star(q)}/\ell_n}.$$
(3.5)

Note that this is similar to (2.54). Indeed, we again sequentially apply the identity $\widehat{u}^{(q)}_j \cdot (\widehat{u}^{(q)}_{j+1})^{3-\tau} h(\widehat{u}^{(q)}_j) = \widehat{u}^{(q)}_{j+1}$, and then (3.5) turns into

$$\mathbb{E}_Y\left[N_{i_{\star(r)}}\left(\mathcal{C}^{(r)}_{t(n^\varrho)}, \text{free}\right) \mid \text{NoBad}\right] \le \widehat{u}^{(q)}_{i_{\star(r)}-\ell} \left(\prod_{i=0}^{i_{\star(r)}-\ell-1} \frac{2}{3-\tau} \frac{L_n^{\star,\text{up}}\left(\widehat{u}^{(q}_{i+1}\right)}{h\left(\widehat{u}^{(q)}_i\right)}\right) \mathrm{e}^{2i^2_{\star(q)}/\ell_n}.$$
(3.6)

Combining this with Markov's inequality and a union bound gives

$$\mathbb{P}_Y\left(\exists \ell \le i_{\star(q)} : H\left(\mathcal{C}^{(r)}_{T_q(\beta_n)-\ell}\right) \ge \widehat{u}^{(q)}_{i_{\star(r)}-\ell} \mid \text{NoBad}\right)$$
$$\le \mathrm{e}^{2i^2_{\star(q)}/\ell_n} \sum_{\ell=0}^{i_{\star(q)}} \left(\prod_{i=0}^{i_{\star(r)}-\ell-1} \frac{2}{3-\tau} \frac{L_n^{\star,\text{up}}\left(\widehat{u}^{(q}_{i+1}\right)}{h\left(\widehat{u}^{(q)}_i\right)}\right).$$
(3.7)

Recall again that $i_{\star(q)}$ is a tight random variable measurable wrt $\mathcal{F}_{\varrho'_n}$ (see (2.34) and (2.36)), and $\ell_n = \mathbb{E}[D_n]n$ is of order n. Thus the first factor on the rhs is $1 + o_{\mathbb{P}_Y}(1)$. Further, in the analysis below (2.56) we have showed that the sum in the rhs of (3.7) is at most the rhs of (2.59). Thus we obtain

$$\mathbb{P}_Y\left(\exists \ell \le i_{\star(q)} : H\left(\mathcal{C}^{(r)}_{T_q(\beta_n)-\ell}\right) \ge \widehat{u}^{(q)}_{i_{\star(q)}-\ell} \mid \text{NoBad}\right) \le C \exp\left\{-C(\log m_q)^\gamma\right\}.$$
(3.8)

Now, to see that $\mathcal{C}^{(r)}_{T_r(\beta_n)}$ and $\mathcal{C}^{(b)}_{T_b(\beta_n)}$ are disjoint, we apply the following procedure: It is easy to see that $H\left(\mathcal{C}^{(r)}_{T_r(\beta_n)-\ell}\right)$ is maximised at $\ell = 0$. Hence, we grow the red cluster first until time $T_r(\beta_n)$, and then stop it. Then, we grow the blue cluster step by step, looking at the pairs of half-edges[12] in $\mathcal{H}\left(\mathcal{C}^{(b)}_1\right), \mathcal{H}\left(\mathcal{C}^{(b)}_2\right), \ldots, \mathcal{H}\left(\mathcal{C}^{(b)}_{T_b(\beta_n)-1}\right)$, and at each step we check whether any of the half edges paired are actually paired to a red half-edge. If this happens for any time before or at $T_b(\beta_n) - 1$, then an early connection happens and the distance is at most $T_b(\beta_n) + T_r(\beta_n)$. (Note that the distance is $T_r(\beta_n) + i$ if we pair a blue half-edge attached to $\mathcal{C}^{(b)}_{i-1}$ to a red half-edge.)

The probability that there is a connection before or at $t(n^{\varrho'_n})$ is of the same order of magnitude as the probability that there is a connection at time $t(n^{\varrho'_n})$, since the total degree in the whole BP is the same order of magnitude as the total degree in the last generation, thus it is enough to investigate the probability that $\mathcal{H}\left(\mathcal{C}^{(b)}_{T_b(\beta_n)-\ell}\right)$ connects to $\mathcal{H}\left(\mathcal{C}^{(r)}_{T_r(\beta_n)}\right)$ for some $\ell \le i_{\star(b)}$. This probability is at most

$$\mathbb{P}_Y\left(\mathcal{C}^{(b)}_{T_b(\beta_n)-\ell} \leftrightarrow \mathcal{C}^{(r)}_{T_r(\beta_n)} \mid H\left(\mathcal{C}^{(b)}_{T_b(\beta_n)-\ell}\right), H\left(\mathcal{C}^{(r)}_{T_r(\beta_n)}\right)\right) \le \frac{H\left(\mathcal{C}^{(b)}_{T_b(\beta_n)-\ell}\right) H\left(\mathcal{C}^{(r)}_{T_r(\beta_n)}\right)}{\ell_n(1 + o(1))}.$$
(3.9)

---

[12] Recall that $\mathcal{H}(A)$, $H(\mathcal{A})$ denote the set and number of half-edges attached to vertices in the set $\mathcal{A}$.

Let us write $\mathcal{D}_n := \left\{ H\left(\mathcal{C}^{(b)}_{T_b(\beta_n)-\ell} \right) \leq \widehat{u}_{i_{\star(b)}-\ell}, \forall \ell \in [i_{\star(b)}] \right\}$. Then by (3.7), $\mathcal{D}_n$ happens $\mathbb{P}_Y$-whp. Using this, we sum the bound on the rhs over $\ell \in [i_{\star(b)}]$, using (2.28), to obtain that

$$\mathbb{P}_Y\left(\mathcal{C}^{(r)}_{T_r(\beta_n)} \cap \mathcal{C}^{(b)}_{T_b(\beta_n)} \neq \varnothing \mid \mathcal{D}_n \right) \leq \frac{\widehat{u}^{(r)}_{i_{\star(r)}}}{\ell_n} \sum_{\ell=1}^{i_{\star(b)}} \widehat{u}^{(q)}_{i_{\star(b)}-\ell} \lesssim \frac{n^{\beta_n(\tau-2)^{b^{(r)}_n(\beta_n)}}}{\ell_n} \sum_{\ell=1}^{i_{\star(b)}} n^{\beta_n(\tau-2)^{b^{(b)}_n(\beta_n)+\ell}},$$

$$(3.10)$$

where we recall that $\lesssim$ means inequality up to multiplicative factors that are of order at most $\exp\{(\log n)^\theta\}$ for some $\theta \in [0, 1)$, as in Definition 1.5. Since $\ell_n = \mathbb{E}[D_n]n$ is of order n, the exponent of n in the dominant term in the numerator is

$$\beta_n\left((\tau-2)^{b^{(r)}_n(\beta_n)} + (\tau-2)^{b^{(b)}_n(\beta_n)+1} \right) < 1,$$

as long as $\beta_n < 1/(\tau-1)$, since $b^{(b)}_n(\beta_n), b^{(r)}_n(\beta_n) \in [0, 1)$. When $\beta_n = 1/(\tau-1)$, the strict inequality still holds as long as $(b^{(b)}_n(\beta_n), b^{(r)}_n(a\beta_n)) \neq (0, 0)$, an event that happens with probability 1 under Assumption 1.7, since $b^{(q)}_n(\beta_n) = 0$ is only possible if $Y^{(n)}_q$ takes values in a measure 0 discrete set, see (1.15). The probability of this event tends to 0 under Assumption 1.7.

For large enough n the multiplicative factors hidden in the $\sim$ sign on the rhs of (3.10) are negligible, thus the rhs of (3.10) tends to zero with n. This finishes the proof of the proposition. $\qquad\blacksquare$

4 Distances in the Graph

Proof of Theorem 1.8 The lower bound is easier, since we can use the first moment method (i.e., Markov's inequality) on the number of paths emanating from $H(\mathcal{R}_{T_r(\beta_n)})$ and $H(\mathcal{B}_{T_b(\beta_n)})$ and connecting to each other to obtain a lower bound. Thus, let us start counting paths of length $z+1$ (that is, z vertices in between) connecting $\mathcal{H}\left(\mathcal{C}^{(q)}_{T_q(\beta_n)} \right)$, for $q \in \{r, b\}$. Starting with (2.52), the restriction now is that $\pi_0 \in \mathcal{C}^{(r)}_{T_r(\beta_n)}$, while $\pi_{z+1} \in \mathcal{C}^{(b)}_{T_b(\beta_n)}$, and there are no restrictions on the in-between vertices. Thus, we obtain a similar formula as in (2.53), except that the restrictions on the vertices π_i are now different:

$$\mathbb{E}_Y\left[N_z\left(\mathcal{C}^{(r)}_{T_r(\beta_n)}, \mathcal{C}^{(b)}_{T_b(\beta_n)} \right) \mid \text{NoBad} \right]$$

$$\leq e^{Cz^2/\ell_n} \sum_{\pi_0 \in \mathcal{C}^{(r)}_{T_r(\beta_n)}} d_{\pi_0} \cdot \prod_{i=1}^{z} \left(\sum_{\pi_i \in [n]} \frac{d_{\pi_i}(d_{\pi_i}-1)}{\ell^\star_n} \right) \cdot \sum_{\pi_{z+1} \in \mathcal{C}^{(b)}_{T_b(\beta_n)}} \frac{d_{\pi_{z+1}}}{\ell^\star_n}. \qquad (4.1)$$

We have to check that in this case $\ell^\star_n = \ell_n(1 - o_\mathbb{P}(1))$ is still satisfied. This follows from the proof of Proposition 3.1 and the fact that Lemma 2.5 holds and thus the total number of used half-edges can be bounded from above as the sum of $\sim \widehat{u}^{(q)}_i$ over $i \leq i_{\star(q)}$ and $q \in \{r, b\}$.

The first and last sum are handled by (3.2) in Proposition 3.1. We would like to estimate the sums within the product sign on the rhs. For this, recall that $D^\star_n$ stands for the (degree-1) of a vertex that a uniformly chosen half-edge in ℓ_n is attached to. $D^\star_n$ then follows distribution

$F_n^\star$, see (1.6). By Claim 2.6, (2.44), the size-biased empirical moment of $D_n^\star$ is

$$\mathbb{E}[D_n^\star] = \sum_{\substack{v \in [n] \\ d_v \leq n^{\beta_n}}} \frac{d_v(d_v - 1)}{\ell_n} \leq n^{\beta_n(3-\tau)} L_n^{\star, \mathrm{up}}(n^{\beta_n}) \lesssim n^{\beta_n(3-\tau)}. \tag{4.2}$$

Applying this inequality on the rhs of (4.1) yields that

$$\mathbb{E}_Y\left[N_z\left(\mathcal{C}_{T_r(\beta_n)}^{(r)}, \mathcal{C}_{T_b(\beta_n)}^{(b)}\right) \mid \mathrm{NoBad}\right] \lesssim n^{-1} n^{\beta_n(\tau-2)^{b_n^{(r)}(\beta_n)}} n^{\beta_n(\tau-2)^{b_n^{(r)}(\beta_n)}} n^{z\beta_n(3-\tau)}. \tag{4.3}$$

By Markov's inequality, the probability that there is at least one path connecting $\mathcal{C}_{T_r(\beta_n)}^{(r)}, \mathcal{C}_{T_b(\beta_n)}^{(b)}$ with $z + 1$ edges can be bounded from above by the expected number of connections, so we obtain the bound

$$\mathbb{P}_Y\left(\mathrm{d}_G\left(\mathcal{C}_{T_r(\beta_n)}^{(r)}, \mathcal{C}_{T_b(\beta_n)}^{(b)}\right) \leq z + 1 \mid \mathrm{NoBad}\right) \lesssim n^{-1+\beta_n\left((\tau-2)^{b_n^{(r)}(\beta_n)} + (\tau-2)^{b_n^{(b)}(\beta_n)} + z(3-\tau)\right)}. \tag{4.4}$$

So, the two clusters are $\mathbb{P}_Y$-whp disjoint as long as this quantity tends to zero. The smallest value of $z \in \mathbb{N}$ when the rhs of (4.4) does *not* tend to zero is

$$z_n^\star := \inf\left\{z \in \mathbb{N} : (\tau - 2)^{b_n^{(r)}(\beta_n)} + (\tau - 2)^{b_n^{(b)}(\beta_n)} + z(3 - \tau) > 1/\beta_n\right\}$$
$$= \left\lceil \frac{1/\beta_n - (\tau - 2)^{b_n^{(r)}(\beta_n)} - (\tau - 2)^{b_n^{(b)}(\beta_n)}}{3 - \tau} \right\rceil. \tag{4.5}$$

Since $z_n^\star$ counts the number of *vertices* needed between $\mathcal{C}_{T_r(\beta_n)}^{(r)}$ and $\mathcal{C}_{T_b(\beta_n)}^{(b)}$, and we would like to count the number of *edges*[13], as long as the number of edges between $\mathcal{C}_{T_r(\beta_n)}^{(r)}$ and $\mathcal{C}_{T_b(\beta_n)}^{(b)}$ is at most $z_n^\star$, the bound in (4.4) tends to zero as $n \to \infty$. This in turn means that $\mathbb{P}_Y$-whp there is no path of length $T_r(\beta_n) + T_b(\beta_n) + z_n^\star$ connecting v_r, v_b. Thus, we obtain that $\mathbb{P}_Y$-whp:

$$\mathrm{d}_G(v_r, v_b) \geq T_r(\beta_n) + T_b(\beta_n) + \left\lceil \frac{1/\beta_n - (\tau - 2)^{b_n^{(r)}(\beta_n)} - (\tau - 2)^{b_n^{(b)}(\beta_n)}}{3 - \tau} \right\rceil + 1. \tag{4.6}$$

This completes the proof of the lower bound on $\mathrm{d}_G(v_r, v_b)$ in Theorem 1.8. For the upper bound on $\mathrm{d}_G(v_r, v_b)$, we expect the existence of a path of length as the rhs of (4.6). To be able to show this, we apply the second moment method. Recall that we have already constructed paths of length $T_q(\beta_n)$ between v_q and a vertex $v_q^\star$, where $d_{v_q^\star}$ as in (2.37). We calculate the expected number and variance of paths of length $z^\star + 1$ connecting $v_r^\star$ to $v_b^\star$, with certain restrictions. Namely, the formula for the variance turns out to be simpler and easier if we count paths where the ith vertex on the path falls into a different (and disjoint) set for all $i \geq 0$. The reason why the variance is easier to calculate is that two possible paths can overlap only in fairly simple ways (see [6, Figure 7]).

Note that since $b_n^{(b)}, b_n^{(r)} \in [0, 1)$,

$$z_n^\star + 2 \leq \left\lceil \frac{1/\beta_n - 2(\tau - 2)}{3 - \tau} \right\rceil + 2 =: M_{\beta_n}. \tag{4.7}$$

---

[13] The number of edges on a path with k in-between vertices is $k + 1$.

Now we divide the set of vertices into M_n many roughly equal disjoint sets. We denote the ith set by Δ_i. By roughly equal we mean that the following inequalities hold for some $0 < c_1 < c_2 < \infty$

$$\nu_i^{\text{new}} := \sum_{v \in \Delta_i} \frac{d_i(d_i - 1)}{\ell_n} \in \left[\frac{c_1}{M_{\beta_n}}, \frac{c_2}{M_{\beta_n}}\right] \cdot \mathbb{E}[D_n^\star] \tag{4.8}$$

$$\kappa_i^{\text{new}} := \sum_{v \in \Delta_i} \frac{d_i(d_i - 1)(d_i - 2)}{\ell_n} \in \left[\frac{c_1}{M_{\beta_n}}, \frac{c_2}{M_{\beta_n}}\right] \cdot \mathbb{E}[D_n^\star(D_n^\star - 1)]. \tag{4.9}$$

This can be done as long as we distribute the vertices in the intervals $[n^{\beta_n(1-\varepsilon)}, n^{\beta_n}]$ in an approximately uniform way. We require that $v_r^\star \in \Delta_0$ and $v_b^\star \in \Delta_{M_{\beta_n}}$ in the partitioning. We will count the paths on vertices $(v_r^\star := v_0, v_1, \ldots, v_z, v_{z+1} := v_b^\star)$ that satisfy the property that the jth vertex falls into Δ_j when $j \leq z/2$ and it falls into $\Delta_{M_{\beta_n}+1-j}$ when $j > z/2$.[14] As a result of the restriction, the proof of [6, Lemma 7.1, (7.5)] applies word by word. This proof bounds the expected number and variance of restricted paths between vertices a, b with k vertices in between. In this proof, we only need to replace ν_i by ν_i^{new} as in (4.8) and κ_i by κ_i^{new} as in (4.9). In our case the degree of vertex $a := v_r^\star$ is $d_a := u_{i_{\star(r)}}^{(r)}$ while the degree of vertex $b := v_b^\star$ is $d_b := u_{i_{\star(b)}}^{(b)}$.

First we need a lower bound on the expected number of paths of length $z + 1$ between $v_r^\star, v_b^\star$. We expect it to be of a similar order of magnitude as the upper bound in (4.3). The differences between a lower and an upper bound (see (4.1)) are the following: (1) the first and last factor in (4.1) in lower bound changes, since we only count the degree of $v_q^\star$, (2) in the middle factor in (4.1) we have to apply the restriction that $\pi_j \in \Delta_j$ for $j < z/2$ while $\pi_j \in \Delta_{M_{\beta_n}-j}$ for $j > z/2$ instead of summing over all vertices in $[n]$.

Using (2.3), combined with (4.8) yields that we have the upper and lower bound

$$\mathbb{E}_Y[N_{z_n^\star}(v_r^\star, v_b^\star)] \overset{\in}{\sim} \frac{n^{\beta_n(\tau-2)b_n^{(b)}(\beta_n)}n^{\beta_n(\tau-2)b_n^{(r)}(\beta_n)}}{\ell_n}n^{z\beta_n(3-\tau)} \cdot \left[(c_1/M_{\beta_n})^z, (c_2/M_{\beta_n})^z\right], \tag{4.10}$$

where $\overset{\in}{\sim}$ means containment in an interval, where an additional factor of order at most $\exp\{\pm(\log n^{\beta_n})^\theta\}$ for some $\theta < 1$, as defined in Definition 1.5, might multiply the prefactors of the interval. This additional factor comes from $v_q^\star$ not being precisely equal to $n^{\beta_n(\tau-2)b_n^{(q)}(\beta_n)}$, as well as $\mathbb{E}[D_n^\star]$ in (4.2) is not precisely equal to $n^{\beta_n(3-\tau)}$.

By the definitions of M_{β_n} and $z_n^\star$ in (4.7) and (4.5) and the bound $\beta_n \geq (\log n)^{-\gamma}$ for some $\gamma < 1$, for $i \in \{1, 2\}$, for all $z \leq z_n^\star$,

$$(c_i/M_{\beta_n})^z \geq \exp\left\{-\gamma \log\log n \cdot (\log n)^\gamma + \log c_i (\log n)^\gamma\right\} \geq \exp\{-(\log n)^\theta\}$$

for any $\theta \in (\gamma, 1)$ and n sufficiently large, so the lower bound on $\mathbb{E}_Y[N_{z^\star}(v_r^\star, v_b^\star)]$ in (4.10) fits the Definition 1.5 when we use that ℓ_n is of order n. Combining this with the upper bound in (4.3), we arrive at the desired

$$\mathbb{E}_Y[N_z(v_r^\star, v_b^\star)] \sim n^{-1}n^{\beta_n(\tau-2)b_n^{(b)}(\beta_n)}n^{\beta_n(\tau-2)b_n^{(r)}(\beta_n)}n^{z\beta_n(3-\tau)}. \tag{4.11}$$

The smallest value of z for which this expression tends to infinity (and not to 0) as $n \to \infty$ is precisely $z_n^\star$ defined in (4.5).

---

[14] This somewhat weird containment is needed since $z + 2 < M_{\beta_n}$ can also occur.

By Chebyshev's inequality,

$$\mathbb{P}_Y(N_{z_n^\star}(v_r^\star, v_b^\star) = 0) \leq \frac{\mathrm{Var}_Y[N_{z_n^\star}(v_r^\star, v_b^\star)]}{\mathbb{E}_Y[N_{z_n^\star}(v_r^\star, v_b^\star)]^2} \tag{4.12}$$

and thus, to show the $\mathbb{P}_Y$-whp existence of at least 1 path of length $z_n^\star + 1$, it is enough to show that the variance is of smaller order than the expectation squared.

Thus next we calculate the variance of $N_{z_n^\star}(v_r^\star, v_b^\star)$. Here we rely on the proof of [6, Lemma 7.1]. Unfortunately here the appendix of the paper does not state the variance independently of the statement of [6, Lemma 7.1]. However, one can word-by-word follow the derivation of the variance, starting from the formula before [6, (A.19)] until [6, (A.25)]. The first occurence where the setting of this paper deviates from that paper is [6, (A.26)]: in this equation, in the last error factor $(1 + \varepsilon)^k$, ε can be set to 0. This can be done since the events in $\mathcal{E}_n'$ defined in [6, (A.15)] hold for $\varepsilon = 0$ under Assumption 1.1.

A crucial observation in the proof there is the inequality in [6, (A.27)] stating that $\kappa_i/v_i^2 \leq \kappa_1/v_1^2$. This allows us to replace every occurrence of κ_i/v_i^2 by κ_1/v_1^2 in that proof. Note that this inequality in our case is not valid, however, with our choice of the Δ_i it is true that

$$\frac{\kappa_i^{\mathrm{new}}}{(v_i^{\mathrm{new}})^2} \leq C \frac{\kappa_1^{\mathrm{new}}}{(v_1^{\mathrm{new}})^2},$$

with $C := c_2/c_1^2$ from (4.8). Thus when following the proof, we are allowed to replace every occurrence of κ_i/v_i^2 by $C\kappa_1^{\mathrm{new}}/(v_1^{\mathrm{new}})^2$. Similarly, $1/v_{k-1}$ can be replaced by $\widetilde{C}/v_1^{\mathrm{new}}$. If we do this replacement in [6, (A.28)], and thus the geometric sums in the formula before [6, (A.29)] yield that in [6, (A.29)] we should replace the two occurrences of $v_{k-1}/(v_{k-1} - 1)$ in [6, (A.29)] by

$$\frac{v_1^{\mathrm{new}}}{v_1^{\mathrm{new}} - \widetilde{C}}$$

From here on, the arguments work word-by-word again and thus we obtain that the arguments as well as the formulas until [6, (A.32)] remain all true when implementing these modifications.

Ultimately, the final estimate for the variance is the sum of the rhs of [6, (A.19),(A.20), (A.29) and (A.32)] with the addition of the prefactor C at places where one sees κ_1/v_1^2 and modifying v_{k-1} in numerators to v_1 and $1/(v_{k-1} - 1)$ to $1/(v_1 - \widetilde{C})$. Using that $d_{v_q^\star} \geq u_{i_{\star(q)}}^{(q)}$, we thus obtain that (ignoring the 'new' superscript everywhere for brevity now):

$$\mathrm{Var}_Y[N_{z_n^\star}(v_r^\star, v_b^\star)] \leq \mathbb{E}_Y[N_{z_n^\star}(v_r^\star, v_b^\star)]$$

$$+ \overline{\mathbb{E}_Y[N_{z_n^\star}(v_r^\star, v_b^\star)]}^2 \left(\frac{v_1}{v_1 - \widetilde{C}} \frac{C\kappa_1}{v_1^2} \left(\frac{1}{u_{i_{\star(r)}}^{(r)}} + \frac{1}{u_{i_{\star(b)}}^{(b)}} \right) + \frac{v_1^2}{(v_1 - \widetilde{C})^2} \frac{C^2 \kappa_1^2}{v_1^4} \frac{1}{u_{i_{\star(r)}}^{(r)} u_{i_{\star(b)}}^{(b)}} \right.$$

$$+ \frac{8(z_n^\star)^2}{\ell_n} + \left(1 + \frac{C\kappa_1 v_1}{v_1^2 u_{i_{\star(r)}}^{(r)}} \right) \left(1 + \frac{C\kappa_1 v_1}{v_1^2 u_{i_{\star(b)}}^{(b)}} \right) \frac{z^\star}{v_1 - C} \left(2 \frac{(z_n^\star)^2 v_1}{\ell_n} \frac{C^2 \kappa_1^2}{v_1^4} \right) \right) \tag{4.13}$$

where $\overline{\mathbb{E}_Y[N_{z_n^\star}(v_r^\star, v_b^\star)]}$ stands for the upper bound on $\mathbb{E}_Y[N_{z_n^\star}(v_r^\star, v_b^\star)]$ in (4.3). Recall that $\overline{\mathbb{E}_Y[N_{z_n^\star}(v_r^\star, v_b^\star)]}$ and $\mathbb{E}_Y[N_{z_n^\star}(v_r^\star, v_b^\star)]$ are all given by (4.11) up to smaller order correction terms. So, in order to show that the rhs of (4.12) tends to zero, it is enough to analyse the factor multiplying $\overline{\mathbb{E}_Y[N_{z_n^\star}(v_r^\star, v_b^\star)]}^2$ in (4.13) and show that it tends to zero as $n \to \infty$. We carry out this now.

By the same method as the one in Claim 2.6, (i.e., using Karamata's theorem) it is not hard to show that

$$\mathbb{E}[D_n^\star(D_n^\star - 1)] = \sum_{\substack{v \in [n] \\ d_v \leq n^{\beta_n}}} \frac{d_v(d_v - 1)(d_v - 2)}{\ell_n} \leq n^{\beta_n(4-\tau)} C L_n^{\star,\mathrm{up}}(n^{\beta_n}) \lesssim n^{\beta_n(4-\tau)}.$$

(4.14)

This together with (4.9) implies that $\kappa_1 \sim n^{\beta_n(4-\tau)}$. From (4.2) and (4.8), $\nu_1 \sim n^{\beta_n(3-\tau)}$, and finally from (2.37), (see also (2.3)) $u_{i_{\star(q)}}^{(q)} \sim n^{\beta_n(\tau-2)^{b_n^{(q)}(\beta_n)}}$. This implies that for $q \in \{r, b\}$

$$\frac{\kappa_1}{\nu_1^2} \frac{1}{u_{i_{\star(q)}}^{(q)}} \sim n^{\beta_n((\tau-2)-(\tau-2)^{b_n^{(q)}(\beta_n)})} \to 0,$$

(4.15)

as $n \to \infty$, since $b_n^{(b)}(\beta_n), b_n^{(r)}(\beta_n) < 1$, and the error terms hidden in the $\sim$ sign are at most $\exp\{\pm(\log n^{\beta_n})^\theta\}$ for $\theta < 1$, and are thus of smaller order of magnitude. Thus, both the first and second term multiplying $\overline{\mathbb{E}_Y[N_{z^\star}(v_r^\star, v_b^\star)]}$ in (4.13) tends to zero as $n \to \infty$. Note that for (4.16) to tend to zero it is crucial that $d_{v_q^\star} > n^{\beta_n(\tau-2)}$, i.e., that $v_q^\star$ is a hub. When distributing the product in the second line of (4.13), using that $\nu_1/(\nu_1 - \widetilde{C})$ is a constant factor, we see that the main contribution comes from the terms

$$\frac{\kappa_1}{\nu_1 u_{i_{\star(q)}}^{(q)}} \cdot \frac{1}{\ell_n} \frac{\kappa_1^2}{\nu_1^4} \sim n^{\beta_n(1-(\tau-2)^{b_n^{(q)}(\beta_n)})} \cdot n^{-1} \cdot n^{2\beta_n(\tau-2)} = n^{\beta_n(\tau-1+\tau-2-(\tau-2)^{b_n^{(q)}(\beta_n)})-1}.$$

(4.16)

Note that since $b_n^{(q)}(\beta_n) \in [0, 1)$, $(\tau - 2) - (\tau - 2)^{b_n^{(q)}} \leq 0$, and thus the exponent is always negative when $\beta_n < 1/(\tau - 1)$. When $\beta_n = 1/(\tau - 1)$, the exponent is always nonpositive and equals 0 if and only if $b_n^{(q)}(\beta_n) = 0$. This is only possible if $Y_q^{(n)}$ takes values in a measure 0 discrete set, see (1.15). The probability of this event tends to 0 under Assumption 1.7.

Combining the estimates in (4.15) and (4.16), we obtain that the variance of $N_{z_n^\star}(v_r^\star, v_b^\star)$ is of smaller order than its expectation squared, hence the rhs of (4.12) tends to zero. This establishes that whp there is a path of length $z_n^\star + 1$ connecting $v_r^\star$ to $v_b^\star$. Thus, we obtain the existence of a path of length as in (1.16). This proves the upper bound on $d_G(v_r, v_b)$ and thus completes the proof of Theorem 1.8. $\qquad\square$

5 Extensions and By-products

In this section prove Theorems 1.13 and 1.15 and sketch the proof of Observation 1.14.

Proof of Theorem 1.13 The proof of the first statement of the theorem, that is, (1.21) follows from the proof of Theorem 1.8 in Section 4. Recall that we write $\Lambda_{\leq z} = \{w \in [n] : d_w \leq n^z\}$. Note that in this case, Claim 2.6 yields that the empirical second moment restricted to degrees at most $n^{\widetilde{\beta}_n}$ is $\lesssim n^{\widetilde{\beta}_n(3-\tau)}$. The proof of Theorem 1.13 is even simpler than the one in Section 4: here, one does not need to estimate the number of half-edges attached to v_1, v_2, since these are given, so there is no need to condition on the sigma algebra $\mathcal{F}_{\varrho_n'}$ or on the good event (NoBad) either. Thus, below in this proof, we can use the 'usual' probability measure $\mathbb{P}$ instead of $\mathbb{P}_Y$. Let us write $\mathcal{P}_{v_1,v_2}(z)$ for a path of length z connecting v_1, v_2. Counting paths restricted to the set $\Lambda_{\widetilde{\beta}_n}$, the restriction in the sum in the middle factor in (4.1) is that

$\pi_i \in \Lambda_{\leq \widetilde{\beta}_n}$, thus, (4.3) turns into

$$\mathbb{E}\left[\# \left\{ \mathcal{P}_{v_1, v_2}(z) \in \Lambda_{\leq \widetilde{\beta}_n} \right\} \right] \lesssim n^{-1} n^{x_1 \widetilde{\beta}_n + x_2 \widetilde{\beta}_n} n^{z \widetilde{\beta}_n (3 - \tau)} \tag{5.1}$$

and (4.4) becomes

$$\mathbb{P}\left(\exists \mathcal{P}_{v_1, v_2}(z) \in \Lambda_{\leq \widetilde{\beta}_n} : |\mathcal{P}_{v_1, v_2}| \leq z + 1 \right) \lesssim n^{-1} n^{\widetilde{\beta}_n (x_1 + x_2 + z(3 - \tau))}. \tag{5.2}$$

From here, the proof of the upper bound in Section 4 can be repeated with β_n replaced by $\widetilde{\beta}_n$, $(\tau - 2)^{b_n^{(q)}(\beta_n)}$ replaced by x_j, yielding (1.21). The proof of the lower bound in Sect. 4 is again valid word-by-word. It is important to note that the second moment method in (4.12) works only if the variance is of smaller order than the expectation squared, which, in turn, is equivalent to the quantities in (4.15) and (4.16) tending to zero, which is the case whenever $x_j > \tau - 2$. (Thus, for non-hub vertices the method does not work.) Next we show (1.22) and (1.24). Note that (1.24) is equivalent to

$$\mathbb{P}\left(\left\{ \# \mathcal{P}_{v_1, v_2}^\star \in \Lambda_{\leq \widetilde{\beta}_n} \right\} \notin \left[n^{\widetilde{\beta}_n f^u(\widetilde{\beta}_n, x_1, x_2)(1 - \varepsilon)}, n^{\widetilde{\beta}_n f^u(\widetilde{\beta}_n, x_1, x_2)(1 + \varepsilon)} \right] \right) \to 0. \tag{5.3}$$

We prove this statement using Chebyshev's inequality. Let us shortly write

$$\zeta_n := \left\{ \# \mathcal{P}_{v_1, v_2}^\star \in \Lambda_{\leq \widetilde{\beta}_n} \right\}. \tag{5.4}$$

First, when we write $\underline{\mathbb{E}[\zeta_n]}$, $\overline{\mathbb{E}[\zeta_n]}$ for the lower and upper bounds on $\mathbb{E}[\zeta_n]$, respectively, then these are handled in (4.11), and equal $\sim n^{\widetilde{\beta}_n f^u(\widetilde{\beta}_n, x_1, x_2)}$ by elementary calculations using the value $z_n^\star$ from (4.5) and f^u from (1.19).

Next, $\mathrm{Var}[\zeta_n]$ is handled in (4.13), where consecutively in (4.15) and (4.16) it is established that

$$\mathrm{Var}[\zeta_n] \leq C \overline{\mathbb{E}[\zeta_n]}^2 \max_{j=1,2} \max \left\{ \frac{\kappa_1}{v_1^2 d_{v_j}}, \frac{\kappa_1^3}{v_1^5 \ell_n d_{v_j}} \right\}. \tag{5.5}$$

Comparing the rhs of (4.15) to the rhs of (4.16) with β_n replaced by $\widetilde{\beta}_n$, it is elementary to check that the dominating expression is the rhs of (4.15) unless $\widetilde{\beta}_n = 1/(\tau - 1)$, in which case both terms are of the same order. Nevertheless, we arrive to

$$\max_{j=1,2} \max \left\{ \frac{\kappa_1}{v_1^2 d_{v_j}}, \frac{\kappa_1^3}{v_1^5 \ell_n d_{v_j}} \right\} \sim n^{-\widetilde{\beta}_n (\min(x_1, x_2) - (\tau - 2))}. \tag{5.6}$$

Note that this is the point where it becomes clear why the assumption $x_1, x_2 > \tau - 2$ was necessary: only in this case can we expect any concentration of the variable ζ_n. Chebyshev's inequality yields that for any $c(n, \varepsilon)$ that depends on n and some $\varepsilon > 0$ to be chosen later

$$\mathbb{P}\left(|\zeta_n - \mathbb{E}[\zeta_n]| \geq \mathbb{E}[\zeta_n] c(n, \varepsilon) \right) \leq \frac{\mathrm{Var}(\zeta_n)}{\mathbb{E}[\zeta_n]^2 c(n, \varepsilon)^2} \lesssim \frac{\overline{\mathbb{E}[\zeta_n]}^2}{\underline{\mathbb{E}[\zeta_n]}^2} \frac{1}{c(n, \varepsilon)^2} n^{-\widetilde{\beta}_n (\min(x_1, x_2) - (\tau - 2))}$$

$$\sim \frac{1}{\widetilde{\beta}_n^2} \frac{1}{c(n, \varepsilon)^2} n^{-\widetilde{\beta}_n (\min(x_1, x_2) - (\tau - 2))}, \tag{5.7}$$

where we have used that $\overline{\mathbb{E}[\zeta_n]}/\underline{\mathbb{E}[\zeta_n]}$ is $1/\widetilde{\beta}_n$ times a factor that can be merged in the $\sim$ sign. By the assumption that $\beta_n (\log n)^\gamma \to \infty$, for some $\gamma < 1$, setting

$$c(n, \varepsilon) := n^{-(1 - \varepsilon) \widetilde{\beta}_n (\min(x_1, x_2) - (\tau - 2))/2} \to 0,$$

the rhs tends to zero for any fix $\varepsilon > 0$. Note that we only have upper and lower bounds on the expected value, thus we obtain that

$$\mathbb{P}(\zeta_n \notin [(1 - c(n, \varepsilon))\underline{\mathbb{E}[\zeta_n]}, (1 + c(n, \varepsilon))\overline{\mathbb{E}[\zeta_n]}]) \to 0. \tag{5.8}$$

The interval in (5.3) is certainly wider than the one excluded here. This finishes the proof of (1.24). Further note that both the lower as well as the upper ends of the interval in (5.8) are still $\sim n^{\widetilde{\beta}_n f^u(\widetilde{\beta}_n, x_1, x_2)}$. This finishes the proof of (1.22).

We turn to the proof of (1.23), which is essentially Markov's inequality. Indeed, when we require for some $\delta < 1$ that one of the vertices on a path must fall in $\Lambda_{\leq \delta \widetilde{\beta}_n}$ (but not its location), we obtain that the expected number of such paths connecting v_1, v_2 is at most

$$\mathbb{E}\left[\# \left\{ \mathcal{P}_{v_1, v_2}(z) : \exists i \leq z, \pi_i \in \Lambda_{\leq \delta \widetilde{\beta}_n} \right\} \right] \lesssim z n^{-1} n^{\widetilde{\beta}_n(x_1 + x_2 + (z-1)(3-\tau) + \delta(3-\tau))}. \tag{5.9}$$

Let us denote by $z_n^\star(\delta)$ the smallest z for which this quantity does not tend to zero, which is exactly

$$z_n^\star(\delta) = \left\lceil \frac{1/\widetilde{\beta}_n - x_1 - x_2}{3 - \tau} + (1 - \delta) \right\rceil$$

On the other hand, without any restriction, the shortest path uses

$$z_n^\star(1) = \left\lceil \frac{1/\widetilde{\beta}_n - x_1 - x_2}{3 - \tau} \right\rceil$$

many in-between vertices, by (1.21). As long as δ is such that $z^\star(\delta) > z^\star(1)$, the rhs of (5.9) tends to 0 for $z = z^\star(1)$. Thus, there will be no connecting paths of length $z^\star(1)$ that have a vertex in $\Lambda_{\leq \delta \widetilde{\beta}_n}$. Finally, the largest δ that we can achieve, $\delta_{\max} := \sup\{\delta : z^\star(\delta) > z^\star(1)\}$ is precisely the lower fractional part of the expression within the upper-integer-part in $z^\star(1)$, in other words, $\delta_{\max} = f^\ell(\widetilde{\beta}_n, x_1, x_2)$, establishing (1.23). □

Sketch proof of Observation 1.14. From the proof of Theorem 1.13 it follows that all shortest paths use degree at least $\sim n^{\widetilde{\beta}_n f^\ell(\widetilde{\beta}_n, x_1, x_2)}$ while the Chebyshev's inequality in (5.7) shows that the number of shortest paths is $\sim n^{\widetilde{\beta}_n f^u(\widetilde{\beta}_n, x_1, x_2)}$. Multiplying these two together yields one part of the observation. The full statement is finished when we show that there is at least one shortest path that actually uses a vertex with degree $\sim n^{\widetilde{\beta}_n f^\ell(\widetilde{\beta}_n, x_1, x_2)(1+\varepsilon)}$ for arbitrary small $\varepsilon > 0$. This can be done using Chebyshev's inequality again in the same way as for ζ_n in (5.4), now counting paths with at least one vertex in $\Lambda_{\leq \widetilde{\beta}_n f^\ell(\widetilde{\beta}_n, x_1, x_2)(1+\varepsilon)}$ and the rest of the vertices in $\Lambda_{\leq \widetilde{\beta}_n}$. In this case, the variance vs expectation squared method carries through the same way. □

Proof of Theorem 1.15 The proof of this theorem is essentially the upper bound – the construction of the connecting path – of the proof of Theorem 1.8. For the first statement, about the segments with increasing degrees, Proposition 2.1 shows that there is a path of length $T_q(\beta_n)$ that connects v_q to a vertex $v_q^\star$ with degree as in (1.25), while Proposition 3.1 ensures that the total degree in the exploration clusters of depth $T_q(\beta_n)$ is the same order of magnitude as the degree of $v_q^\star$, and that the two exploration clusters are disjoint.

By Lemma 2.5, the $(T_q(\beta_n) - k)$th vertex on this path has degree at least $u_{i_{\star(q)} - k}^{(q)}$ and at most $\widehat{u}_{i_{\star(q)} - k}^{(q)}$, both of them $\sim n^{\beta_n(\tau - 2)^{b_n^{(q)}(\beta_n) + k}}$.

Thus, when connecting the two clusters $\mathcal{C}_{T_r(\beta_n)}^{(r)}$ and $\mathcal{C}_{T_b(\beta_n)}^{(b)}$, Theorem 1.13 can be applied to show that there is a path that connects $v_r^\star, v_b^\star$ of length as in (1.26), using vertices of degree

at least $n^{\beta_n f_n^\ell}$. By the bound on the total degrees in Proposition 3.1, we can identify all the vertices in $\mathcal{C}^{(q)}_{T_q(\beta_n)}$, and apply Theorem 1.13 once more to see that no shorter path is possible between $\mathcal{C}^{(r)}_{T_r(\beta_n)}$ and $\mathcal{C}^{(b)}_{T_b(\beta_n)}$ than the one with length (1.26), and any of these shortest paths uses vertices with degrees at least $n^{\beta_n f_n^\ell}$. □

6 Comparison of Typical Distances

In this section we compare the result of Theorem 1.8 applied to the $\beta_n = 1/(\tau - 1)$ setting to [31, Theorem 1.2] in more detail. Here we argue that the two formulations - namely the one in [31, Theorem 1.2] and the one in (1.28) - are indeed the same, by describing the core idea of the proof of [31, Theorem 1.2], and relate quantities (events, random variables, etc.) appearing in that proof to quantities in this paper. In this section $\approx$ means equality up to a $(1 + o_{\mathbb{P}}(1))$ factor.

The proof of [31, Theorem 1.2] goes through a minimisation problem, where the two BFS clusters of v_r and v_b should connect the first time such that a coupling (to two branching processes) should be maintained. More precisely, suppose k_1 is a random variable that is measurable w.r.t. $\{\mathcal{C}^{(q)}_s\}^m_{s=1}$, for some m (in this paper we take $m = t(n^{\varrho'_n})$). Suppose we run the BFS started from v_r for k_1 steps, and from v_b for $k - k_1 - 1$ steps. There are $Z^{(r)}_{k_1+1}$, $Z^{(b)}_{k-k_1}$ many half edges attached to the vertices in the two clusters, respectively. The distance between v_r, v_b is then larger than k if these sets of half-edges do not connect to each other, and the probability of this event is approximately

$$\mathbb{P}\left(H\left(\mathcal{C}^{(r)}_{k_1}\right) \cap H\left(\mathcal{C}^{(b)}_{k-k_1-1}\right) = \varnothing \mid Z^{(r)}_{k_1+1}, Z^{(b)}_{k-k_1}\right) \approx \exp\left\{-cZ^{(r)}_{k_1+1}Z^{(b)}_{k-k_1}/\ell_n\right\} \quad (6.1)$$

A branching process approximation similar to the one in Section 2.1 is performed to approximate the numerator in the exponent. However, this BP approximation is only valid until none of the colors have more half-edges than $n^{(1-\varepsilon)/(\tau-1)}$ for some small $\varepsilon > 0$, i.e., they do not reach the highest-degree vertices in the graph yet. This criterion is established in [31, Proposition 3.2]. The set $\mathcal{T}^{i,n}_m$ in [31, (3.3)] exactly describes those values of ℓ for which $\{\ell \le T_q + 1\}$, $q \in \{r, b\}$ holds (where $T_q = T_q(1/(\tau - 1))$, defined in (1.15), or in (2.38), is the time to reach the hubs). The $+1$ is added to T_q since the half-edges attached to vertices in the BFS cluster at time T_q can be described as the next generation, they have size $Z^{(q)}_{T_q+1}$.

Now, from (6.1), we see that $\{d_G(v_r, v_b) > k\}$ happens whp if $Z^{(r)}_{k_1+1} \cdot Z^{(b)}_{k-k_1} = o(n)$ and also that both $k_1 + 1 \in \mathcal{T}^{r,n}_m$ and $k - k_1 \in \mathcal{T}^{b,n}_m$ holds. These latter conditions are described in the set of indices $\mathcal{B}_n$ in [31, (4.57)]:

$$\mathcal{B}_n := \{k_1 \in \mathbb{N} : k_1 + 1 \le T_r + 1, k - k_1 \le T_b + 1\}.$$

Using the BP approximation similar as in (1.11), and the rhs of (6.1),

$$\mathbb{P}(d_G(v_r, v_b) > k)$$
$$\approx \max_{k_1 \in \mathcal{B}_n} \exp\left\{-C \exp\left\{(\tau - 2)^{-(k_1+1)}Y^{(n)}_r + (\tau - 2)^{-(k-k_1)}Y^{(n)}_b - \log n\right\}\right\}.$$

With the event

$$\mathcal{E}_{n,k}(\delta) := \left\{\exists k_1 \in \mathcal{B}_n, (\tau - 2)^{-(k_1+1)}Y^{(n)}_r + (\tau - 2)^{-(k-k_1)}Y^{(n)}_b < (1 - \delta)\log n\right\}$$

it is obvious from (6.1) that for any $\delta > 0$, $\lim_{n \to \infty} \mathbb{P}(d_G(v_r, v_b) > k \mid \mathcal{E}_{n,k}(\delta)) \to 1$, while $\mathbb{P}(d_G(v_r, v_b) > k \mid \mathcal{E}_{n,k}^c(\delta)) \to 0$. Hence, we get that

$$\mathbb{P}_Y(d_G(v_r, v_b) > k) \approx \mathbb{P}\left(\min_{k_1 \in \mathcal{B}_n}(\tau - 2)^{-(k_1+1)}Y_r^{(n)} + (\tau - 2)^{-(k-k_1)}Y_b^{(n)} < \log n\right). \tag{6.2}$$

The paper shows that $\min_{k_1 \in \mathcal{B}_n}$ can be replaced by $\min_{k_1 \leq k}$ in the minimum in (6.2) above.

Next we show that the formulation of (6.2) gives the same distances as our statement for typical distances in Theorem 1.8, that is,

$$d_G(v_r, v_b) = T_r + T_b + 2 - \mathbb{1}\{(\tau - 2)^{b_n^{(b)}} + (\tau - 2)^{b_n^{(r)}} > \tau - 1\}. \tag{6.3}$$

To be able to show that the two formulation are the same, we use (6.2) to show that

(a) $\mathbb{P}(d_G(v_r, v_b) > T_r + T_b) \to 1$,
(b) $\mathbb{P}(d_G(v_r, v_b) > T_r + T_b + 1) \to \mathbb{P}(\mathbb{1}\{(\tau - 2)^{b_n^{(b)}} + (\tau - 2)^{b_n^{(r)}} > \tau - 1\} = 0)$ and finally that
(c) $\mathbb{P}(d_G(v_r, v_b) > T_r + T_b + 2) \to 0$.

From (2.38) and for $\beta_n = 1/(\tau - 1)$, it is an elementary calculation to check that that for any $i \in \mathbb{Z}$, $q \in \{r, b\}$

$$(\tau - 2)^{-(T_q+1+i)}Y_q^{(n)} = \log n \frac{(\tau - 2)^{b_n^{(r)}-i}}{\tau - 1}. \tag{6.4}$$

First, we check that (6.2) gives (a). For this we set $k = T_r + T_b$, and let us write $k_1 := T_r - \ell$ for some $\ell \in \mathbb{Z}$, then $k - k_1 = T_b + \ell$. Hence, we can rewrite (6.2) using (6.4) with $i = -\ell$ and $i = \ell - 1$, and get

$$\mathbb{P}_Y(d_G(v_r, v_b) > T_r + T_b) \approx \mathbb{P}\left(\min_\ell \frac{(\tau - 2)^{b_n^{(r)}+\ell} + (\tau - 2)^{b_n^{(b)}+1-\ell}}{\tau - 1} < 1\right).$$

It is clear now that setting $\ell = 0$ shows that the inequality is satisfied for all $b_n^{(r)}, b_n^{(b)} \in [0, 1)$, since the expression after the $\min_\ell$, for $\ell = 0$ is at most $(1 + (\tau - 2))/(\tau - 1) = 1$. Moreover, note that for $\ell = 0$, $k_1 = T_r$ and $k - k_r = T_b$, so both $k_1 + 1 \leq T_r + 1$ and $k - k_1 \leq T_b + 1$ hold, hence, we found an index k_1 in $\mathcal{B}_n$.

Next, we check that (6.2) gives (b). For this, we set $k = T_r + T_b + 1$, again write $k_1 = T_r - \ell$, so that $k - k_1 = T_b + 1 + \ell$. Hence, we can rewrite (6.2) using (6.4) with $i = -\ell$ and $i = \ell$ and get

$$\mathbb{P}_Y(d_G(v_r, v_b) > T_r + T_b + 1) \approx \mathbb{P}\left(\min_\ell \frac{(\tau - 2)^{b_n^{(r)}+\ell} + (\tau - 2)^{b_n^{(b)}-\ell}}{\tau - 1} < 1\right).$$

It is clear now that setting $\ell = 0$ yields (b). Moreover, note that for $\ell = 0$ $k_1 \in \mathcal{B}_n$ holds as well. We argue that $\ell = 0$ is indeed the minimizer of the expression after the min. Wlog we can assume that $\ell \geq 1$, the case when $\ell \leq -1$ can be treated similarly. Then, we need to show that for all $\ell \geq 1$,

$$(\tau - 2)^{b_n^{(r)}+\ell} + (\tau - 2)^{b_n^{(b)}-\ell} > (\tau - 2)^{b_n^{(r)}} + (\tau - 2)^{b_n^{(b)}}. \tag{6.5}$$

Rearranging this inequality yields

$$(\tau - 2)^{b_n^{(b)}}\left((\tau - 2)^{-\ell} - 1\right) > (\tau - 2)^{b_n^{(r)}}(1 - (\tau - 2)^\ell)$$

Since for all $\ell > 1$,

$$(\tau - 2)^{b_n^{(b)}} \left((\tau - 2)^{-\ell} - 1 \right)$$
$$> (\tau - 2)^{-\ell+1} - (\tau - 2) \geq 1 - (\tau - 2)^\ell \geq (\tau - 2)^{b_n^{(r)}} \left(1 - (\tau - 2)^\ell \right),$$

the claim is established.

Finally, we check (c). For this, we set $k = T_r + T_b + 2$, again write $k_1 = T_r - \ell$, so that $k - k_1 = T_b + 2 + \ell$. Hence, we can rewrite (6.2) using (6.4) with $i = -\ell$ and $i = \ell + 1$ and get

$$\mathbb{P}_Y(d_G(v_r, v_b) > T_r + T_b + 2) \approx \mathbb{P} \left(\min_\ell \frac{(\tau - 2)^{b_n^{(r)}+\ell} + (\tau - 2)^{b_n^{(b)}-\ell-1}}{\tau - 1} < 1 \right). \tag{6.6}$$

We need to show that no $\ell \in \mathbb{Z}$ satisfies this minimisation problem and thus the probability tends to 0. For this, we use again that $b_n^{(r)}, b_n^{(b)} \in [0, 1)$ implies that

$$(\tau - 2)^{b_n^{(r)}+\ell} + (\tau - 2)^{b_n^{(b)}-\ell-1} > (\tau - 2)^{\ell+1} + (\tau - 2)^{-\ell},$$

and it is elementary to show again that the rhs is at least $\tau - 1$ for all $\ell \in \mathbb{Z}$ and $\tau \in (2, 3)$, thus, the inequality on the rhs of (6.6) cannot be satisfied.

These calculations show that the statement of Theorem 1.8 yields - through a non-trivial rewrite - the statement of [31, Theorem 1.2]. The final formula of [31, Theorem 1.2], i.e., the distribution of the fluctuation of the typical distance around $2 \log \log n / |\log(\tau - 2)|$ is then obtained by solving analytically the minimisation problem on the rhs of (6.2) with $k_1 \in \mathcal{C}_n^{(b)}$ replaced by $k_1 \leq k$.

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. Albert, R., Barabási, A.-L.: Statistical mechanics of complex networks. Rev. Mod. Phys. **74**(1), 47–97 (2002)
3. Albert, R., Jeong, H., Barabási, A.-L.: Error and attack tolerance of complex networks. Nature **406**(6794), 378–382 (2000)
4. Amaral, L.A.N., Scala, A., Barthélémy, M., Stanley, H.E.: Classes of small-world networks. Proc. Natl. Acad. Sci. USA **97**, 11149–11152 (2000)
5. Barabási, A.-L., Albert, R.: Emergence of scaling in random networks. Science **286**(5439), 509–512 (1999)
6. Baroni, E., Hofstad, R.v d, Komjáthy., J.: Fixed speed competition on the configuration model with infinite variance degrees: unequal speeds. Electron. J. Probab. **20**(116), 1–48 (2015)
7. Bhamidi, S., Hofstad, R.v d, Hooghiemstra, G.: First passage percolation on random graphs with finite mean degrees. Ann. Appl. Probab. **20**(5), 1907–1965 (2010)

8. Bhamidi, S., Hofstad, R. v. d., Hooghiemstra, G.: Universality for first passage percolation on sparse random graphs. to appear. Ann. Probab. arXiv:1210.6839 (2012)

9. Bingham, N.H., Goldie, C.M., Teugels, J.L.: Regular Variation, vol. 27. Cambridge University Press, Cambridge (1989)

10. Broder, A., Kumar, R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., Tomkins, A., Wiener, J.: Graph structure in the web. Comput. Netw. **33**(1), 309–320 (2000)

11. Chung, F., Lu, L.: The diameter of sparse random graphs. Adv. Appl. Math. **26**(4), 257–279 (2001)

12. Chung, F., Lu, L.: The average distance in a random graph with given expected degrees. Internet Math. **1**(1), 91–113 (2004)

13. Clauset, A., Shalizi, C., Newman, M.E.J.: Power-law distributions in empirical data. SIAM Rev. **51**(4), 661–703 (2009)

14. Cohen, R., Erez, K., ben Avraham, D., Havlin, S.: Breakdown of the internet under intentional attack. Phys. Rev. Lett. **86**, 3682–3685 (2001)

15. Cohen, R., Havlin, S.: Scale-free networks are ultrasmall. Phys. Rev. Lett. **90**, 058701 (2003)

16. Davies, P.L.: The simple branching process: a note on convergence when the mean is infinite. J. Appl. Probab. **15**(3), 466–480 (1978)

17. Dodds, P., Muhamad, R., Watts, D.: An experimental study of search in global social networks. Science **301**(5634), 827–829 (2003)

18. Dommers, S., Hofstad, R v d, Hooghiemstra, G.: Diameters in preferential attachment graphs. J. Stat. Phys **139**(72), 107 (2010)

19. Dorogovtsev, S.N., Mendes, J.F.F., Samukhin, A.N.: Metric structure of random networks. Nucl. Phys. B **653**(3), 307–338 (2003)

20. Eckhoff, M., Mörters, P., et al.: Vulnerability of robust preferential attachment networks. Electron. J. Probab. **19** (2014)

21. Faloutsos, M., Faloutsos, P., Faloutsos, C.: On power-law relationships of the internet topology. ACM SIGCOMM Comput. Commun. Rev. **29**(4), 251–262 (1999)

22. Federico, L., van der Hofstad, R.: Critical window for connectivity in the configuration model. Comb. Probab. Comput. pp. 1–21 (2017)

23. Fernholz, D., Ramachandran, V.: The diameter of sparse random graphs. Random Struct. Algorithms **31**(4), 482–516 (2007)

24. Fronczak, A., Fronczak, P., Hołyst, P.: How to calculate the main characteristics of random graphs-a new approach. arXiv:cond-mat/0308629 [cond-mat.stat-mech] (2003)

25. Fronczak, A., Fronczak, P., Hołyst, J.A.: Average path length in random networks. Phys. Rev. E **70**, 056110 (2004)

26. Gao, P., Wormald, N.: Enumeration of graphs with a heavy-tailed degree sequence. Adv. Math. **287**, 412–450 (2016)

27. Guimera, R., Mossa, S., Turtschi, A., Amaral, L.N.: The worldwide air transportation network: anomalous centrality, community structure, and cities' global roles. Proc. Natl. Acad. Sci. USA **102**(22), 7794–7799 (2005)

28. Hofstad, R.V.D.: Random Graphs and Complex Networks, vol. 1. Cambridge University Press, Cambridge (2016)

29. van der Hofstad, R.: Random Graphs and Complex Networks, vol. 2. Cambridge University Press, Cambridge (2016)

30. 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)

31. 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)

32. Holme, P., Kim, B.J., Yoon, C.N., Han, S.K.: Attack vulnerability of complex networks. Phys. Rev. E **65**, 056109 (2002)

33. Hołyst, J.A., Sienkiewicz, J., Fronczak, A., Fronczak, P., Suchecki, K.: Universal scaling of distances in complex networks. Phys. Rev. E **72**, 026108 (2005)

34. Janson, S., Luczak, M.J.: A new approach to the giant component problem. Random Struct. Algorithms **34**(2), 197–216 (2009)

35. Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N., Barabási, A.-L.: The large-scale organization of metabolic networks. Nature **407**(6804), 651–654 (2000)

36. Jordano, P., Bascompte, J., Olesen, J.M.: Invariant properties in coevolutionary networks of plantanimal interactions. Ecol. Lett. **6**(1), 69–81 (2003)

37. Kumar, R., Raghavan, P., Rajagopalan, S., Tomkins, A.: Trawling the web for emerging cyber-communities. Comput. Netw. **31**(1116), 1481–1493 (1999)

38. Levin, D.A., Peres, Y., Wilmer, E.L.: Markov Chains and Mixing Times. American Mathematical Society, Providence (2009)
39. Liljeros, F., Edling, C.R., Amaral, L.A.N., Stanley, H.E.: The web of human sexual contacts. Nature **411**, 907 (2001)
40. McKay, B.D., Wormald, N.C.: Asymptotic enumeration by degree sequence of graphs with degreeso(n1/2). Combinatorica **11**(4), 369–382 (1991)
41. Milgram, S.: The small world problem. Psychol Today **1**, 60–67 (1967)
42. Molloy, M., Reed, B.: A critical point for random graphs with a given degree sequence. Random Struct. Algorithms **6**(2–3), 161–180 (1995)
43. Molloy, M., Reed, B.: The size of the giant component of a random graph with a given degree sequence. Comb. Probab. Comput. **7**(3), 295–305 (1998)
44. Montoya, J.M., Pimm, S.L., Solé, R.V.: Ecological networks and their fragility. Nature **442**(7100), 259–264 (2006)
45. Mossa, S., Barthélémy, M., Stanley, H.E., Amaral, L.A.: Truncation of power law behavior in scale-free network models due to information filtering. Phys. Rev. Lett. **88**, 138701 (2002)
46. Nanavati, A.A., Gurumurthy, S., Das, G., Chakraborty, D., Dasgupta, K., Mukherjea, S., Joshi, A.: On the structural properties of massive telecom call graphs: findings and implications. In: Proceedings of the 15th ACM International Conference on Information and Knowledge Management, CIKM '06, pp. 435–444, New York, NY, USA . ACM (2006)
47. Newman, M .E .J.: Scientific collaboration networks. I and II. Phys. Rev. E **64**(1), 016131–016132 (2001)
48. Newman, M.E.J.: The structure of scientific collaboration networks. Proc. Natl. Acad. Sci. USA **98**(2), 404–409 (2001)
49. M. E. J, Newman: The structure and function of complex networks. SIAM Rev. **45**(2), 167–256 (2003). (electronic)
50. Newman, M.E.J.: Power laws, Pareto distributions and Zipf's law. Contemp. Phys. **46**(5), 323–351 (2005)
51. Newman, M.E.J., Strogatz, S., Watts, D.: Random graphs with arbitrary degree distribution and their application. Phys. Rev. E **64**(026118), 1–17 (2000)
52. Norros, I., Reittu, H.: On a conditionally Poissonian graph process. Adv. in Appl. Probab. **38**(1), 59–75 (2006)
53. Norros, I., Reittu, H.: Network models with a 'soft hierarchy': a random graph construction with loglog scalability. IEEE Netw. **22**(2) (2008)
54. Ostilli, M.: Fluctuation analysis in complex networks modeled by hidden-variable models: necessity of a large cutoff in hidden-variable models. Phys. Rev. E **89**, 022807 (2014)
55. Pool, I d S, Kochen, M.: Contacts and influence. Soc. Netw. **1**(5), 51 (1978)
56. Seneta, E.: The simple branching process with infinite mean. J. Appl. Probab. **10**(1), 206–212 (1973)
57. Seneta, E.: Regularly varying functions in the theory of simple branching processes. Adv. Appl. Probab. **6**(3), 408–420 (1974)
58. Strogatz, S.H.: Exploring complex networks. Nature **410**(6825), 268–276 (2001)
59. Travers, J., Milgram, S.: An experimental study of the small world problem. Sociometry **32**, 425–443 (1969)
60. Ugander, J., Karrer, B., Backstrom, L., Marlow, C.: The anatomy of the Facebook social graph. arXiv:1111.4503 (2011)
61. Watts, D.J.: Small Worlds: The Dynamics of Networks Between Order and Randomness Princeton Studies in Complexity. Princeton University Press, Princeton (1999)
62. Watts, D.J., Strogatz, S.H.: Collective dynamics of 'small-world' networks. Nature **393**, 440–442 (1998)