CONNECTIVITY OF RANDOM GRAPHS AFTER CENTRALITY-BASED VERTEX REMOVAL

REMCO VAN DER HOFSTAD [ID],* ** AND
MANISH PANDEY [ID],* *** *Eindhoven University of Technology*

Abstract

Centrality measures aim to indicate who is important in a network. Various notions of 'being important' give rise to different centrality measures. In this paper, we study how important the central vertices are for the *connectivity structure* of the network, by investigating how the removal of the most central vertices affects the number of connected components and the size of the giant component. We use *local convergence techniques* to identify the limiting number of connected components for locally converging graphs and centrality measures that depend on the vertex's neighbourhood. For the size of the giant, we prove a general upper bound. For the matching lower bound, we specialise to the case of *degree centrality* on one of the most popular models in network science, the *configuration model*, for which we show that removal of the highest-degree vertices destroys the giant most.

Keywords: Strictly local centrality measures; centrality-based vertex removal; number of connected components; size of giant; configuration model

2020 Mathematics Subject Classification: Primary 60G99
Secondary 05C80; 60E15

1. Introduction and main results

1.1. Introduction

Complex networks are everywhere. Prominent examples include social networks, the Internet, the World Wide Web, transportation networks, etc. In any network, it is of great interest to be able to quantify who is 'important' and who is less so. This is what *centrality measures* aim to do.

There are several well-known centrality measures in networks [38, Chapter 7], such as degree centrality, PageRank centrality, and betweenness centrality. PageRank, first proposed in [39], can be visualised as the stationary distribution of a random walk with uniform restarts. Betweenness centrality was first defined by Anthonisse [2]. For a survey of centrality measures, see Boldi and Vigna [8] and Borgatti [12].

968 R. VAN DER HOFSTAD AND M. PANDEY

There are many ways in which we can compare the effectiveness of centrality measures. Here, one can think of importance for spreading diseases or information [42], for being an information bridge between different communities, or for being an important source of information in a network of scientific papers [41].

In this paper, we compare centrality measures by investigating the rate of *disintegration* of the network by the removal of central vertices. This approach quantifies the notion that a centrality measure is more effective when the network disintegrates more upon the removal of the most central vertices. This work is motivated by the analysis and simulations performed in [35], where the number of connected components and the size of the giant were compared after the removal of the most central vertices based on several centrality measures from a simulation perspective.

Central questions. Our key questions are as follows.

- How will the number of connected components grow with the removal of the most/least central vertices?

- Does vertex removal with respect to centrality measures preserve the existence of local limits?

- What are the subcritical and supercritical regimes of the giant component for vertex removals with respect to centrality measure?

- What is the proportion of vertices in the giant component?

Main innovation of this paper. To answer the above questions, we rely on the theory of *local convergence* [1, 6]; see also [23, Chapter 2] for an extensive overview. We show that when considering *strictly local* centrality measures, i.e. measures that depend on a fixed-radius neighbourhood of a vertex, the number of connected components after the vertex removal procedure converges, and the local limit of the removed graph can be determined in terms of the original local limit. Thus, this answers the first two questions rather generally, assuming local convergence.

It is well known that the *giant* in a random graph is not determined by its local limit (even though it frequently 'almost' is: see [22]). While the upper bound on the giant is always equal to the survival probability of the local limit, the matching *lower bound* can be different. For instance, take two disjoint unions of graphs with the same size and local limit. Then, the survival probability of the local limit remains the same but the proportion of vertices in the largest connected component is reduced by a factor of 2, which is strictly less than the survival probability if it is positive. To give an example where we can prove that the giant of the vertex-removed graph equals the survival probability of its local limit, we restrict our attention to the configuration model with a given degree distribution, and degree centrality. There, we identify the giant of the vertex-removed graph, and also show that the giant is smaller when removing more degree-central vertices.

1.2. Preliminaries

In this section, we define centrality measures and then give an informal introduction to local convergence, as these play a central role in this paper. We will assume that $|V(G)| = n$, and for convenience assume that $V(G) = \{1, \ldots, n\} \equiv [n]$.

1.2.1. *Centrality measures.* In this section, we define *centrality measures*.

Definition 1.1. (*Centrality measures in undirected graphs.*) For a *finite* (undirected) graph $G = (V(G), E(G))$, a *centrality measure* is a function $R \colon V(G) \to \mathbb{R}_{\geq 0}$, where we consider v to be central if $R(v)$ is large.

Centrality measures can be extended to *directed* graphs, but in this paper we restrict to undirected graphs as in [35]. Commonly used examples of centrality measures are as follows.

Degree centrality. In degree centrality, we rank the vertices according to their degrees and then divide the rank by the total number of vertices. We *randomly* assign ranks to vertices with the same degree, to remove ties.

PageRank centrality. PageRank is a popularly known algorithm for measuring centrality in the World Wide Web [13].

Definition 1.2. (*PageRank centrality.*) Consider a finite (undirected) graph G. Let $e_{j,i}$ be the number of edges between j and i. Denote the degree of vertex $i \in [n]$ by d_i. Fix a damping factor or teleportation parameter $c \in (0, 1)$. Then PageRank is the unique probability vector $\boldsymbol{\pi}_n = (\pi_n(i))_{i \in [n]}$ that satisfies that, for every $i \in [n]$,

$$\pi_n(i) = c \sum_{j \in [n]} \frac{e_{j,i}}{d_j} \pi_n(j) + \frac{1-c}{n}, \quad i \in [n].$$

Suppose $\mathbf{1} = (1, 1, \ldots, 1)$ is the all-one vector. Then, with $P = (p_{i,j})_{i,j \in V(G)}$ and $p_{i,j} = e_{i,j}/d_i$ the transition matrix for the random walk on G,

$$\boldsymbol{\pi}_n = \left(\frac{1-c}{n} \right) \mathbf{1} (I - cP)^{-1}. \tag{1.1}$$

In order to use local convergence techniques [19], it is useful to work with the *graph-normalised* PageRank, given by

$$\boldsymbol{R}_n = n \boldsymbol{\pi}_n, \quad \text{for which } \frac{1}{n} \sum_{j \in [n]} R_n(j) = 1.$$

Since $c \in (0, 1)$, (1.1) implies

$$\boldsymbol{\pi}_n = \left(\frac{1-c}{n} \right) \mathbf{1} \sum_{k=0}^{\infty} c^k P^k. \tag{1.2}$$

Equation (1.2) is sometimes called *power-iteration* for PageRank [3, 7, 9]. It will also be useful to consider the finite-radius PageRank, $\boldsymbol{R}_n^{(N)}$, for $N \in \mathbb{N}$, defined by

$$\boldsymbol{R}_n^{(N)} := \left(\frac{1-c}{n} \right) \mathbf{1} \sum_{k=0}^{N} c^k P^k, \tag{1.3}$$

which approximates PageRank by a finite number of powers in (1.2). PageRank has been intensely studied on various random graph models [5, 15, 16, 19, 31, 32, 34], with the main

focus being to prove or disprove that PageRank has the same power-law exponent as the in-degree distribution.

Other popular centrality measures. *Closeness centrality* measures the average distance of a vertex from a randomly chosen vertex. The higher the average, the lower the centrality index, and *vice versa*. Interestingly, Evans and Chen [18] predict that closeness centrality is closely related to degree centrality, at least for locally tree-like graphs. *Betweenness centrality* measures the extent to which vertices are important for *interconnecting* different vertices, and is given by

$$b_v = \sum_{u \neq v \neq w} \frac{\sigma_{u,v}(w)}{\sigma_{u,w}},$$

where $\sigma_{u,v}$ is the number of shortest paths from vertex u to vertex v, and $\sigma_{u,v}(w)$ is the number of shortest paths between u and v containing w.

In this paper we mainly work with *strictly local* centrality measures.

Definition 1.3. (*Strictly local centrality measures.*) For a *finite* (undirected) graph $G = (V(G), E(G))$, a *strictly local centrality measure* is a centrality measure $R: V(G) \to \mathbb{R}_{\geq 0}$, such that there exists an $r \in \mathbb{N}$, and for each vertex $v \in V(G)$, $R(v)$ depends only on the graph through the neighbourhood

$$B_r^{(G)}(v) := \{u \in V(G): \mathrm{dist}_G(u, v) \leq r\},$$

where $\mathrm{dist}_G(u, v)$ denotes the graph distance between u and v in G.

PageRank is very well approximated by its strictly local version as in (1.3) [4, 8, 19].

1.2.2. *Local convergence of undirected random graphs.* In this section, we informally introduce local convergence of random graphs, which describes what a graph locally looks like from the perspective of a uniformly chosen vertex, as the number of vertices in a graph goes to infinity. For example, the sparse Erdős–Rényi random graph, which is formed by bond-percolation on the complete graph, locally looks like a Poisson branching process, as n tends to infinity [23, Chapter 2]. Before moving further, we discuss the following notation.

Notation 1.1. (*Probability convergence.*) Suppose $(X_n)_{n \geq 1}$, $(Y_n)_{n \geq 1}$ are two sequences of random variables and X is a random variable.

(1) We write $X_n \xrightarrow{\mathbb{P}/d} X$ when X_n converges in probability/distribution to X.

(2) We write $X_n = o_{\mathbb{P}}(Y_n)$ when $X_n/Y_n \xrightarrow{\mathbb{P}} 0$.

Now let us give the informal definition of local convergence. We will rely on two types of local convergence, namely, local weak convergence and local convergence in probability. Let o_n denote a uniformly chosen vertex from $V(G_n)$. Local weak convergence means that

$$\mathbb{P}\big(B_r^{(G_n)}(o_n) \cong (H, o')\big) \to \bar{\mu}\big(B_r^{(\bar{G})}(o) \cong (H, o')\big), \tag{1.4}$$

for all rooted graphs (H, o'), where a rooted graph is a graph with a distinguished vertex in the vertex set $V(H)$ of H. In (1.4), $(\bar{G}, o) \sim \bar{\mu}$ is a *random* rooted graph, which is called the local weak limit. For local convergence in probability, instead, we require that

$$\frac{1}{|V(G_n)|} \sum_{v \in V(G_n)} \mathbb{1}_{\{B_r^{(G_n)}(v) \cong (H, o')\}} \xrightarrow{\mathbb{P}} \mu\big(B_r^{(G)}(o) \cong (H, o')\big) \tag{1.5}$$

holds for all rooted graphs (H, o'). In (1.5), $(G, o) \sim \mu$ is a random rooted graph which is called the local limit in probability (and bear in mind that μ can possibly be a random measure on rooted graphs). Both (1.4) and (1.5) describe the convergence of the proportions of vertices around which the graph locally looks like a certain specific graph. We discuss this definition more formally in Section 2.1. We now turn to our main results.

1.3. Main results

In this section, we state our main results. In Section 1.3.1 we discuss our results that hold for general strictly local centrality measures on locally converging random graph sequences. In Section 1.3.2 we investigate the size of the giant after vertex removal. Due to the non-local nature of the giant, there we restrict to *degree centrality* on the *configuration model*.

1.3.1. *Strictly local centrality measures.* We first define our vertex removal procedure based on centrality.

Definition 1.4. (*Vertex removal based on centrality threshold.*) Let G be an arbitrary graph. Define $G(R, r)$ to be the graph obtained after removing all the vertices v for which $R(v) > r$. We call this the *r-killed graph of G*.

Theorem 1.1. (*Continuity of vertex removal.*) *Let R be a strictly local centrality measure and $(G_n)_{n \geq 1}$ a sequence of random graphs that converges locally weakly/locally in probability to $(G, o) \sim \mu$. Then $(G_n(R, r))_{n \geq 1}$ converges locally weakly/locally in probability to $(G(R, r), o)$, respectively.*

Remark 1.1. (*Extension to 'almost' local centrality measures.*) We believe that the proof for Theorem 1.1 can be extended to 'almost' local centrality measures like PageRank, but the proof is much more involved.

Theorem 1.1 means that vertex removal with respect to a strictly local centrality threshold is continuous with respect to the local convergence.

Let $C_r(o)$ denote the connected component containing the root in the r-killed graph, $(G(R, r), o) \sim \mu$. As a corollary to Theorem 1.1, we bound the limiting value of the proportion of vertices in the giant component in $G_n(R, r)$ in probability by the survival probability of $(G(R, r), o) \sim \mu$, which is $\mu(|C_r(o)| = \infty)$.

Corollary 1.1. (*Upper bound on giant.*) *Let $v(C_1(G_n(R, r)))$ denote the number of vertices in the giant component of $G_n(R, r)$. Under the conditions of Theorem 1.1,*

$$\lim_{n \to \infty} \mathbb{P}(v(C_1(G_n(R, r))) \leq n(\zeta + \varepsilon)) = 1,$$

for all $\varepsilon > 0$, where $\zeta = \mu(|C_r(o)| = \infty)$.

The intuition for Corollary 1.1 is as follows. For any integer k, the set of vertices with component sizes larger than k, if non-empty, contains the vertices in the largest connected component. Thus the proportion of vertices with component size larger than k is an upper bound for the proportion of vertices in the giant component. The former, when divided by n, is the probability that a uniformly chosen vertex has a component size greater than k, which converges due to local convergence properties of the graph sequence. Sending k to infinity gives us the required upper bound.

Let $K_n^r(R)$ denote the number of connected component in the killed graph $G_n(R, r)$. As another corollary to Theorem 1.1, we give the convergence properties for $K_n^r(R)$.

Corollary 1.2. (Number of connected components.) *Under the conditions of Theorem 1.1, we have the following.*

(a) *If* $(G_n)_{n \geq 1}$ *converges locally in probability to* $(G, o) \sim \mu$, *then*

$$\frac{K_n^r(R)}{n} \overset{\mathbb{P}}{\to} \mathbb{E}_\mu \left[\frac{1}{|C_r(o)|} \right].$$

(b) *If* $(G_n)_{n \geq 1}$ *converges locally weakly to* $(\bar{G}, \bar{o}) \sim \bar{\mu}$, *then*

$$\frac{\mathbb{E}[K_n^r(R)]}{n} \to \mathbb{E}_{\bar{\mu}} \left[\frac{1}{|C_r(o)|} \right].$$

The limit in Corollary 1.2 follows from the identity

$$\frac{K_n}{n} = \frac{1}{n} \sum_{v \in V(G)} \frac{1}{|C(v)|} = \mathbb{E}_n \left[\frac{1}{|C(o_n)|} \right],$$

where o_n is a uniformly chosen vertex and $|C(v)|$ is the connected component containing v. The function $f(G, v) = 1/|C(v)|$ is continuous and bounded on the metric space of rooted graphs. Thus we can apply local convergence results on this function.

1.3.2. *Degree centrality and configuration model.* In this section, we restrict to degree centrality in the configuration model. Before starting with our main results, we introduce the configuration model.

Configuration model. The configuration model was introduced by Bollobás [10]; see also [21, Chapter 7] and the references therein for an extensive introduction. It is one of the simplest possible models for generating a random graph with a given degree distribution. Written as $CM_n(\boldsymbol{d})$, it is a random graph on n vertices having a given degree sequence $\boldsymbol{d}$, where $\boldsymbol{d} = (d_1, d_2, \ldots, d_n) \in \mathbb{N}^n$. The giant in the configuration model has attracted considerable attention in, for example, [11], [22], [30], [36], and [37]. It is also known how the tail of the limiting degree distribution influences the size of giant [17]. Further, the diameter and distances in the supercritical regime have been studied in [24], [25], and [26], while criteria for the graph to be simple appear in [14], [28], and [29]. In this paper we assume that the degrees satisfy the following usual conditions.

Condition 1.1. (Degree conditions.) *Let* $\boldsymbol{d} = (d_1, d_2, \ldots, d_n)$ *denote a degree sequence. Let* $n_j = \{v: d_v = j\}$ *denote the number of vertices with degree j. We assume that there exists a probability distribution* $(p_j)_{j \geq 1}$ *such that the following hold:*

(a) $\lim_{n \to \infty} n_j / n = p_j$,

(b) $\lim_{n \to \infty} \sum_{j \geq 1} j n_j / n = \sum_{j \geq 1} j p_j < \infty$.

Let D be a non-negative integer random variable with probability mass function $(p_j)_{j \geq 1}$. These regularity conditions ensure that the sequence of graphs converges locally in probability to a unimodular branching process, with degree distribution D. See [23, Chapter 4] for more details.

Generalised vertex removal based on degrees. We wish to study the effect of the removal of the α proportion of vertices with the highest or lowest degrees on the giant of the configuration model. For this, we define α-sequences with respect to a probability mass function as follows.

Definition 1.5. (α-sequence.) Fix $\alpha \in (0, 1)$. Let $\boldsymbol{r} := (r_i)_{i \geq 1}$ be a sequence of elements of $[0, 1]$ satisfying

$$\mathbb{E}[r_D] := \sum_{i \geq 1} p_i r_i = \alpha.$$

Then $\boldsymbol{r}$ is called an α-sequence with respect to $\boldsymbol{p} = (p_j)_{j \geq 1}$.

Suppose $(G_n)_{n \geq 1}$ is a sequence of random graphs converging locally in probability to $(G, o) \sim \mu$, and let the limiting degree distribution be D, with probability mass function as $\boldsymbol{p} = (p_j)_{j \geq 1}$. Suppose $\boldsymbol{r} = (r_j)_{j \geq 1}$ is an α-sequence with respect to $\boldsymbol{p}$; then we define vertex removal based on α-sequences as follows.

Definition 1.6. (*Vertex removal based on α-sequences.*) Remove $\lfloor n r_i p_i \rfloor$ vertices of degree i from G_n uniformly at random, for each $n \geq 1$. This gives us the *vertex-removed graph according to the α-sequence $\boldsymbol{r} = (r_j)_{j \geq 1}$, denoted by $(G_{n,\boldsymbol{r}})_{n \geq 1}$.*

Remark 1.2. In $G_{n,\boldsymbol{r}}$, we asymptotically remove an α proportion of vertices because, due to Condition 1.1 and the dominated convergence theorem,

$$\lim_{n \to \infty} \sum_{j \geq 1} r_j \frac{n_j}{n} = \sum_{j \geq 1} r_j p_j = \alpha.$$

Results for the configuration model. Let $(G_n)_{n \geq 1}$ be a sequence of random graphs satisfying $G_n \sim \mathrm{CM}_n(\boldsymbol{d})$. Throughout the paper, we shall assume that

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

Indeed, for $\nu < 1$, we already know that there is no giant to start with, and it cannot appear by removing vertices.

Definition 1.7. (*r-set.*) Suppose $\boldsymbol{r} = (r_j)_{j \geq 1}$ is an α-sequence with respect to $\boldsymbol{p} = (p_j)_{j \geq 1}$. Let $X = [0, 1]^{\mathbb{N}}$ be the set of all sequences in $[0, 1]$. Define the set

$$S(\boldsymbol{r}) = \left\{ (\boldsymbol{r}^{(n)})_{n \geq 1} \in X^{\mathbb{N}} : \lim_{n \to \infty} r_i^{(n)} = r_i \ \forall i \geq 1 \right\}.$$

An $\boldsymbol{r}$-set $S(\boldsymbol{r})$ can be thought of as the set of those sequences in $[0, 1]^{\mathbb{N}}$ which converge to $\boldsymbol{r}$ component-wise. Let $v(C_1(\boldsymbol{r}))$ and $e(C_1(\boldsymbol{r}))$ denote the number of vertices and edges in the giant component of $G_{n,\boldsymbol{r}}$. The following theorem describes the giant in $G_{n,\boldsymbol{r}^{(n)}}$.

Theorem 1.2. (Existence of giant after vertex removal.) *Let $\boldsymbol{r}$ be an α-sequence. Let $(\boldsymbol{r}^{(n)})_{n \geq 1}$ be a sequence from the $\boldsymbol{r}$-set $S(\boldsymbol{r})$. Then the graph $G_{n,\boldsymbol{r}^{(n)}}$ has a giant component if and only if $\nu_{\boldsymbol{r}} > 1$, where*

$$\nu_{\boldsymbol{r}} := \frac{\mathbb{E}[D(D-1)(1-r_D)]}{\mathbb{E}[D]}. \tag{1.6}$$

For $\nu_{\mathbf{r}} > 1$,

$$\frac{v(C_1(\mathbf{r}^{(n)}))}{n} \xrightarrow{\mathbb{P}} \rho(\mathbf{r}) := 1 - \alpha - 2\mathbb{E}[Dr_D]\eta_{\mathbf{r}} - \sum_{i \geq 1}(1 - r_i)p_i\eta_{\mathbf{r}}^i, \tag{1.7}$$

$$\frac{e(C_1(\mathbf{r}^{(n)}))}{n} \xrightarrow{\mathbb{P}} e(\mathbf{r}) := \frac{\mathbb{E}[D]}{2}(1 - \eta_{\mathbf{r}}^2) - \mathbb{E}[Dr_D](1 - \eta_{\mathbf{r}}). \tag{1.8}$$

In the above equations, $\eta_{\mathbf{r}} \in (0, 1]$ is the smallest solution of

$$g_{\mathbf{r}}'(\eta_{\mathbf{r}}) = \frac{\mathbb{E}[D]}{\beta_{\mathbf{r}}}\eta_{\mathbf{r}}, \tag{1.9}$$

where $\beta_{\mathbf{r}} = \mathbb{E}[Dr_D] + 1 - \alpha$ and $g_{\mathbf{r}}(\cdot)$ is the generating function for a random variable dependent on $\mathbf{r}$ given by

$$g_{\mathbf{r}}(s) = \frac{\sum_{i=1}^{\infty}(1 - r_i)p_i s^i + \mathbb{E}[Dr_D]s}{\beta_{\mathbf{r}}}. \tag{1.10}$$

Remark 1.3. (*Class of sequences with the same limiting proportion.*) For the class of sequences which converge to a given α-sequence component-wise, we always have the same limiting proportion of vertices/edges in the giant component. Thus, choosing our sequence $(\mathbf{r}^{(n)})_{n \geq 1}$ appropriately, it is always possible to remove an exact α proportion of the vertices from each G_n having the same limiting graph.

We next investigate the effect of removing the α proportion of vertices with the highest/lowest degrees. For that we first define quantiles.

Definition 1.8. (*Degree centrality quantiles.*) For each $\alpha \in (0, 1)$, let k_α, the top α-quantile, satisfy

$$\mathbb{P}(D > k_\alpha) < \alpha \quad \text{and} \quad \mathbb{P}(D \geq k_\alpha) \geq \alpha.$$

Similarly, for each $\alpha \in (0, 1)$, let l_α, the bottom α-quantile, satisfy

$$\mathbb{P}(D < l_\alpha) < \alpha \quad \text{and} \quad \mathbb{P}(D \leq l_\alpha) \geq \alpha.$$

Definition 1.9. (*α-sequences corresponding to top and bottom removal.*) Let k be the top α-quantile for the degree distribution D. Define $\bar{\mathbf{r}}(\alpha)$ to have coordinates equal to zero until the kth coordinate, which is $(\alpha - \mathbb{P}(D > k))/p_k$, and ones thereafter. Then $\bar{\mathbf{r}}(\alpha)$ is the *α-sequence corresponding to the top α-removal.*

Similarly, let l be the lower α-quantile for the degree distribution D. Define $\underline{\mathbf{r}}(\alpha)$ to have coordinates equal to one until the lth coordinate, which is $(\alpha - \mathbb{P}(D < l))/p_l$, and zeros thereafter. Then $\underline{\mathbf{r}}(\alpha)$ is the *α-sequence corresponding to the bottom α-removal.*

Corollary 1.3. (Highest/lowest α-proportion removal.) *Let $\bar{\alpha}_c$ and $\underline{\alpha}_c$ be defined as*

$$\bar{\alpha}_c = \inf\{\alpha > 0 \colon \mathbb{E}[D(D - 1)\mathbb{1}\{D \leq k_\alpha\}] > \mathbb{E}[D] + k_\alpha(k_\alpha - 1)(\alpha - \mathbb{P}(D > k_\alpha))\},$$

$$\underline{\alpha}_c = \inf\{\alpha > 0 \colon \mathbb{E}[D(D - 1)\mathbb{1}\{D \geq l_\alpha\}] > \mathbb{E}[D] + l_\alpha(l_\alpha - 1)(\alpha - \mathbb{P}(D < l_\alpha))\}.$$

(1) *Let $v(\bar{C}_1^\alpha)$ and $e(\bar{C}_1^\alpha)$ be the number of vertices and edges in the largest connected component of the top α-proportion degree vertex-removed graph, respectively. If $\alpha \geq \bar{\alpha}_c$, then there is no giant component, i.e. $v(\bar{C}_1^\alpha)/n \xrightarrow{\mathbb{P}} 0$. If $\alpha < \bar{\alpha}_c$, then the giant exists and*

$$\frac{v(\bar{C}_1^\alpha)}{n} \xrightarrow{\mathbb{P}} \rho(\bar{r}(\alpha)) > 0 \quad and \quad \frac{e(\bar{C}_1^\alpha)}{n} \xrightarrow{\mathbb{P}} e(\bar{r}(\alpha)) > 0.$$

(2) *Let $v(\underline{C}_1^\alpha)$ and $e(\underline{C}_1^\alpha)$ be the number of vertices and edges in the largest connected component of the lowest α proportion degree vertex-removed graph, respectively. If $\alpha \geq \underline{\alpha}_c$, then there is no giant, i.e. $v(\underline{C}_1^\alpha)/n \xrightarrow{\mathbb{P}} 0$. If $\alpha < \underline{\alpha}_c$, then the giant exists and*

$$\frac{v(\underline{C}_1^\alpha)}{n} \xrightarrow{\mathbb{P}} \rho(\underline{r}(\alpha)) > 0 \quad and \quad \frac{e(\underline{C}_1^\alpha)}{n} \xrightarrow{\mathbb{P}} e(\underline{r}(\alpha)) > 0.$$

We next give bounds on the size of the giant in the vertex-removed graph.

Theorem 1.3. (Bounds for proportion of giant.) *Let r be an α-sequence. Then*

$$\mathbb{E}[D(\eta_r - r_D)](1 - \eta_r) \leq \rho(r) \leq \mathbb{E}[D(1 - r_D)](1 - \eta_r).$$

In the above equation, $\eta_r \in (0, 1]$ satisfies (1.9). *Furthermore,*

$$\rho(r) \leq \frac{\mathbb{E}[D(1 - r_D)]^2}{\mathbb{E}[D]} \quad and \quad e(r) \leq \frac{\mathbb{E}[D(1 - r_D)]^2}{2\mathbb{E}[D]} \tag{1.11}$$

and

$$\rho(r) \leq 1 - \alpha - \frac{2\mathbb{E}[Dr_D]^2}{\mathbb{E}[D]} < 1 - \alpha - \frac{2\alpha^2}{\mathbb{E}[D]}. \tag{1.12}$$

Additionally, for α-sequences which are positively correlated with degree, we have

$$\rho(r) \leq 1 - \alpha - 2\alpha^2\mathbb{E}[D]. \tag{1.13}$$

Remark 1.4. $\bar{r}(\alpha)$ is positively correlated with degree. Thus, for top α-proportion removal (1.13), we have

$$\rho(\bar{r}(\alpha)) \leq 1 - \alpha - 2\alpha^2\mathbb{E}[D].$$

The next logical question is which α-sequence destroys the graph the most. Intuitively, the size of the giant should be the smallest if we remove the top α-proportion degree vertices and the largest when we remove the bottom α proportion degree vertices, i.e. $\rho(\bar{r}(\alpha)) \leq \rho(r) \leq \rho(\underline{r}(\alpha))$ and $e(\bar{r}(\alpha)) \leq e(r) \leq e(\underline{r}(\alpha))$. To prove this, we first define a partial order over the set of all α-sequences that is able to capture which vertex removal is more effective for the destruction of the giant in the configuration model. For this, we recall that, for two non-negative measures μ_1 and μ_2 on the real line, we say that μ_1 is stochastically dominated by μ_2, and write $\mu_1 \preceq_{\text{st}} \mu_2$, if $\mu_1([K, \infty)) \leq \mu_2([K, \infty))$ for all $K \in \mathbb{R}$. We next extend this to a partial order over the set of all α-sequences.

Definition 1.10. (*Stochastically dominant sequences.*) Let $r = (r_j)_{j \geq 1}$ and $r' = (r'_j)_{j \geq 1}$ be α_1- and α_2-sequences, respectively. Then the sequences $q = (p_j r_j)_{j \geq 1}$ and $q' = (p_j r'_j)_{j \geq 1}$ form finite non-negative measures. We say that r' stochastically dominates r, or $r \preceq_p r'$, if $q \preceq_{\text{st}} q'$.

The following theorem answers which are the best and worst ways to remove α-proportion vertices from a configuration model, and compares other α-sequences.

Theorem 1.4. (Comparison between α-sequences.) *Let $\boldsymbol{r} = (r_j)_{j \geq 1}$ and $\boldsymbol{r}' = (r'_j)_{j \geq 1}$ be two α-sequences such that $\boldsymbol{r}' \preccurlyeq_{\boldsymbol{p}} \boldsymbol{r}$. Then*

$$\rho(\bar{\boldsymbol{r}}(\alpha)) \leq \rho(\boldsymbol{r}) \leq \rho(\boldsymbol{r}') \leq \rho(\underline{\boldsymbol{r}}(\alpha)) \quad and \quad e(\bar{\boldsymbol{r}}(\alpha)) \leq e(\boldsymbol{r}) \leq e(\boldsymbol{r}') \leq e(\underline{\boldsymbol{r}}(\alpha)).$$

As a result, the critical α for the bottom and top vertex removal satisfy $\underline{\alpha}_c > \bar{\alpha}_c$.

The following corollary is stronger than the theorem (as the corollary immediately implies the theorem), but it follows from the proof of Theorem 1.4.

Corollary 1.4. (General comparison.) *Suppose $\boldsymbol{r}, \boldsymbol{r}' \in [0, 1]^{\mathbb{N}}$ satisfy $\boldsymbol{r} \preccurlyeq_{\boldsymbol{p}} \boldsymbol{r}'$. Then $\rho(\boldsymbol{r}') \leq \rho(\boldsymbol{r})$ and $e(\boldsymbol{r}') \leq e(\boldsymbol{r})$.*

1.4. Overview of the proof

In this section, we give an overview to the proofs. The theorems in this paper are of two types: the first concerns arbitrary strictly local centrality measures on a sequence of locally converging random graphs, whereas the second concerns degree centrality on the configuration model. We discuss these separately in the next two sections.

1.4.1. *Strictly local centrality measures.* In this section, we discuss the proof structure for Theorem 1.1, which involves the following steps.

(1) Proof that the vertex removal function with respect to a strictly local centrality measure is *continuous* with respect to the rooted graph topology. As a result, if a function on the set of rooted graphs is continuous, then this function composed with the function of vertex removal with respect to the strictly local centrality R is also continuous.

(2) Next we use the above observations and the local convergence of the sequence $(G_n)_{n \geq 1}$ to complete the proof of Theorem 1.1, and Corollaries 1.1 and 1.2 are immediate consequences.

1.4.2. *Configuration model and degree centrality.* In this section, we discuss a construction, motivated by Janson [27], and how it helps in the study of vertex removal procedure on the configuration model.

Let $G = (V, E)$ be a graph. Suppose $V' \subset V$ is the set of vertices that we wish to remove. We will start by constructing an intermediate graph, which we call the vertex- *exploded graph*, and denote it by $\tilde{G}(V')$. To obtain $\tilde{G}(V')$ from G, we use the following steps.

(1) Index the vertices of the set $V' = \{v_1, v_2, \ldots, v_m\}$, where $m = |V'|$.

(2) *Explode* the first vertex from the set V' into d_{v_1} vertices of degree 1.

(3) Label these new degree-1 vertices *red*.

(4) Do this for each vertex of V'.

The half-edges incident to the exploded vertices are redistributed to form the half-edges incident to the red vertices. Thus, after explosion, a degree-k vertex falls apart into k degree-1 vertices. This is illustrated in Figure 1, where a degree-k vertex is exploded.

Remark 1.5. (*Significance of vertex explosion.*) The vertex explosion ensures that the exploded graph is again a configuration model. Removing the degree-1 vertices of the exploded graph
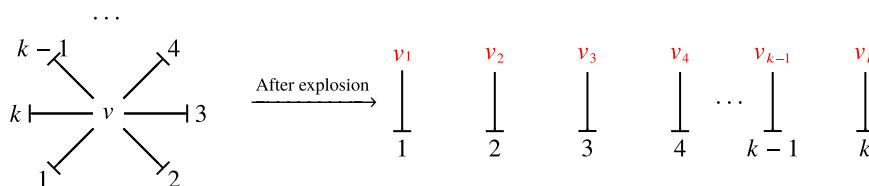
FIGURE 1. Explosion of a degree-k vertex.

does not non-trivially change the component structure of components having linear size. The removal of the red vertices, which are degree-1 vertices, from the exploded graph gives the vertex-removed graph.

We denote the number of vertices and degree sequence of $\tilde{G}(V')$ by $\tilde{n}$ and its degree sequence by $\tilde{\boldsymbol{d}}$. The removal of red vertices, together with the (single) edge adjacent to them, from $\tilde{G}(V')$ gives the V'-removed graph. Next, we show that the vertex-exploded graph remains a configuration model, if vertices are removed with respect to their degree sequence. This means that we only look at the degree sequence to decide which vertices to remove, independently of the graph. It is this property that makes the configuration model amenable for vertex removal based on degree centrality.

Lemma 1.1. (Exploded configuration model is still a configuration model.) *Let $G_n = \mathrm{CM}_n(\boldsymbol{d})$ be a configuration model with degree sequence $\boldsymbol{d} = (d_i)_{i \in [n]}$. Let $\tilde{G}_n$ be the vertex-exploded graph formed from G_n, where vertices are exploded with respect to the degree sequence. Then $\tilde{G}_n$ has the same distribution as $\mathrm{CM}_{\tilde{n}}(\tilde{\boldsymbol{d}})$.*

Proof. A degree sequence in a configuration model can be viewed as half-edges incident to vertices. These half-edges are matched via a uniform perfect matching. Suppose one explodes a vertex of degree i in $\mathrm{CM}_n(\boldsymbol{d})$. Then we still have its i half-edges, but these half-edges are now incident to i newly formed red vertices of degree 1. Since we are just changing the vertex to which these half-edges are incident, the half-edges in the vertex-exploded graph are still matched via a uniform perfect matching, so that the newly formed graph is also a configuration model, but now with $\tilde{n}$ vertices and degree sequence $\tilde{\boldsymbol{d}}$. $\qquad\square$

This lemma is useful because of the following result by Janson and Luczak [30] to study the giant component in a configuration model.

Theorem 1.5. (Janson and Luczak [30].) *Consider $\mathrm{CM}_n(\boldsymbol{d})$ satisfying regularity Condition 1.1 with limiting degree distribution D. Let $g_D(x) := \sum_j p_j x^j$ be the probability generating function of D. Assume $p_1 > 0$.*

(i) *If $\mathbb{E}[D(D-2)] > 0$, then there exists unique $\eta \in (0, 1)$ such that $g'_D(\eta) = \mathbb{E}[D]\eta$ and*

    (a) *$v(C_1)/n \xrightarrow{\mathbb{P}} 1 - g_D(\eta)$,*

    (b) *$v_j(C_1)/n \xrightarrow{\mathbb{P}} p_j(1 - \eta^j)$ for all $j \geq 0$,*

    (c) *$e(C_1)/n \xrightarrow{\mathbb{P}} (\mathbb{E}[D]/2)(1 - \eta^2)$.*

(ii) *If $\mathbb{E}[D(D-2)] \leq 0$, then $v(C_1)/n \xrightarrow{\mathbb{P}} 0$ and $e(C_1)/n \xrightarrow{\mathbb{P}} 0$.*

Here C_1 denotes the largest component of $\mathrm{CM}_n(\boldsymbol{d})$.

Intuition. The local limit of the configuration model is a unimodular branching process with root offspring distribution $(p_k)_{k\geq 1}$. A unimodular branching process is a branching process in which the offspring distribution of future generations, generations coming after the root, is given by the size-biased distribution of the root's offspring distribution minus one, denoted by $(p_k^*)_{k\geq 1}$, that is,

$$p_k^* = \frac{(k+1)p_{k+1}}{\sum_{j\geq 1} jp_j}.$$

In terms of this local limit, we can interpret the limits in Theorem 1.5 as follows. First, η is the extinction probability of a half-edge. We think of this as the extinction probability of a half-edge in the configuration model. Then the proportion of surviving degree-k vertices is given by $p_k(1-\eta^k)$, which is the proportion of degree-k vertices times the probability that one of its k half-edges survives as in Theorem 1.5(i)(b). Summing this over all possible values of k gives Theorem 1.5(i)(a). Then η, the extinction probability of the root of the branching process when it has degree 1, satisfies the equation

$$\eta = \sum_k p_k^*\eta^k = \sum_{k\geq 1} \frac{(k+1)p_{k+1}}{\mathbb{E}[D]}\eta^k = \frac{g_D'(\eta)}{\mathbb{E}[D]}.$$

Similarly, for Theorem 1.5(i)(c), an edge does not survive when both its half-edges go extinct, which occurs with probability η^2. Thus

$$\frac{e(C_1)}{n} = \frac{1}{n}\sum_{e\in G_n} \mathbb{1}_{\{e\in C_1\}} = \frac{\ell_n}{n}\left(\frac{1}{\ell_n}\sum_{e\in G_n} \mathbb{1}_{\{e\in C_1\}}\right) \to \frac{\mathbb{E}[D]}{2}(1-\eta^2),$$

where $\ell_n = \sum_v d_v$ is the number of half-edges. Thus $\lim_{n\to\infty} \ell_n/n = \mathbb{E}[D]$, due to Condition 1.1(b).

Using this theorem on the vertex-exploded graph, we obtain the size of giant (vertices and edges). Removal of the red vertices in the vertex-exploded graph gives the required vertex-removed graph. This fact is used to complete the proof of Theorem 1.2. The remaining results are proved by carefully investigating the limiting proportion of vertices and edges in the giant, and how they relate to the precise α-sequence.

1.5. Discussion and open problems

In this section, we discuss the motivation of this paper and some open problems. The problem of vertex removal with respect to centrality measures was motivated by Mocanu *et al.* [35], which showed by simulation how the size of the giant and the number of connected components behave for vertex removal, based on different centrality measures. These plots were used to compare the effectiveness of these centrality measures.

A related problem is whether we can compare the size of giants in configuration models if their limiting degree distributions have some stochastic ordering. This question can be answered using the notion of ε-transformation as discussed in Definition 3.1 below. Let $\rho_{\mathrm{CM}}(\boldsymbol{p})$ denote the size of giant for the configuration model case when the limiting degree distribution is $\boldsymbol{p}$.

Theorem 1.6. (Stochastic ordering and giant in configuration model.) *If* $\boldsymbol{p} \preccurlyeq_{\mathrm{st}} \boldsymbol{q}$, *then* $\rho_{\mathrm{CM}}(\boldsymbol{p}) \leq \rho_{\mathrm{CM}}(\boldsymbol{q})$.

A similar result is proved in [33] for increasing concave ordering (a stochastic order different from the one used in this paper), but it requires a strong additional assumption that we can remove.

Open problems. We close this section with some related open problems.

(1) What is the size of the configuration model giant for the vertex-removed graph in the case of strictly local centrality measures? Which strictly local centrality measure performs the best? Can we extend this to non-local centrality measures?

(2) *Strictly local approximations to non-local centrality measures.* PageRank can be very well approximated by the strictly local approximation in (1.3). Can betweenness, closeness, and other important centrality measures also be well approximated by strictly local centrality measures? What graph properties are needed for such an approximation to be valid? Is it true on configuration models, as predicted for closeness centrality in [18] and suggested for both closeness and betweenness in [20]? If so, then one can use these strictly local centrality measures for vertex removal procedures.

(3) *Stochastic monotonicity for the number of connected components.* If one starts removing vertices from a graph, the number of connected components first increases and then decreases. Where does it start decreasing?

(4) *Extension of results to directed graphs.* Can we extend the results of this paper to directed graphs? For the directed graphs, there are several connectivity notions, while vertices have *two* degrees, making degree centrality less obvious.

2. Proofs: strictly local centrality measures

In this section, we define *local convergence* of random graphs and then prove the results stated in Section 1.3.1.

2.1. Local convergence of undirected random graphs

To define local convergence, we introduce isomorphisms of rooted graphs, a metric space on them, and finally convergence in it. For more details, refer to [23, Chapter 2].

Definition 2.1. (*Rooted (multi)graph, rooted isomorphism, and neighbourhood.*)

(1) We call a pair (G, o) a *rooted (multi)graph* if G is a locally finite, connected (multi)graph and o is a distinguished vertex of G.

(2) We say that two multigraphs G_1 and G_2 are *isomorphic,* written as $G_1 \cong G_2$, if there exists a bijection $\phi \colon V(G_1) \to V(G_2)$ such that for any $v, w \in V(G_1)$, the number of edges between v, w in G_1 equals the number of edges between $\phi(v), \phi(w)$ in G_2.

(3) We say that the rooted (multi)graphs (G_1, o_1) and (G_2, o_2) are *rooted isomorphic* if there exists a graph-isomorphism between G_1 and G_2 that maps o_1 to o_2.

(4) For $r \in \mathbb{N}$, we define $B_r^{(G)}(o)$, the (closed) r-ball around o in G or r-neighbourhood of o in G, as the subgraph of G spanned by all vertices of graph distance at most r from o. We think of $B_r^{(G)}(o)$ as a rooted (multi)graph with root o.

We next define a metric space over the set of rooted (multi)graphs.

Definition 2.2. (*Metric space on rooted (multi)graphs.*) Let $\mathcal{G}_{\star}$ be the space of all rooted (multi)graphs. Let (G_1, o_1) and (G_2, o_2) be two rooted (multi)graphs. Let

$$R^{\star} = \sup\left\{r \geq 0 \mid B_r^{(G_1)}(o_1) \cong B_r^{(G_2)}(o_2)\right\}.$$

The metric $d_{\mathcal{G}_{\star}} : \mathcal{G}_{\star} \times \mathcal{G}_{\star} \to \mathbb{R}_{\geq 0}$ is defined by

$$d_{\mathcal{G}_{\star}}((G_1, o_1), (G_2, o_2)) := \frac{1}{R^{\star} + 1}.$$

We next define local weak convergence and local convergence in probability for any sequence from the space of rooted random graphs.

Definition 2.3. (*Local weak convergence of rooted (multi)graphs.*) A sequence of random (multi)graphs $(G_n)_{n \geq 1}$ is said to *converge locally weakly* to the random rooted graph (G, o), a random variable taking values in $\mathcal{G}_{\star}$, having law μ, if, for every bounded continuous function $f : \mathcal{G}_{\star} \to \mathbb{R}$, as $n \to \infty$,

$$\lim_{n \to \infty} \mathbb{E}[f(G_n, o_n)] = \mathbb{E}_{\mu}[f(G, o)],$$

where o_n is a uniformly chosen vertex from the vertex set of G_n.

Definition 2.4. (*Local convergence in probability of rooted (multi)graphs.*) A sequence of random (multi)graphs $(G_n)_{n \geq 1}$ is said to *converge locally in probability* to the random rooted graph (G, o), a random variable taking values in $\mathcal{G}_{\star}$, having law μ, when, for every $r > 0$, and for every $(H, o') \in \mathcal{G}_{\star}$, as $n \to \infty$,

$$\frac{1}{|V(G_n)|} \sum_{v \in V(G_n)} \mathbb{1}_{\{B_r^{(G_n)}(v) \cong (H, o')\}} \xrightarrow{\mathbb{P}} \mu\left(B_r^{(G)}(o) \cong (H, o')\right).$$

2.2. Local convergence proofs

In this section, we prove the results stated in Section 1.3.1. Let us take a strictly local centrality measure R. This means that it only depends on a finite fixed neighbourhood, say $N \in \mathbb{N}$. The following lemma shows that the map $(G, o) \to (G(R, r), o)$ is continuous.

Lemma 2.1. (Continuity of centrality-based vertex removal.) *Let R be a strictly local centrality measure and let $h : \mathcal{G}_{\star} \to \mathbb{R}$ be a continuous bounded map, where $\mathcal{G}_{\star}$ is the space of rooted graphs, with the standard metric, d, as in Definition 2.2. Fix $r > 0$. Suppose $f : \mathcal{G}_{\star} \to \mathbb{R}$ satisfies*

$$f(G, o) = h(G(R, r), o).$$

Then $f(\cdot)$ is a bounded continuous function.

Proof. $f(\cdot)$ is bounded because $h(\cdot)$ is bounded. Thus we only need to show that $f(\cdot)$ is continuous, that is, for all $\varepsilon > 0$, there exists $\delta > 0$ such that

$$d_{\mathcal{G}_{\star}}((G_1, o_1), (G_2, o_2)) < \delta \implies |h(G_1(R, r), o_1) - h(G_2(R, r), o_2)| < \varepsilon.$$

Recall that $d_{\mathcal{G}_{\star}}$ denotes the metric on $\mathcal{G}_{\star}$. By continuity of $h(\cdot)$, for all $\varepsilon > 0$, there exists $\delta' > 0$ such that

$$d_{\mathcal{G}_{\star}}((G_1, o_1), (G_2, o_2)) < \delta' \implies |h(G_1, o_1) - h(G_2, o_2)| < \varepsilon.$$

Thus it is enough to show that, for all $\delta' > 0$, there exists $\delta > 0$, such that

$$d_{\mathcal{G}_\star}((G_1, o_1), (G_2, o_2)) < \delta \implies d_{\mathcal{G}_\star}((G_1(R, r), o_1), (G_2(R, r), o_2)) < \delta'. \qquad (2.1)$$

Let

$$\delta = \frac{\delta'}{1 - N\delta'}.$$

Then we claim that (2.1) holds. Here we have used the definition of metric space over rooted graphs along with the fact that for all $l \geq 0$

$$B_{N+l}^{(G_1)}(o_1) \cong B_{N+l}^{(G_2)}(o_2) \implies B_l^{(G_1(R,r))}(o_1) \cong B_l^{(G_2(R,r))}(o_2).$$

This is because the centrality measure $R(v)$ only depends on $B_N^{(G)}(v)$. This proves the lemma. $\square$

Proof of Theorem 1.1. Suppose that (G_n, o_n) converges locally weakly to $(G, o) \sim \mu$, which means that for all continuous and bounded functions $f \colon \mathcal{G}_\star \to \mathbb{R}$,

$$\lim_{n \to \infty} \mathbb{E}[f(G_n, o_n)] = \mathbb{E}[f(G, o)]. \qquad (2.2)$$

Using $f(\cdot)$ from Lemma 2.1 in (2.2), we get

$$\lim_{n \to \infty} \mathbb{E}[h(G_n(R, r), o_n)] = \lim_{n \to \infty} \mathbb{E}[f(G_n, o_n)] = \mathbb{E}[f(G, o)] = \mathbb{E}[h(G(R, r), o)].$$

Thus $(G_n(R, r), o_n)$ converges locally weakly to $(G(R, r), o)$. Similarly, suppose that (G_n, o_n) converges locally in probability to $(G, o) \sim \mu$, which means for all continuous and bounded function $f \colon \mathcal{G}_\star \to \mathbb{R}$,

$$\mathbb{E}[f(G_n, o_n) \mid G_n] \xrightarrow{\mathbb{P}} \mathbb{E}[f(G, o)]. \qquad (2.3)$$

Using $f(\cdot)$ from Lemma 2.1 in (2.3), we get

$$\mathbb{E}[h(G_n(R, r), o_n) \mid G_n] = \mathbb{E}[f(G_n, o_n) \mid G_n] \xrightarrow{\mathbb{P}} \mathbb{E}[f(G, o)] = \mathbb{E}[h(G(R, r), o)].$$

Thus $(G_n(R, r), o_n)$ converges locally in probability to $(G(R, r), o)$. $\square$

Proof of Corollaries 1.1 and 1.2. These proofs follow immediately from the upper bound on the giant given in [23, Corollary 2.28], and the convergence of the number of connected components in [23, Corollary 2.22]. $\square$

3. Proofs: Degree centrality and configuration model

In this section, we restrict ourselves to the configuration model and investigate all possible α-proportion vertex removals, which can be performed with respect to degree centrality, and then compare them on the basis of the size of the giant. For two sequences $(f(n))_{n \geq 1}$ and $(g(n))_{n \geq 1}$, we write $f(n) = o(g(n))$ when $\lim_{n \to \infty} f(n)/g(n) = 0$.

3.1. Giant in vertex-removed configuration model: Proof of Theorem 1.2

For the proof we use the vertex explosion construction as in Section 1.4.2. We apply the following steps on G_n.

(1) Choose $\lfloor n_k r_k^{(n)} \rfloor$ vertices uniformly from the set of vertices of degree k and explode them into k vertices of degree 1.

(2) Label these newly formed degree-1 vertices red.

(3) Repeat this for each $k \in \mathbb{N}$.

This newly constructed graph is $\tilde{G}(V')$ from Section 1.4.2 for an appropriately chosen V'. We shall denote it by $\tilde{G}_n$. Let $\tilde{n}_k$, $\tilde{n}$ and n_+ denote the number of degree-k vertices, total number of vertices, and number of vertices with a red label in the vertex-exploded graph $\tilde{G}_n$, respectively. Then

$$\tilde{n}_i = \begin{cases} n_+ + n_1 - \lfloor n_1 r_1^{(n)} \rfloor & \text{for } i = 1, \\ n_i - \lfloor n_i r_i^{(n)} \rfloor & \text{for } i \neq 1, \end{cases} \tag{3.1}$$

$$n_+ = \sum_{i=1}^{\infty} i \lfloor n_i r_i^{(n)} \rfloor = n \sum_{i=1}^{\infty} i \frac{\lfloor n_i r_i^{(n)} \rfloor}{n}, \tag{3.2}$$

$$\tilde{n} \sum_{i=1}^{\infty} \tilde{n}_k = n + \sum_{i=1}^{\infty} (i-1) \lfloor n_i r_i^{(n)} \rfloor = n \left(1 + \sum_{i=1}^{\infty} (i-1) \frac{\lfloor n_i r_i^{(n)} \rfloor}{n} \right). \tag{3.3}$$

Then, by applying the generalised dominated convergence theorem on (3.2) and (3.3),

$$n_+ = \mathbb{E}[Dr_D] n + o(n), \quad \tilde{n} = \beta_r n + o(n). \tag{3.4}$$

Recall that $\beta_r = \mathbb{E}[Dr_D] + 1 - \alpha$. Due to (3.1) and (3.4),

$$\frac{\tilde{n}_i}{n} = \begin{cases} \mathbb{E}[Dr_D] + p_1(1 - r_1) + o(1) & \text{for } i = 1, \\ p_i(1 - r_i) + o(1) & \text{for } i \neq 1. \end{cases} \tag{3.5}$$

We define

$$\tilde{p}_i := \begin{cases} \beta_r^{-1}(\mathbb{E}[Dr_D] + p_1(1 - r_1)) & \text{for } i = 1, \\ \beta_r^{-1} p_i(1 - r_i) & \text{for } i \neq 1. \end{cases}$$

Then, by (3.5), for every $i \geq 1$,

$$\frac{\tilde{n}_i}{\tilde{n}} \to \tilde{p}_i. \tag{3.6}$$

Let $\tilde{D}$ be a random variable with probability mass function $(\tilde{p}_j)_{j \geq 1}$.

By Lemma 1.1, $\tilde{G}_n$ is also a configuration model. Equation (3.6) shows that the regularity Condition 1.1(a) holds for $\tilde{d}$. Similarly, Condition 1.1(b) also holds. Using Theorem 1.5, the giant in $\tilde{G}_n$ exists if and only if $\mathbb{E}[\tilde{D}(\tilde{D} - 2)] > 0$. We compute

$$\mathbb{E}[\tilde{D}(\tilde{D} - 2)] = \frac{1}{\beta_r} \sum_{j \geq 1} j(j-2) p_j (1 - r_j) - \frac{1}{\beta_r} \mathbb{E}[Dr_D]$$

$$= \frac{1}{\beta_r} \mathbb{E}[D(D-2)(1 - r_D)] - \frac{1}{\beta_r} \mathbb{E}[Dr_D]$$

$$= \frac{1}{\beta_r} \mathbb{E}[D(D-1)(1 - r_D)] - \frac{1}{\beta_r} \mathbb{E}[D]. \tag{3.7}$$

Recall ν_r from (1.6). By (3.7),

$$\mathbb{E}[\tilde{D}(\tilde{D}-2)] > 0 \iff \nu_r > 1.$$

Thus $\tilde{G}_n$ has a giant if and only if $\nu_r > 1$. Recall that the removal of red vertices (having degree 1) from $\tilde{G}_n$ gives $G_{n,r^{(n)}}$. Since we remove a fraction $n_+/\tilde{n}_1$ of all degree-1 vertices, due to the law of large numbers for a hypergeometric distribution, we asymptotically remove the same fraction of degree-1 vertices in the giant component. In more detail, by Theorem 1.5(i(b)) and $j=1$, $\tilde{C}_1$ contains $\tilde{p}_1(1-\eta_r)\tilde{n}(1+o_\mathbb{P}(1))$ degree-1 vertices, where $\eta_r \in (0,1]$ satisfies $\beta_r g'_{\tilde{D}}(\eta_r) = \mathbb{E}[D]\eta_r$, and the proportion of red vertices is $n_+/\tilde{n}$. Hence the number of red vertices removed from $\tilde{C}_1$ is

$$\frac{n_+}{\tilde{n}_1} \times \tilde{p}_1(1-\eta_r)\tilde{n}(1+o_\mathbb{P}(1)) = n\mathbb{E}[Dr_D](1-\eta_r)(1+o_\mathbb{P}(1)).$$

This means that, as $n \to \infty$,

$$v(C_1(r^{(n)})) = v(\tilde{C}_1) - n(1-\eta_r)\mathbb{E}[Dr_D](1+o_\mathbb{P}(1)). \tag{3.8}$$

See Janson [27, (3.5)], where a similar argument is used for the computation of the proportion of vertices in the giant component after site percolation. By Theorem 1.5(i(a)),

$$\frac{v(\tilde{C}_1)}{\tilde{n}} \xrightarrow{\mathbb{P}} 1 - g_{\tilde{D}}(\eta_r). \tag{3.9}$$

Here we recall that $v(C_1(r))$ is the number of vertices in the largest component of $G_{n,r}$. We define $v(\tilde{C}_1)$ to be the number of vertices in the largest component of $\tilde{G}_n$ and $v_1(\tilde{C}_1)$ to be the number of degree-1 vertices in the largest component of $\tilde{G}_n$.

Due to (3.8), we conclude that there is a giant in $G_{n,r^{(n)}}$ if and only if there is a giant in $\tilde{G}_n$. Thus the giant exists if $\nu_r > 1$ and does not exist if $\nu_r \leq 1$. By (3.4), $\tilde{n}/n \to \beta_r$. This means that (3.9) and (3.8) give us

$$\frac{v(C_1(r^{(n)}))}{n} \xrightarrow{\mathbb{P}} \beta_r(1 - g_{\tilde{D}}(\eta_r)) - (1-\eta_r)\mathbb{E}[Dr_D]. \tag{3.10}$$

Let

$$\rho(r) := \beta_r(1 - g_{\tilde{D}}(\eta_r)) - (1-\eta_r)\mathbb{E}[Dr_D] = 1 - \alpha - \mathbb{E}[Dr_D]\eta_r - \beta_r g_{\tilde{D}}(\eta_r).$$

By the definition of a generating function,

$$\beta_r g_{\tilde{D}}(\eta_r) = \sum_{j \geq 1} \tilde{p}_j \eta_r^j = \sum_{j \geq 1} p_i(1-r_i)\eta_r^i + \mathbb{E}[Dr_D]\eta_r.$$

Thus

$$\rho(r) = 1 - \alpha - 2\mathbb{E}[Dr_D]\eta_r - \sum_{i \geq 1} p_i(1-r_i)\eta_r^i.$$

Removing a red vertex removes exactly one edge. Hence the number of edges in the giant component of the vertex-removed graph is the number of edges in the vertex-exploded graph

minus the number of red vertices in the giant component of the vertex-exploded graph. Now, using the law of large numbers, as in (3.8), we get

$$e(C_1(\boldsymbol{r}^{(n)})) = e(\tilde{C}_1) - v_1(\tilde{C}_1) \times \frac{n_+}{\tilde{n}_1}(1 + o_{\mathbb{P}}(n)). \tag{3.11}$$

Dividing by n and taking limits in probability on both sides of (3.11) completes the proof of Theorem 1.2.

Proof of Corollary 1.3. Let $k^{(n)}$ be the top α-quantile for the degree distribution D_n, which has probability mass function $(n_j/n)_{j \geq 1}$. Define $\bar{\boldsymbol{r}}^{(n)}(\alpha)$ to be

$$\bar{\boldsymbol{r}}^{(n)}(\alpha) = \left(0, 0, 0, \ldots 0, \frac{n(\alpha - \mathbb{P}(D_n > k^{(n)}))}{n_{k^{(n)}}}, 1, 1, \ldots\right).$$

Due to Theorem 1.2, it is enough to show that $(\bar{\boldsymbol{r}}^{(n)}(\alpha))_{n \geq 1} \in S(\bar{\boldsymbol{r}}(\alpha))$. This follows by Condition 1.1. The proof for the bottom α-quantile is identical. This completes the proof. ☐

3.2. Preliminary bounds

In this section, we prove bounds on the size of the giant in Theorem 1.3. We also discuss a bound on $\eta_{\boldsymbol{r}}$, interpreted as the half-edge extinction probability. Recall from (1.9) that $\eta_{\boldsymbol{r}}$ is the smallest solution in $(0,1]$ of (1.9).

Lemma 3.1. (Lower bound for $\eta_{\boldsymbol{r}}$.) *For every α-sequence $\boldsymbol{r}$,*

$$\eta_{\boldsymbol{r}} \geq \frac{\mathbb{E}[Dr_D]}{\mathbb{E}[D]}.$$

Proof. By (1.9) and (1.10),

$$\sum_{j \geq 1} jp_j(1 - r_j)\eta_{\boldsymbol{r}}^{j-1} + \mathbb{E}[Dr_D] = \eta_{\boldsymbol{r}}\mathbb{E}[D],$$

which implies that

$$\mathbb{E}[D(1 - r_D)\eta_{\boldsymbol{r}}^{D-1}] = \mathbb{E}[(\eta_{\boldsymbol{r}} - r_D)D].$$

In the above equality, obviously the left-hand side is non-negative. Thus the right-hand side is also non-negative, as required. ☐

Remark 3.1. (*Lower bound for η dependent only on α.*) The lower bound given above is non-trivial, since using $D \geq 1$ a.s. gives

$$\eta_{\boldsymbol{r}} \geq \frac{\mathbb{E}[Dr_D]}{\mathbb{E}[D]} \geq \frac{\alpha}{\mathbb{E}[D]} > 0.$$

This lower bound is independent of the α-sequence used.

Proof of Theorem 1.3. By (1.7), Lemma 3.1 and Remark 3.1,

$$\rho(\boldsymbol{r}) = 1 - \alpha - 2\mathbb{E}[Dr_D]\eta_{\boldsymbol{r}} - \sum_{j \geq 1} p_i(1 - r_i)\eta_{\boldsymbol{r}}^i$$

$$\leq 1 - \alpha - 2\mathbb{E}[Dr_D]\eta_{\boldsymbol{r}} \leq 1 - \alpha - \frac{2(\mathbb{E}[Dr_D])^2}{\mathbb{E}[D]}. \tag{3.12}$$

Since $D \geq 1$ almost surely, we have $\mathbb{E}[Dr_D] \geq \mathbb{E}[r_D] = \alpha$. Using this fact in (3.12) proves (1.12). If the α-sequence, $\boldsymbol{r}$, is positively correlated with the degree sequence, we have $\mathbb{E}[Dr_D] \geq \alpha \mathbb{E}[D]$. Using this fact in (3.12) proves (1.13). The mean-value theorem implies that there exists $\zeta \in (\eta_{\boldsymbol{r}}, 1)$ such that $1 - g_{\boldsymbol{r}}(\eta_{\boldsymbol{r}}) = g_{\boldsymbol{r}}'(\zeta)(1 - \eta_{\boldsymbol{r}})$. Thus, also using (3.10), we obtain

$$\rho(\boldsymbol{r}) = \beta_{\boldsymbol{r}}(1 - g_{\boldsymbol{r}}(\eta_{\boldsymbol{r}})) - \mathbb{E}[Dr_D](1 - \eta_{\boldsymbol{r}}) = (\beta_{\boldsymbol{r}} g_{\boldsymbol{r}}'(\zeta) - \mathbb{E}[Dr_D])(1 - \eta_{\boldsymbol{r}}). \qquad (3.13)$$

Since $g_{\boldsymbol{r}}(\cdot)$ is a generating function of a non-negative random variable, $g_{\boldsymbol{r}}'(\cdot)$ is an increasing function, so that

$$\frac{\mathbb{E}[D]}{\beta_{\boldsymbol{r}}} \eta_{\boldsymbol{r}} = g_{\boldsymbol{r}}'(\eta_{\boldsymbol{r}}) \leq g_{\boldsymbol{r}}'(\zeta) \leq g_{\boldsymbol{r}}'(1) = \frac{\mathbb{E}[D]}{\beta_{\boldsymbol{r}}}.$$

Substituting this inequality in (3.13), we get

$$\mathbb{E}[D(\eta_{\boldsymbol{r}} - r_D)](1 - \eta_{\boldsymbol{r}}) \leq \rho(\boldsymbol{r}) \leq \mathbb{E}[D(1 - r_D)](1 - \eta_{\boldsymbol{r}}). \qquad (3.14)$$

Using the lower bound on $\eta_{\boldsymbol{r}}$ (Lemma 3.1) in (3.14), we obtain a bound for $\rho(\boldsymbol{r})$, in (1.11). Next, we obtain a bound for $e(\boldsymbol{r})$. By Theorem 1.2,

$$e(\boldsymbol{r}) = \left(\frac{\mathbb{E}[D]}{2}(1 + \eta_{\boldsymbol{r}}) - \mathbb{E}[Dr_D] \right)(1 - \eta_{\boldsymbol{r}}) = -\frac{\mathbb{E}[D]}{2}\eta_{\boldsymbol{r}}^2 + \mathbb{E}[Dr_D]\eta_{\boldsymbol{r}} + \frac{\mathbb{E}[D]}{2} - \mathbb{E}[Dr_D].$$

Define the polynomial $P(\cdot)$ by

$$P(x) := -\frac{\mathbb{E}[D]}{2}x^2 + \mathbb{E}[Dr_D]x + \frac{\mathbb{E}[D]}{2} - \mathbb{E}[Dr_D].$$

Note that $P(\cdot)$ is a quadratic polynomial with a negative leading coefficient. Thus it is maximal at $x^{\star} = \mathbb{E}[Dr_D]/\mathbb{E}[D] \in (0, 1)$. By Lemma 3.1,

$$\eta_{\boldsymbol{r}} \in \left[\frac{\mathbb{E}[Dr_D]}{\mathbb{E}[D]}, 1 \right].$$

Thus

$$0 = P(1) < e(\boldsymbol{r}) = P(\eta_{\boldsymbol{r}}) \leq P(x^{\star}).$$

This gives (1.11), which completes the proof. □

3.3. ε-transformations and half-edge extinction probability

In this section, we study the effect of a small perturbation in the α-sequence on the size of the giant after vertex removal. For this we define the notion of an 'ε-transformation' that moves mass to the left.

Definition 3.1. (ε-*transformation.*) Suppose $\boldsymbol{r} = (r_j)_{j \geq 1}$ is an α-sequence. Fix $k, l \in \mathbb{N}$. We call $\boldsymbol{r}^{k,l}(\varepsilon) = (r_j^{k,l}(\varepsilon))_{j \geq 1}$ an ε-transformation on the co-ordinates $(k, k + l)$ if it satisfies $r_i^{k,l}(\varepsilon) = r_i$ for $i \notin \{k, k + l\}$, while

$$r_k^{k,l}(\varepsilon) = r_k + \frac{\varepsilon}{p_k} \quad \text{and} \quad r_{k+l}^{k,l}(\varepsilon) = r_{k+l} - \frac{\varepsilon}{p_{k+l}}.$$

Note that $\boldsymbol{r}^{k,l}(\varepsilon)$ is also an α-sequence for $\varepsilon \leq p_{k+l}r_{k+l}$. Thus the domain of ε is $[0, p_{k+l}r_{k+l}]$.

Let $\rho^{k,l}(\varepsilon)$ denote the limiting proportion of vertices in the giant after the vertex removal procedure according to $\boldsymbol{r}^{k,l}(\varepsilon)$.

Remark 3.2. (*Supercritical regime*.) Throughout this section we assume that we start with the supercritical regime, i.e. $\nu_{\boldsymbol{r}} > 1$; recall $\nu_{\boldsymbol{r}}$ from (1.6).

Remark 3.3. (*Graph remains supercritical after ε-transformation*.) Starting in the supercritical regime, the sequence of random graphs remains in the supercritical regime even after the ε-transformation. In other words, $\nu_{\boldsymbol{r}} > 1 \implies \nu_{\boldsymbol{r}^{k,l}(\varepsilon)} > 1$. Indeed, we can compute that

$$\nu_{\boldsymbol{r}^{k,l}(\varepsilon)} = \frac{\mathbb{E}[D(D-1)(1 - r_D^{k,l}(\varepsilon))]}{\mathbb{E}[D]}$$

$$= \frac{\mathbb{E}[D(D-1)(1 - r_D)]}{\mathbb{E}[D]} + \frac{k(k-1)(-\varepsilon) + (k+l)(k+l-1)\varepsilon}{\mathbb{E}[D]}$$

$$> \nu_{\boldsymbol{r}}.$$

3.3.1. *Monotonicity of the half-edge extinction probability.* In this section, we investigate the effect of the ε-transformation on the half-edge extinction probability of the configuration model after vertex explosion. Let $\tilde{D}$ with probability mass function $(\tilde{p}_j)_{j\geq 1}$ be as described in Section 3.1. $\tilde{D}$ was computed for the α-sequence $\boldsymbol{r}$. Let $\tilde{D}_\varepsilon$ be the corresponding $\tilde{D}$ for the α-sequence $\boldsymbol{r}^{k,l}(\varepsilon)$. Let $g_\varepsilon(\cdot)$ and $g(\cdot)$ be the generating functions of $\tilde{D}_\varepsilon$ and $\tilde{D}$, respectively. Let $\eta_\varepsilon, \eta \in (0, 1)$ satisfy

$$g_\varepsilon'(\eta_\varepsilon) = \frac{\mathbb{E}[D]}{\beta_\varepsilon}\eta_\varepsilon \quad \text{and} \quad g'(\eta) = \frac{\mathbb{E}[D]}{\beta}\eta,$$

where

$$\beta_\varepsilon = \mathbb{E}[Dr_D^{k,l}(\varepsilon)] + 1 - \alpha \quad \text{and} \quad \beta = \mathbb{E}[Dr_D] + 1 - \alpha.$$

Then η and η_ε are the half-edge extinction probabilities in the configuration model with limiting degree distributions $\tilde{D}$ and $\tilde{D}_\varepsilon$, respectively. By the definition of $r_D^{k,l}(\varepsilon)$,

$$\mathbb{E}[Dr_D^{k,l}(\varepsilon)] = \mathbb{E}[Dr_D] - l\varepsilon.$$

This means that $\beta_\varepsilon = \beta - l\varepsilon$. We can also relate g_ε and g by noting that

$$g_\varepsilon(s) = \mathbb{E}[s^{\tilde{D}_\varepsilon}] = \beta_\varepsilon^{-1}\left(\sum_{j\geq 1} p_j(1 - r_j^{k,l}(\varepsilon))s^j + \mathbb{E}[Dr_D^{k,l}(\varepsilon)]s\right),$$

so that

$$\beta_\varepsilon g_\varepsilon(s) = \sum_{j\geq 1} p_j(1 - r_j)s^j + \mathbb{E}[Dr_D]s - l\varepsilon s + \sum_{j\in\{k,k+l\}} p_j(r_j - r_j^{k,l}(\varepsilon))s^j.$$

In turn, this means that

$$\beta_\varepsilon g_\varepsilon(s) = \beta g(s) - \varepsilon(s^k - s^{k+l} + ls). \tag{3.15}$$

Recall that the proportion of vertices in the giant component of the graph obtained after vertex removal with respect to $\boldsymbol{r}^{k,l}(\varepsilon)$ is $\rho^{k,l}(\varepsilon)$. By (3.10),

$$\rho^{k,l}(\varepsilon) = 1 - \alpha + (\mathbb{E}[Dr_D] - l\varepsilon)\eta_\varepsilon - \beta_\varepsilon g_\varepsilon(\eta_\varepsilon). \tag{3.16}$$

Substituting (3.15) into (3.16) gives

$$\rho^{k,l}(\varepsilon) = 1 - \alpha - \beta g(\eta_\varepsilon) + \varepsilon\left(\eta_\varepsilon^k - \eta_\varepsilon^{k+l}\right) + \mathbb{E}[Dr_D]\eta_\varepsilon. \tag{3.17}$$

The following lemma shows that η_ε is well-defined for ε such that $v_{\boldsymbol{r}^{k,l}(\varepsilon)} > 1$.

Lemma 3.2. (Uniqueness.) *Let $g(\cdot)$ be the generating function for D, let $p_1 > 0$, and let $\eta \in (0, 1]$ satisfy*

$$g'(\eta) = \mathbb{E}[D]\eta. \tag{3.18}$$

Then η exists and is unique if $v = \mathbb{E}[D(D-1)]/\mathbb{E}[D] > 1$. If $v \le 1$, then $\eta = 1$ is the only solution to (3.18).

The above lemma follows from Janson and Luczak [30, Theorem 2.3]. For $p_1 = 0$, we define $\eta = 0$ instead. Then η can be interpreted as the extinction probability of a branching process with offspring distribution $p_k^\star = (k+1)p_k/\mathbb{E}[D]$. The next lemma is needed in the rest of this section.

Lemma 3.3. (Negative polynomial.) *For $s \in (0, 1)$,*

$$(k+l)s^{k+l-1} - ks^{k-1} - l < 0.$$

Proof. Let $P(x) := (k+l)x^{k+l-1} - kx^{k-1} - l$. Then

$$P'(x) = (k+l)(k+l-1)x^{k+l-2} - k(k-1)x^{k-2}.$$

This means that

$$P'(x) > 0 \iff x^l > \frac{k(k-1)}{(k+l)(k+l-1)}.$$

The only minimiser of $P(\cdot)$ in (0,1), denoted by γ, equals

$$\gamma = \left(\frac{k(k-1)}{(k+l)(k+l-1)}\right)^{1/l}.$$

Since there is no maximum in the interval $(0, 1)$, the maximum occurs either in 0 or 1. We have $P(0) = -l$ and $P(1) = 0$, so that $P(x) < 0$ for all $x \in (0, 1)$. $\square$

In the next lemma, we compute the derivative of η_ε with respect to ε, and show that it is negative.

Lemma 3.4. (Half-edge extinction probability is decreasing in ε.) *Suppose $v_{\boldsymbol{r}} > 1$. For all $\varepsilon \in [0, p_{k+l}r_{k+l}]$,*

$$\frac{\partial \eta_\varepsilon}{\partial \varepsilon} = \frac{(k+l)\eta_\varepsilon^{k+l-1} - k\eta_\varepsilon^{k-1} - l}{\mathbb{E}[D] - \beta_\varepsilon g_\varepsilon''(\eta_\varepsilon)} < 0.$$

Proof. Differentiating (3.15) gives

$$\beta_\varepsilon g_\varepsilon'(s) = \beta g'(s) - \varepsilon\left(ks^{k-1} - (k+l)s^{k+l-1} + l\right).$$

By definition of η_ε, we have $\beta_\varepsilon g_\varepsilon'(\eta_\varepsilon) = \mathbb{E}[D]\eta_\varepsilon$. Thus

$$\beta g'(\eta_\varepsilon) - \varepsilon\left(k\eta_\varepsilon^{k-1} - (k+l)\eta_\varepsilon^{k+l-1} + l\right) = \mathbb{E}[D]\eta_\varepsilon. \tag{3.19}$$

Define $h\colon \mathbb{R}^2 \to \mathbb{R}$, as

$$h(x, y) = \beta g'(x) - y\big(kx^{k-1} - (k+l)x^{k+l-1} + l\big) - \mathbb{E}[D]x.$$

Due to Lemma 3.3, for $x \in [0, 1)$, we have

$$\frac{\partial h(x, y)}{\partial y} = (k+l)x^{k+l-1} - kx^{k-1} - l < 0.$$

Also, $h(\eta_\varepsilon, \varepsilon) = 0$ for every $\varepsilon > 0$. The implicit function theorem [40, Theorem 9.28] on $h\colon \mathbb{R}^2 \to \mathbb{R}$ implies that $\varepsilon \mapsto \eta_\varepsilon$ is differentiable. Thus we differentiate (3.19) with respect to ε, to obtain

$$\mathbb{E}[D]\frac{\partial \eta_\varepsilon}{\partial \varepsilon} = \beta g''(\eta_\varepsilon)\frac{\partial \eta_\varepsilon}{\partial \varepsilon} - \big(k\eta_\varepsilon^{k-1} - (k+l)\eta_\varepsilon^{k+l-1} + l\big)$$

$$- \varepsilon\big(k(k-1)\eta_\varepsilon^{k-2} - (k+l)(k+l-1)\eta_\varepsilon^{k+l-2}\big)\frac{\partial \eta_\varepsilon}{\partial \varepsilon},$$

which gives

$$\frac{\partial \eta_\varepsilon}{\partial \varepsilon} = \frac{(k+l)\eta_\varepsilon^{k+l-1} - k\eta_\varepsilon^{k-1} - l}{\mathbb{E}[D] - \beta g''(\eta_\varepsilon) + \varepsilon\big(k(k-1)\eta_\varepsilon^{k-2} - (k+l)(k+l-1)\eta_\varepsilon^{k+l-2}\big)}. \tag{3.20}$$

Differentiating (3.15) twice with respect to s, and putting $s = \eta_\varepsilon$, we obtain

$$\beta_\varepsilon g_\varepsilon''(\eta_\varepsilon) = \beta g''(\eta_\varepsilon) - \varepsilon\big(k(k-1)\eta_\varepsilon^{k-2} - (k+l)(k+l-1)\eta_\varepsilon^{k+l-2}\big).$$

This means that

$$\beta g''(\eta_\varepsilon) = \beta_\varepsilon g_\varepsilon''(\eta_\varepsilon) + \varepsilon\big(k(k-1)\eta_\varepsilon^{k-2} - (k+l)(k+l-1)\eta_\varepsilon^{k+l-2}\big). \tag{3.21}$$

Substituting (3.21) in (3.20), we get

$$\frac{\partial \eta_\varepsilon}{\partial \varepsilon} = \frac{(k+l)\eta_\varepsilon^{k+l-1} - k\eta_\varepsilon^{k-1} - l}{\mathbb{E}[D] - \beta_\varepsilon g_\varepsilon''(\eta_\varepsilon)}, \tag{3.22}$$

as required. We next show that $\partial \eta_\varepsilon / \partial \varepsilon < 0$. Let $h(x) = \beta_\varepsilon g_\varepsilon'(x) - \mathbb{E}[D]x$. Notice that $h(\eta_\varepsilon) = h(1) = 0$. By Rolle's theorem, there exists $\phi \in (\eta_\varepsilon, 1)$ such that $h'(\phi) = 0$. This implies that

$$g_\varepsilon''(\phi) = \frac{\mathbb{E}[D]}{\beta_\varepsilon}.$$

Since $g_\varepsilon(\cdot)$ is a generating function, $g_\varepsilon''(\cdot)$ is increasing, so that

$$\beta_\varepsilon g_\varepsilon''(\eta_\varepsilon) < \beta_\varepsilon g_\varepsilon''(\phi) = \mathbb{E}[D].$$

Thus

$$\mathbb{E}[D] - \beta_\varepsilon g_\varepsilon''(\eta) > 0,$$

which means that (3.22) is a well-defined expression, which means that η_ε is differentiable and the derivative is given by (3.22). By Lemma 3.3, the numerator in (3.22) is negative, which completes the proof. $\qquad\square$

Lemma 3.4 immediately implies that $\varepsilon \mapsto \partial \eta_\varepsilon / \partial \varepsilon$ is continuous.

Corollary 3.1. (Continuity of the derivative.) *Suppose $v_r > 1$. Then $\partial \eta_\varepsilon / \partial \varepsilon$ is a continuous function.*

We next show that the derivative of the half-edge extinction probability is bounded, which is also a consequence of Lemma 3.4.

Corollary 3.2. (Bound on the derivative.) *Suppose $v_r > 1$. There exists $\psi_{k,l} < 0$ such that $\psi_{k,l} < \partial \eta_\varepsilon / \partial \varepsilon < 0$ for all $\varepsilon \in [0, p_{k+l} r_{k+l}]$.*

Proof. Let $h(\varepsilon) = \partial \eta_\varepsilon / \partial \varepsilon$. Notice that ε can only take values in $[0, p_{k+l} r_{k+l}]$, by Definition 3.1. This means that the domain of function $h(\cdot)$ is compact. Further, by Corollary 3.1, $h(\cdot)$ is continuous. This means that $h(\cdot)$ is bounded below. By Lemma 3.4 we already know that $h(\varepsilon) < 0$ for all ε. Thus there exists $\psi_{k,l} < 0$ such that $\psi_{k,l} < h(\varepsilon) < 0$ for all $\varepsilon \in [0, p_{k+l} r_{k+l}]$. □

3.4. Derivative of vertex/edge limiting proportion in the giant of $G_{n,r}(\varepsilon)$

In this section, we compute the derivative of the limiting proportion of vertices $\varepsilon \mapsto \rho^{k,l}(\varepsilon)$ and edges $\varepsilon \mapsto e^{k,l}(\varepsilon)$ in the giant component in $G_{n,r}(\varepsilon)$, and show that these derivatives are positive. In the next proposition, we compute $\partial e^{k,l}(\varepsilon)/\partial \varepsilon$ and its sign.

Proposition 3.1. (Monotonicity of edge proportion.) *Suppose $v_r > 1$. For all $\varepsilon \in [0, p_{k+l} r_{k+l}]$, and α-sequence r,*

$$\frac{\partial e^{k,l}(\varepsilon)}{\partial \varepsilon} = -\frac{\partial \eta_\varepsilon}{\partial \varepsilon} \mathbb{E}\big[D(\eta_\varepsilon - r_D^{k,l}(\varepsilon))\big] + l(1 - \eta_\varepsilon) > 0.$$

Proof. By Theorem 1.2,

$$e^{k,l}(\varepsilon) = \frac{\mathbb{E}[D]}{2}(1 - \eta_\varepsilon^2) - (\mathbb{E}[Dr_D] - l\varepsilon)(1 - \eta_\varepsilon). \tag{3.23}$$

Differentiating (3.23) with respect to ε, we obtain

$$\frac{\partial e^{k,l}(\varepsilon)}{\partial \varepsilon} = -\frac{\partial \eta_\varepsilon}{\partial \varepsilon}(\mathbb{E}[D(\eta_\varepsilon - r_D)] + l\varepsilon) + l(1 - \eta_\varepsilon)$$

$$= -\frac{\partial \eta_\varepsilon}{\partial \varepsilon} \mathbb{E}\big[D(\eta_\varepsilon - r_D^{k,l}(\varepsilon))\big] + l(1 - \eta_\varepsilon),$$

as required. By Lemma 3.4, $\partial \eta_\varepsilon / \partial \varepsilon < 0$. By Lemma 3.1, for the α-sequence $r^{k,l}(\varepsilon)$,

$$\eta_\varepsilon \geq \frac{\mathbb{E}[Dr_D^{k,l}(\varepsilon)]}{\mathbb{E}[D]} \implies \mathbb{E}\big[D(\eta_\varepsilon - r_D^{k,l}(\varepsilon))\big] > 0.$$

This completes the proof. □

In the next proposition, we compute $\partial \rho^{k,l}(\varepsilon)/\partial \varepsilon$.

Proposition 3.2. (Derivative of vertex proportion.) *Suppose $v_r > 1$. For all $\varepsilon \in [0, p_{k+l} r_{k+l}]$, and α-sequence r,*

$$\frac{\partial \rho^{k,l}(\varepsilon)}{\partial \varepsilon} = -A_\varepsilon \frac{\partial \eta_\varepsilon}{\partial \varepsilon} + B_\varepsilon,$$

where

$$A_\varepsilon = \mathbb{E}\big[D\big(\eta_\varepsilon - r_D^{k,l}(\varepsilon)\big)\big] + \big(k\eta_\varepsilon^{k-1} - (k+l)\eta_\varepsilon^{k+l-1}\big)\varepsilon, \tag{3.24}$$

$$B_\varepsilon = \eta_\varepsilon^k(1 - \eta_\varepsilon^l) + \big(k\eta_\varepsilon^{k-1} - (k+l)\eta_\varepsilon^{k+l-1}\big)\varepsilon. \tag{3.25}$$

Consequently,

$$\frac{\partial \rho^{k,l}(\varepsilon)}{\partial \varepsilon}\bigg|_{\varepsilon=0} = -\mathbb{E}[D(\eta - r_D)]\frac{\partial \eta_\varepsilon}{\partial \varepsilon}\bigg|_{\varepsilon=0} + \eta^k(1 - \eta^l) > 0.$$

Proof. Taking a derivative of (3.17) with respect to ε gives

$$\frac{\partial \rho^{k,l}(\varepsilon)}{\partial \varepsilon} = -\beta g'(\eta_\varepsilon)\frac{\partial \eta_\varepsilon}{\partial \varepsilon} + \varepsilon\big(k\eta_\varepsilon^{k-1} - (k+l)\eta_\varepsilon^{k+l-1}\big) + \big(\eta_\varepsilon^k - \eta_\varepsilon^{k+l}\big) + \mathbb{E}[Dr_D]\frac{\partial \eta_\varepsilon}{\partial \varepsilon}$$

$$= (\mathbb{E}[Dr_D] - \beta g'(\eta_\varepsilon))\frac{\partial \eta_\varepsilon}{\partial \varepsilon} + \eta_\varepsilon^k(1 - \eta_\varepsilon^l) + \varepsilon\big(k\eta_\varepsilon^{k-1} - (k+l)\eta_\varepsilon^{k+l-1}\big).$$

Thus

$$\frac{\partial \rho^{k,l}(\varepsilon)}{\partial \varepsilon} = (\mathbb{E}[Dr_D] - \beta g'(\eta_\varepsilon))\frac{\partial \eta_\varepsilon}{\partial \varepsilon} + B_\varepsilon. \tag{3.26}$$

By (3.19), we have

$$\beta g'(\eta_\varepsilon) = \mathbb{E}[D]\eta_\varepsilon + \varepsilon\big(k\eta_\varepsilon^{k-1} - (k+l)\eta_\varepsilon^{k+l-1} + l\big)$$

$$= A_\varepsilon - \mathbb{E}\big[Dr_D^{k,l}(\varepsilon)\big] + l\varepsilon = A_\varepsilon - \mathbb{E}[Dr_D].$$

Substituting this value in (3.26) proves the claim. □

Proposition 3.2 implies that $\partial \rho^{k,l}(\varepsilon)/\partial \varepsilon\big|_{\varepsilon=0} > 0$. We next extend this to all ε.

Proposition 3.3. (Monotonicity of vertex proportion.) *Suppose $v_r > 1$. Then, for all $\varepsilon \in [0, p_{k+l}r_{k+l}]$,*

$$\frac{\partial \rho^{k,l}(\varepsilon)}{\partial \varepsilon} > 0.$$

Proof. By Lemma 3.1, we have

$$\mathbb{E}\big[D\big(\eta_\varepsilon - r_D^{k,l}(\varepsilon)\big)\big] > 0 \implies A_\varepsilon > \big(k\eta_\varepsilon^{k-1} - (k+l)\eta_\varepsilon^{k+l-1}\big)\varepsilon, \tag{3.27}$$

where A_ε is as in (3.24). By Lemma 3.4, $\partial \eta_\varepsilon/\partial \varepsilon < 0$. Thus, by Proposition 3.2 and (3.27),

$$\frac{\partial \rho^{k,l}(\varepsilon)}{\partial \varepsilon} > \left(1 - \frac{\partial \eta_\varepsilon}{\partial \varepsilon}\right)\big(k\eta_\varepsilon^{k-1} - (k+l)\eta_\varepsilon^{k+l-1}\big)\varepsilon + \eta_\varepsilon^k(1 - \eta_\varepsilon^l).$$

We have $\eta_\varepsilon \in (0, 1)$. By Lemma 3.3, $k\eta_\varepsilon^{k-1} - (k+l)\eta_\varepsilon^{k+l-1} > -l$. This means

$$\frac{\partial \rho^{k,l}(\varepsilon)}{\partial \varepsilon} > \eta_\varepsilon^k(1 - \eta_\varepsilon^l) - l\varepsilon\left(1 - \frac{\partial \eta_\varepsilon}{\partial \varepsilon}\right).$$

Thus

$$\varepsilon < \frac{\eta_\varepsilon^k(1 - \eta_\varepsilon^l)}{l(1 - \partial\eta_\varepsilon/\partial\varepsilon)} \implies \frac{\partial\rho^{k,l}(\varepsilon)}{\partial\varepsilon} > 0. \tag{3.28}$$

By Lemma 3.4 and Remark 3.1,

$$\frac{\alpha}{\mathbb{E}[D]} \leq \eta_\varepsilon \leq \eta < 1.$$

By Corollary 3.2, we know that there exists $\psi_{k,l} < 0$, such that $\psi_{k,l} < \partial\eta_\varepsilon/\partial\varepsilon < 0$. Define $\tau(\boldsymbol{r})$ as

$$\tau(\boldsymbol{r}) := \frac{\alpha^k(1 - \eta)}{\mathbb{E}[D]^k l(1 - \psi_{k,l})} < \frac{\eta_\varepsilon^k(1 - \eta_\varepsilon^l)}{l(1 - \partial\eta_\varepsilon/\partial\varepsilon)}.$$

Thus $\rho^{k,l}(\varepsilon)$ has a positive derivative in $(0, \tau(\boldsymbol{r}))$, because of (3.28). If one iterates this on the α-sequence $\boldsymbol{r}' = \boldsymbol{r}^{k,l}(\tau(\boldsymbol{r}))$, we find that $\rho^{k,l}(\varepsilon)$ has a positive derivative in $(0, \tau(\boldsymbol{r}) + \tau(\boldsymbol{r}'))$. Next, we show $\tau(\boldsymbol{r}) \leq \tau(\boldsymbol{r}')$. For that we note the following facts.

(1) Both $\boldsymbol{r}$ and $\boldsymbol{r}'$ have the same α, while $\mathbb{E}[D]$ is a constant.

(2) We are using the same k, l for both ε-transformations $\boldsymbol{r}$ and $\boldsymbol{r}'$, which means that the exponents in the expression of $\tau(\cdot)$ are the same.

(3) $\psi_{k,l}$ depends only on k and l, and is thus the same for both cases.

This means that $\tau(\boldsymbol{r})$ and $\tau(\boldsymbol{r}')$ differ only because of their dependence on $\eta = \eta(\boldsymbol{r})$ and $\eta = \eta(\boldsymbol{r}')$. Further, $\varepsilon \mapsto \eta_\varepsilon$ is a decreasing function. Thus $\tau(\boldsymbol{r}') \geq \tau(\boldsymbol{r})$. As a result, $\rho^{k,l}(\varepsilon)$ has a positive derivative in the neighbourhood $(0, 2\tau(\boldsymbol{r}))$. Iterate this to cover the full domain of the function. $\qquad \square$

3.4.1. *Proof of main result.* In this section, we prove Theorem 1.4. We start by showing that we can go from $\boldsymbol{r}$ to $\boldsymbol{r}'$ by using ε-transformations when $\boldsymbol{r}' \preccurlyeq_p \boldsymbol{r}$.

Lemma 3.5. (*Stochastic ordering and ε-transformations.*) *Let $\boldsymbol{r} = (r_j)_{j \geq 1}$ and $\boldsymbol{r}' = (r_j')_{j \geq 1}$ be two α-sequences such that $\boldsymbol{r}' \preccurlyeq_p \boldsymbol{r}$. Then $\boldsymbol{r}'$ can be obtained from $\boldsymbol{r}$ through a series of ε-transformations.*

The proof of Lemma 3.5, which is straightforward yet technically involved, is deferred to Appendix B. With these tools, we can now prove our main result in Theorem 1.4.

Proof of Theorem 1.4. Recall $\bar{r}(\alpha)$ and $\underline{r}(\alpha)$ from Definition 1.9. Let $\boldsymbol{r} = (r_i)_{i \geq 1}$ be an arbitrary α-sequence. Then, by definition of stochastic dominance, $\underline{r}(\alpha) \preccurlyeq_p \boldsymbol{r} \preccurlyeq_p \bar{r}(\alpha)$. By Propositions 3.1 and 3.3, the proportion of edges and vertices in the giant increases after each ε-transformation. Thus we get the desired result from Lemma 3.5. $\qquad \square$

We next prove Corollary 1.4, which relies on the following lemma that allows us to compare α-sequences with different α.

Lemma 3.6. (*Stochastic ordering.*) *Fix $\varepsilon > 0$. Suppose $\boldsymbol{r}$ is an α-sequence and $\boldsymbol{r}'$ is an $(\alpha + \varepsilon)$-sequence such that $\boldsymbol{r} \preccurlyeq_p \boldsymbol{r}'$. Then there exists an ε-sequence, say $\boldsymbol{\delta} = (\delta_i)_{i \geq 1}$, satisfying $\boldsymbol{r} + \boldsymbol{\delta} \preccurlyeq_p \boldsymbol{r}'$.*

The proof of Lemma 3.6 is again deferred to Appendix B.

Proof of Corollary 1.4. We know that $\mathbb{E}[r'_D] \leq \mathbb{E}[r_D]$, since $\boldsymbol{r} \preccurlyeq_{\boldsymbol{p}} \boldsymbol{r}'$. Furthermore, if $\mathbb{E}[r'_D] = \mathbb{E}[r_D]$, then the result follows due to Theorem 1.4. Thus, suppose $\mathbb{E}[r'_D] < \mathbb{E}[r_D]$, which means that there exists $\varepsilon > 0$ such that $\mathbb{E}[r_D] = \mathbb{E}[r'_D] + \varepsilon$. By Lemma 3.6, there exists an ε-sequence, say $\boldsymbol{\delta}$, such that $\boldsymbol{r} + \boldsymbol{\delta} \preccurlyeq_{\boldsymbol{p}} \boldsymbol{r}'$. Notice that both $\boldsymbol{r} + \boldsymbol{\delta}$ and $\boldsymbol{r}'$ are α-sequences. Thus, from Theorem 1.4, we have $\rho(\boldsymbol{r}') \leq \rho(\boldsymbol{r} + \boldsymbol{\delta})$. Also, by the definition of vertex removal according to an α-sequence, it follows that $\rho(\boldsymbol{r} + \boldsymbol{\delta}) \leq \rho(\boldsymbol{r})$, because we are removing $\boldsymbol{\delta}$ proportion more vertices after $\boldsymbol{r}$-removal in $(\boldsymbol{r} + \boldsymbol{\delta})$-removal. Thus we get the desired result. $\qquad\square$

Appendix A. Stochastic ordering and giant in configuration model

In this section, we prove Theorem 1.6. For that we first show Lemma A.1, which essentially proves Theorem 1.6 for a small subcase. Lemma A.2 finishes the proof by showing that this was sufficient. First we define ε-transformations for probability measures.

Definition A.1. (ε-*transformations for probability measures.*) Let $\boldsymbol{p} = (p_i)_{i \geq 0}$ be a probability mass function on the set of non-negative integers. For $\varepsilon \leq p_k$, define

$$\boldsymbol{p}(\varepsilon) := (p_1, p_2, \ldots, p_{k-1}, p_k - \varepsilon, p_{k+1}, \ldots, p_{k+l-1}, p_{k+l} + \varepsilon, p_{k+l+1}, \ldots).$$

Before moving to our first lemma, we recall that $\rho_{\mathrm{CM}}(\boldsymbol{p})$ is the size of giant in the configuration model with limiting degree distribution having probability mass function $\boldsymbol{p} = (p_i)_{i \geq 1}$. Recall that $\rho_{\mathrm{CM}}(\boldsymbol{p}) = 1 - g_D(\eta)$, where η satisfies $g'_D(\eta) = \mathbb{E}[D]\eta$.

Lemma A.1. *For all* $\varepsilon \leq p_k$, $\rho_{\mathrm{CM}}(\boldsymbol{p}) \leq \rho_{\mathrm{CM}}(\boldsymbol{p}(\varepsilon))$.

Proof. Let $g(\cdot)$ and $g_\varepsilon(\cdot)$ be the generating functions corresponding to $\boldsymbol{p}$ and $\boldsymbol{p}(\varepsilon)$, respectively. Let η and η_ε satisfy $g'(\eta) = \mathbb{E}[D]\eta$ and $g'_\varepsilon(\eta_\varepsilon) = (\mathbb{E}[D] + l\varepsilon)\eta_\varepsilon$. It is enough to show that $g_\varepsilon(\eta_\varepsilon) \leq g(\eta)$, since $\rho_{\mathrm{CM}}(\boldsymbol{p}) = 1 - g(\eta)$. For this, we note that

$$g_\varepsilon(s) = \sum_{i \geq 0} p_i(\varepsilon)s^i = g(s) - \varepsilon s^k(1 - s^l). \tag{A.1}$$

Thus $g_\varepsilon(s) \leq g(s)$ for all $s \in [0, 1]$. As a result, it is enough to show that $\eta_\varepsilon \leq \eta$. We differentiate (A.1) with respect to s to obtain

$$g'_\varepsilon(s) = g'(s) + (k + l)\varepsilon s^{k+l-1} - k\varepsilon s^{k-1}. \tag{A.2}$$

Substitution of η_ε in (A.2) gives

$$(\mathbb{E}[D] + l\varepsilon)\eta_\varepsilon = g'(\eta_\varepsilon) + (k + l)\varepsilon\eta_\varepsilon^{k+l-1} - k\varepsilon\eta_\varepsilon^{k-1}. \tag{A.3}$$

Differentiating (A.3) with respect to ε, we get

$$\frac{\partial \eta_\varepsilon}{\partial \varepsilon} = \frac{k\eta_\varepsilon^{k-1} - (k + l)\eta_\varepsilon^{k+l-1} + l\eta_\varepsilon}{g''(\eta_\varepsilon) - (\mathbb{E}[D] + l\varepsilon) + \varepsilon(k + l)(k + l - 1)\eta_\varepsilon^{k+l-2} - \varepsilon k(k - 1)\eta_\varepsilon^{k-2}}. \tag{A.4}$$

Differentiation of (A.1) twice gives

$$g''_\varepsilon(s) = g''(s) - \varepsilon k(k - 1)s^{k-2} + \varepsilon(k + l)(k + l - 1)s^{k+l-2}. \tag{A.5}$$

Substituting (A.5) in (A.4) gives

$$\frac{\partial \eta_\varepsilon}{\partial \varepsilon} = \frac{k\eta_\varepsilon^{k-1} - (k + l)\eta_\varepsilon^{k+l-1} + l\eta_\varepsilon}{g''_\varepsilon(\eta_\varepsilon) - (\mathbb{E}[D] + l\varepsilon)}. \tag{A.6}$$

Let $h(x) = g'_\varepsilon(x) - (\mathbb{E}[D] + l\varepsilon)x$. We know that $h(\eta_\varepsilon) = h(1) = 0$. Due to Rolle's theorem, there exists $\theta \in (\eta_\varepsilon, 1)$ so that $h'(\theta) = 0$, so that $g''_\varepsilon(\theta) = \mathbb{E}[D] + l\varepsilon$. Since $g''_\varepsilon(\cdot)$ is an increasing function, $g''_\varepsilon(\eta_\varepsilon) \le g''_\varepsilon(\theta) = \mathbb{E}[D] + l\varepsilon$. Thus

$$g''_\varepsilon(\eta) - (\mathbb{E}[D] + l\varepsilon) < 0.$$

As a result, the denominator in (A.6) is negative. We claim that the numerator in (A.6) as a polynomial in η is always positive.

Let $P(x) := ks^{k-2} - (k+l)s^{k+l-2} + l$. It is enough to show that $P(\cdot)$ is positive in $(0, 1)$. Modifying the proof of Lemma 3.3, we see that γ, the only maximiser of $P(\cdot)$ in $(0, 1)$, is given by

$$\gamma = \left(\frac{k(k-2)}{(k+l)(k+l-2)} \right)^{1/l}.$$

Since there is no minimiser in the interval $(0, 1)$, the minimum occurs either in 0 or in 1. We have $P(0) = l$ and $P(1) = 0$, so that $P(x) > 0$ for all $x \in (0, 1)$. Thus the right-hand side of (A.6) is always negative. This completes the proof. $\qquad\square$

Lemma A.2. *Let $\boldsymbol{p} \preccurlyeq_{\mathrm{st}} \boldsymbol{q}$. Then $\boldsymbol{q}$ can be reached from $\boldsymbol{p}$ by a series of ε-transformations.*

Proof. Define the sets I and J that partition $\mathbb{N}$ as

$$I := \{i \in \mathbb{N} \colon p_i < q_i\} \quad \text{and} \quad J := \{i \in \mathbb{N} \colon p_i \ge q_i\}.$$

Applying our ε-transformation on (i, j) with $\varepsilon = \varepsilon_{i,j}$ for all $i \in I$ and $j \in J \cap [i+1, \infty)$ on the α-sequence $\boldsymbol{p}$, we obtain the α-sequence, say $\boldsymbol{p}_\varepsilon$, which satisfies

$$\boldsymbol{p}_\varepsilon = \left(\left(p_i + \sum_{j>i, j \in J} \varepsilon_{i,j} \right)_{i \in I}, \left(p_j - \sum_{i<j, i \in I} \varepsilon_{i,j} \right)_{j \in J} \right).$$

Then $\boldsymbol{p}_\varepsilon = \boldsymbol{q}$, if $(\varepsilon_{i,j})_{(i,j) \in I \times J}$ satisfies

$$\sum_{j>i, j \in J} \varepsilon_{i,j} = q_i - p_i \quad \text{for all } i \in I \quad \text{and} \quad \sum_{i<j, i \in I} \varepsilon_{i,j} = p_j - q_j \quad \text{for all } j \in J. \tag{A.7}$$

Thus it is enough to prove that the solution to (A.7) exists. We construct this solution recursively to complete this proof. Let $I = \{i_1, i_2, \ldots\}$, where $i_1 = \min I$ and i_k is defined recursively as

$$i_k = \min I \setminus \{i_1, i_2, \ldots, i_{k-1}\},$$

and the minimum $\min I$ of a set of integers I is its minimal element. Let $J = \{j_1, j_2, \ldots\}$, where $j_1 = \min J$ and j_k is defined recursively as

$$j_k = \min J \setminus \{j_1, j_2, \ldots, j_{k-1}\}.$$

Notice that $1 \in I$ by stochastic domination. We next define some notation to complete our proof.

Notation A.1. (*Sums of differences.*) Let $s_0 = t_0 = 0$, and define $s_n = \sum_{k=1}^n (q_{i_k} - p_{i_k})$ and $t_n = \sum_{k=1}^n (p_{j_k} - q_{j_k})$.

TABLE 1. Values of $\varepsilon_{i,j}$.

$i \backslash j$	$j_{l_{i'}+1}$	$j_{l_{i'}+2}$	$\cdots$	$j_{l_{i'+1}}$	$j_{l_{i'+1}+1}$	$\cdot\ \cdot$
$i_{n_{i'}+1}$	$s_{n_{i'}+1} - t_{l_{i'}}$	0	$\cdots$	0	0	$\cdot\ \cdot$
$i_{n_{i'}+2}$	$q_{n_{i'}+2} - p_{n_{i'}+2}$	0	$\cdots$	0	0	$\cdot\ \cdot$
$i_{n_{i'}+3}$	$q_{n_{i'}+3} - p_{n_{i'}+3}$	0	$\cdots$	0	0	$\cdot\ \cdot$
$\cdot$	$\cdot$	$\cdot$	$\cdots$	$\cdot$	$\cdot$	$\cdot\ \cdot$
$\cdot$	$\cdot$	$\cdot$	$\cdots$	$\cdot$	$\cdot$	$\cdot\ \cdot$
$i_{n_{i'+1}}$	$q_{n_{i'+1}} - p_{n_{i'+1}}$	0	$\cdots$	$\cdot$	$\cdot$	$\cdot\ \cdot$
$i_{n_{i'+1}+1}$	$t_{l_{i'}+1} - s_{n_{i'+1}}$	$t_{l_{i'}+2} - t_{l_{i'}+1}$	$\cdots$	$t_{l_{i'+1}} - t_{l_{i'+1}-1}$	$s_{n_{i'+1}+1} - t_{l_{i'+1}}$	$0\ \ \cdot$

We define $\varepsilon_{i,j}$ recursively to satisfy (A.7) in the following way. We define $n_0 = l_0 = 0$. We first prove that there exists $0 \le n_1 < j_1 - 1$ such that

$$s_{n_1} < t_1 \quad \text{and} \quad s_{n_1+1} \ge t_1,$$

and there exists $1 \le l_1$ such that

$$s_{n_1} < t_{l_1} \quad \text{and} \quad s_{n_1} \ge t_{l_1+1}.$$

Indeed, by stochastic domination,

$$s_{j_1-1} - t_1 = \sum_{k=1}^{j_1-1} (q_k - p_k) - (p_{j_1} - q_{j_1}) = \sum_{k=1}^{j_1} q_k - \sum_{k=1}^{j_1} p_k > 0.$$

We have $n_1 + 1 = \min\{i \in \mathbb{N}: s_i \ge t_1\}$, and $s_{j_1-1} \ge t_1$. Thus $n_1 < j_1 - 1$. Since $j_1 = \min J$, this means that $\{1, 2, 3, \ldots, j_1 - 1\} \subset I$. Thus $i_{n_1} = n_1 < j_1 - 1$ and $i_{n_1+1} = n_1 + 1 < j_1$. Having obtained n_1 and l_1, we define $\varepsilon_{i,j}$ for $1 \le i \le i_{n_1+1}$ and $j_1 \le j \le j_{l_1+1}$ by substituting $i' = 0$ in Table 1. This defines $\varepsilon_{i,j}$ for $1 \le i \le i_{n_1+1}$ and $j_1 \le j \le j_{l_1+1}$. Further, we set $\varepsilon_{i,j} = 0$ for all $i > n_1 + 1$ and $j < j_{l_1}$.

We next iterate this argument, and prove that, similarly, there exists $n_2 > n_1$ satisfying

$$s_{n_2} - s_{n_1} < t_{l_1+1} - t_{l_1} \quad \text{and} \quad s_{n_2+1} - s_{n_1} \ge t_{l_1+1} - t_{l_1},$$

and there exists $l_2 > l_1$ such that

$$s_{n_2} - s_{n_1} < t_{l_2} - t_{l_1} \quad \text{and} \quad s_{n_2+1} - s_{n_1} \ge t_{l_2+1} - t_{l_1}.$$

As before, by stochastic domination, $i_{n_2+1} < j_{l_1+1}$. Having obtained n_2 and l_2, we define $\varepsilon_{i,j}$ for $i_{n_1+1} \le i \le i_{n_2+1}$ and $j_{l_1+1} \le j \le j_{l_2+1}$ by substituting $i' = 1$ in Table 1.

We next distinguish different cases, depending on whether I and J are finite or infinite.

Suppose first that both I and J are infinite sets. In that case the recursion never stops and we keep on defining (n_s, j_s), and, in turn, $\varepsilon_{i,j}$ for all values of (i, j), hence getting the solution for (A.7).

Suppose next that one of I or J is finite. Then, due to stochastic domination, the set I will have to be the one to be finite. Suppose $\max I = k$. Then there will exist $m \in \mathbb{N}$ such that $i_{n_m+1} = k$. Then we define $\varepsilon_{i,j}$ in Table 2. To complete the proof, we need to show that

$$s_{n_m+1} = \lim_{n \to \infty} t_n. \tag{A.8}$$

TABLE 2. Values of $\varepsilon_{i,j}$.

$i \backslash j$	$j_{l_{m-1}+1}$	$j_{l_{m-1}+2}$	$\cdots$	j_n	j_{n+1}	j_{n+2}
$i_{n_{m-1}+2}$	$q_{n_1+2} - p_{n_1+2}$	0	$\cdots$	0	0	$\cdots$
$i_{n_{m-1}+3}$	$q_{n_1+3} - p_{n_1+3}$	0	$\cdots$	0	0	$\cdots$
$\cdot$	$\cdot$	$\cdot$	$\cdots$	$\cdot$	$\cdot$	$\cdot$
$\cdot$	$\cdot$	$\cdot$	$\cdots$	$\cdot$	$\cdot$	$\cdot$
i_{n_m}	$q_{n_2} - p_{n_2}$	0	$\cdots$	0	0	$\cdots$
$k = i_{n_m+1}$	$t_{l_{m-1}+1} - s_{n_m}$	$t_{l_{m-1}+2} - t_{l_{m-1}+1}$	$\cdots$	$t_n - t_{n-1}$	$t_{n+1} - t_n$	$\cdots$

This is because the sum of the kth row is $\lim_{n \to \infty} t_n - s_{n_m}$. If (A.8) is true, then the kth row sum becomes $q_k - p_k$, which is as desired in (A.7). Now we show (A.8). Since I and J partition the natural numbers, we have $I \cup J = \mathbb{N}$ and thus

$$\sum_{i \in I \cup J} p_i = \sum_{j \in I \cup J} q_j = 1, \quad \text{so that} \quad \sum_{i \in I} p_i + \sum_{i \in J} p_i = \sum_{i \in I} q_i + \sum_{j \in J} q_j.$$

In turn this implies that

$$\sum_{i \in I} (p_i - q_i) = \sum_{j \in J} (q_j - p_j).$$

Notice that, by definition, $\lim_{n \to \infty} t_n = \sum_{i \in I} (p_i - q_i)$ and $s_{n_m+1} = \sum_{j \in J} (q_j - p_j)$. This finishes the proof. $\square$

Proof of Theorem 1.6. Lemmas A.1 and A.2 complete the proof. $\square$

Appendix B. Characterisation of stochastic ordering in α-sequences

In this section, we prove Lemmas 3.5 and 3.6. By definition of stochastic domination, $\boldsymbol{r} \preccurlyeq_{\boldsymbol{p}} \boldsymbol{r}' \iff \boldsymbol{pr} \preccurlyeq_{st} \boldsymbol{pr}'$. Thus Lemma 3.5 follows from Lemma A.2. Therefore we only prove Lemma 3.6.

Proof of Lemma 3.6. We define our ε-sequence $(\delta_i)_{i \geq 1}$ recursively. First we define

$$k_1 := \begin{cases} \min\{n \in \mathbb{N} : r'_n < r_n\} & \text{if the minimum is attained,} \\ \infty & \text{otherwise.} \end{cases}$$

We know that $k_1 > 1$, due to $\boldsymbol{r} \preccurlyeq_{\boldsymbol{p}} \boldsymbol{r}'$. Define $\delta_i = r'_i - r_i$ for each $i < k_1$. If $k_1 = \infty$, we define $\delta_i = r'_i - r_i$ for all $i \in \mathbb{N}$, so that, by definition of k_1, $\delta_i = r'_i - r_i \geq 0$, and thus it satisfies the required condition. Suppose instead that $k_1 < \infty$. Then we define

$$l_1 = \min\left\{ n > k_1 : \sum_{i=k_1}^{n} p_i(r'_i - r_i) > 0 \right\}.$$

Since $\boldsymbol{r} \preccurlyeq_{\boldsymbol{p}} \boldsymbol{r}'$ and $k_1 < \infty$, by definition we have $\sum_{i=k_1}^{\infty} p_i(r'_i - r_i) > 0$. As a result, there exists m such that $\sum_{i=k_1}^{m} p_i(r'_i - r_i) > 0$, which means that the set whose minimum we want to compute is non-empty. By the well ordering principle for natural numbers, we can conclude

that the minimum is attained and thus $l_1 < \infty$. Having obtained $l_1 < \infty$, we define δ_i for all $1 \le i \le l_1$ as

$$\delta_i := \begin{cases} r'_i - r_i & \text{for } i < k_1, \\ 0 & \text{for } k_1 \le i < l_1, \\ \frac{1}{p_{l_1}} \sum_{i=k_1}^{l_1} (r'_i - r_i) p_i & \text{for } i = l_1. \end{cases}$$

Next we define k_2, similarly to k_1, as

$$k_2 := \begin{cases} \min\{n > l_1 : r'_n < r_n\} & \text{if the minimum is attained,} \\ \infty & \text{otherwise.} \end{cases}$$

If $k_2 = \infty$, then we define $\delta_i = r'_i - r_i$ for all $i > l_1$ and it again satisfies the required condition. Suppose instead that $k_2 < \infty$. Then we define l_2 similarly to l_1 as

$$l_2 = \min\left\{ n > k_2 : \sum_{i=k_2}^{n} p_i(r'_i - r_i) > 0 \right\}.$$

Since $\boldsymbol{r} \preccurlyeq_{\boldsymbol{p}} \boldsymbol{r}'$, it can be shown that the above minimum is always attained and $l_2 < \infty$. We define δ_i for all $k_2 \le i \le l_2$ as

$$\delta_i := \begin{cases} r'_i - r_i & \text{for } l_1 \le i < k_2, \\ 0 & \text{for } k_2 \le i < l_2, \\ \frac{1}{p_{l_1}} \sum_{i=k_1}^{l_1} (r'_i - r_i) p_i & \text{for } i = l_1. \end{cases}$$

We keep on defining k_i and l_i similarly, unless $k_i = \infty$ for some $i \ge 1$. In this way, we get the desired ε-sequence iteratively. $\qquad\square$

Funding information

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] ALDOUS, D. AND STEELE, J. M. (2004). The objective method: probabilistic combinatorial optimization and local weak convergence. In *Probability on Discrete Structures* (Encyclopaedia Math. Sci. **110**), ed. H. Kesten, pp. 1–72. Springer, Berlin.
[2] ANTHONISSE, J. M. (1971). The rush in a directed graph. *Stichting Mathematisch Centrum, Mathematische Besliskunde* **BN 9/71**, 1–10.
[3] AVRACHENKOV, K. AND LEBEDEV, D. (2006). PageRank of scale-free growing networks. *Internet Math.* **3**, 207–231.
[4] AVRACHENKOV, K., LITVAK, N., NEMIROVSKY, D. AND OSIPOVA, N. (2007). Monte Carlo methods in PageRank computation: when one iteration is sufficient. *SIAM J. Numer. Anal.* **45**, 890–904.

[5] BANERJEE, S. AND OLVERA-CRAVIOTO, M. (2022). PageRank asymptotics on directed preferential attachment networks. *Ann. Appl. Prob.* **32**, 3060–3084.

[6] BENJAMINI, I. AND SCHRAMM, O. (2001). Recurrence of distributional limits of finite planar graphs. *Electron. J. Prob.* **6**, 13.

[7] BIANCHINI, M., GORI, M. AND SCARSELLI, F. (2005). Inside PageRank. *ACM Trans. Internet Technol.* **5**, 92–128.

[8] BOLDI, P. AND VIGNA, S. (2014). Axioms for centrality. *Internet Math.* **10**, 222–262.

[9] BOLDI, P., SANTINI, M. AND VIGNA, S. (2005). PageRank as a function of the damping factor. In *Proceedings of the 14th International Conference on World Wide Web (WWW '05)*, pp. 557–566. ACM, New York.

[10] BOLLOBÁS, B. (1980). A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combinatorics* **1**, 311–316.

[11] BOLLOBÁS, B. AND RIORDAN, O. (2015). An old approach to the giant component problem. *J. Combinatorial Theory B* **113**, 236–260.

[12] BORGATTI, S. P. (2005). Centrality and network flow. *Soc. Netw.* **27**, 55–71.

[13] BRIN, S. AND PAGE, L. (1998). The anatomy of a large-scale hypertextual web search engine. *Comput. Netw. ISDN Syst.* **30**, 107–117.

[14] BRITTON, T., DEIJFEN, M. AND MARTIN-LÖF, A. (2006). Generating simple random graphs with prescribed degree distribution. *J. Statist. Phys.* **124**, 1377–1397.

[15] CHEN, N., LITVAK, N. AND OLVERA-CRAVIOTO, M. (2014). PageRank in scale-free random graphs. In *Algorithms and Models for the Web Graph (WAW 2014)* (Lecture Notes Comput. Sci. **8882**), ed. A. Bonato *et al.*, pp. 120–131. Springer, Cham.

[16] CHEN, N., LITVAK, N. AND OLVERA-CRAVIOTO, M. (2017). Generalized PageRank on directed configuration networks. *Random Structures Algorithms* **51**, 237–274.

[17] DEIJFEN, M., ROSENGREN, S. AND TRAPMAN, P. (2018). The tail does not determine the size of the giant. *J. Statist. Phys.* **173**, 736–745.

[18] EVANS, T. S. AND CHEN, B. (2022). Linking the network centrality measures closeness and degree. *Commun. Phys.* **5**, 172.

[19] GARAVAGLIA, A., HOFSTAD, R. V. D. AND LITVAK, N. (2020). Local weak convergence for PageRank. *Ann. Appl. Prob.* **30**, 40–79.

[20] HINNE, M. (2011). Local approximation of centrality measures. Master's thesis, Radboud University Nijmegen, The Netherlands.

[21] HOFSTAD, R. V. D. (2017). *Random Graphs and Complex Networks*, Vol. 1 (Cambridge Series in Statistical and Probabilistic Mathematics **43**). Cambridge University Press.

[22] HOFSTAD, R. V. D. (2021). The giant in random graphs is almost local. Available at arXiv:2103.11733.

[23] HOFSTAD, R. V. D. (2023+). *Random Graphs and Complex Networks*, Vol. 2. In preparation. Available at http://www.win.tue.nl/~rhofstad/NotesRGCNII.pdf.

[24] HOFSTAD, R. V. D., HOOGHIEMSTRA, G. AND VAN MIEGHEM, P. (2005). Distances in random graphs with finite variance degrees. *Random Structures Algorithms* **27**, 76–123.

[25] HOFSTAD, R. V. D., HOOGHIEMSTRA, G. AND ZNAMENSKI, D. (2007). Distances in random graphs with finite mean and infinite variance degrees. *Electron. J. Prob.* **12**, 703–766.

[26] HOFSTAD, R. V. D., HOOGHIEMSTRA, G. AND ZNAMENSKI, D. (2007). A phase transition for the diameter of the configuration model. *Internet Math.* **4**, 113–128.

[27] JANSON, S. (2009). On percolation in random graphs with given vertex degrees. *Electron. J. Prob.* **14**, 87–118.

[28] JANSON, S. (2009). The probability that a random multigraph is simple. *Combinatorics Prob. Comput.* **18**, 205–225.

[29] JANSON, S. (2014). The probability that a random multigraph is simple, II. *J. Appl. Prob.* **51A**, 123–137.

[30] JANSON, S. AND LUCZAK, M. J. (2009). A new approach to the giant component problem. *Random Structures Algorithms* **34**, 197–216.

[31] JELENKOVIĆ, P. R. AND OLVERA-CRAVIOTO, M. (2010). Information ranking and power laws on trees. *Adv. Appl. Prob.* **42**, 1057–1093.

[32] LEE, J. AND OLVERA-CRAVIOTO, M. (2020). PageRank on inhomogeneous random digraphs. *Stochastic Process. Appl.* **130**, 2312–2348.

[33] LESKELÄ, L. AND NGO, H. (2015). The impact of degree variability on connectivity properties of large networks. In *Algorithms and Models for the Web Graph (WAW 2015)* (Lecture Notes Comput. Sci. **9479**), ed. D. Gleich *et al.*, pp. 66–77. Springer, Cham.

[34] LITVAK, N., SCHEINHARDT, W. R. W. AND VOLKOVICH, Y. (2007). In-Degree and PageRank: why do they follow similar power laws? *Internet Math.* **4**, 175–198.

[35] MOCANU, D. C., EXARCHAKOS, G. AND LIOTTA, A. (2018). Decentralized dynamic understanding of hidden relations in complex networks. *Scientific Reports* **8**, 1571.

[36]   MOLLOY, M. AND REED, B. (1995). A critical point for random graphs with a given degree sequence. *Random Structures Algorithms* **6**, 161–179.

[37]   MOLLOY, M. AND REED, B. (1998). The size of the giant component of a random graph with a given degree sequence. *Combinatorics Prob. Comput.* **7**, 295–305.

[38]   NEWMAN, M. (2018). *Networks*, 2nd edn. Oxford University Press.

[39]   PAGE, L., BRIN, S., MOTWANI, R. AND WINOGRAD, T. (1999). The PageRank citation ranking: bringing order to the web. Technical Report 1999-66, Stanford InfoLab. Previous number = SIDL-WP-1999-0120.

[40]   RUDIN, W. (1964). *Principles of Mathematical Analysis*, 2nd edn. McGraw-Hill, New York.

[41]   SENANAYAKE, U., PIRAVEENAN, M. AND ZOMAYA, A. (2015). The Pagerank-Index: going beyond citation counts in quantifying scientific impact of researchers. *PLoS One* **10**, e0134794.

[42]   WEI, X., ZHAO, J., LIU, S. AND WANG, Y. (2022). Identifying influential spreaders in complex networks for disease spread and control. *Scientific Reports* **12**, 5550.