THE ASYMPTOTIC VARIANCE OF THE GIANT COMPONENT OF CONFIGURATION MODEL RANDOM GRAPHS

BY FRANK BALL AND PETER NEAL

University of Nottingham and Lancaster University

For a supercritical configuration model random graph, it is well known that, subject to mild conditions, there exists a unique giant component, whose size R_n is $O(n)$, where n is the total number of vertices in the random graph. Moreover, there exists $0 < \rho \le 1$ such that $R_n/n \xrightarrow{p} \rho$ as $n \to \infty$. We show that for a sequence of *well behaved* configuration model random graphs with a deterministic degree sequence satisfying $0 < \rho < 1$; there exists $\sigma^2 > 0$, such that $\text{var}(\sqrt{n}(R_n/n - \rho)) \to \sigma^2$ as $n \to \infty$. Moreover, an explicit, easy to compute, formula is given for σ^2. This provides a key stepping stone for computing the asymptotic variance of the size of the giant component for more general random graphs.

1. Introduction. The theoretical treatment of random graphs goes back to Erdős and Rényi (1959). The Erdős–Rényi random graph is constructed as follows. Suppose that there are n vertices, labeled $1, 2, \ldots, n$. An edge exists between vertices i and j with probability $\mu/(n-1)$, independent of the remainder of the graph, where it is assumed that $0 \le \mu \le n-1$. Therefore, for any vertex i, its degree is $\text{Bin}(n-1, \mu/(n-1))$ [a binomial random variable with $n-1$ trials and success probability $\mu/(n-1)$] and as $n \to \infty$ with μ fixed, the vertex degree distribution converges to $\text{Po}(\mu)$ (a Poisson random variable with mean μ). It is well known that for large n, the order of magnitude of the size of the largest connected component, R_n, of the Erdős–Rényi random graph depends upon the threshold parameter μ. If $\mu < 1$ (subcritical), $R_n = O_p(\log n)$, if $\mu = 1$ (critical), $R_n = O_p(n^{\frac{2}{3}})$ and if $\mu > 1$ (supercritical), there is a constant $\theta(\mu) > 0$ so that $R_n \overset{p}{\sim} \theta(\mu)n$. [Here, $R_n = O_p(f(n))$ means that there exists a constant $C < \infty$ such that $P(R_n \le Cf(n)) \to 1$ as $n \to \infty$ and $R_n \overset{p}{\sim} f(n)$ means that $R_n/f(n)$ converges in probability to 1 as $n \to \infty$.] If $R_n \overset{p}{\sim} Cn$, for some $C > 0$, we say that a giant component exists. The size of the second largest component is $O_p(\log n)$ in the supercritical case, so there exists a unique giant component. For a sequence of Erdős–Rényi random graphs $\{\mathcal{G}_n\}$, indexed by the total number of vertices n, as $n \to \infty$ with fixed $\mu > 1$, it can be shown that

$$\frac{1}{n}R_n \xrightarrow{p} \rho \qquad \text{as } n \to \infty,$$

where ρ is the nonzero solution to $\rho = 1 - \exp(-\mu\rho)$ and $\xrightarrow{p}$ denotes convergence in probability. Furthermore, a central limit theorem exists, with

$$(1.1) \qquad \sqrt{n}\left(\frac{1}{n}R_n - \rho\right) \xrightarrow{D} N(0, \sigma_{\mathrm{ER}}^2) \qquad \text{as } n \to \infty,$$

where $\sigma_{\mathrm{ER}}^2 = \frac{\rho(1-\rho)}{(1-\mu(1-\rho))^2}$ and $\xrightarrow{D}$ denotes convergence in distribution. See, for example, Durrett (2007), Chapter 2, for a more detailed discussion of Erdős–Rényi random graphs.

Many graphs observed in the social sciences, epidemiology and computing do not have a Poisson (vertex) degree distribution, although the Erdős–Rényi random graph provides a useful representation of the Reed–Frost epidemic in an homogeneously mixing population; see Barbour and Mollison (1990). Therefore, there has been considerable interest in random graphs with arbitrary degree distributions. Random graphs with mixed Poisson degree distribution arise by letting each vertex i have a connectivity parameter C_i with the probability of an edge existing between vertices i and j being proportional to $C_i C_j$. Mixed Poisson random graphs have been studied in, for example, Chung and Lu (2002) and Britton, Deijfen and Martin-Löf (2006), Section 3, with a central limit theorem for the size of the giant component given by Neal (2007), Theorem 5.2. The mixed Poisson distribution has the property that its variance is greater than or equal to its mean, with equality if and only if the distribution is Poisson. By contrast, the configuration model considered in this paper allows for an arbitrary but specified vertex degree distribution.

In Molloy and Reed (1995), the configuration model with a deterministic degree sequence was considered. The configuration model was introduced in Bollobás (1980) and we refer the reader to Bollobás (2001), Section 2.4, for further references. Suppose that D_i^n is the degree of vertex i and suppose that $S_n = \sum_{i=1}^n D_i^n$ is even. Then to vertex i assign D_i^n half-edges and pair up half-edges uniformly at random to form edges between vertices. This we term the Molloy–Reed (MR) random graph. It is shown in Molloy and Reed (1995) that if $\frac{1}{n}\sum_{j=1}^n D_j^n(D_j^n - 2) \to \kappa > 0$ as $n \to \infty$ and there exists $\boldsymbol{\pi} = (\pi_0, \pi_1, \ldots)$, satisfying $\sum_{i=0}^\infty \pi_i = 1$, such that for $i = 0, 1, \ldots, \sum_{j=1}^n 1_{\{D_j^n=i\}}/n \to \pi_i$ as $n \to \infty$, there exists a giant component of size $O(n)$, subject to there existing $\delta > 0$ such that for all sufficiently large n, $\Delta_n = \max_{\{1 \le i \le n\}} D_i^n \le n^{1/4-\delta}$. In Molloy and Reed (1998), it was shown that $R_n \overset{p}{\sim} \rho n$, where ρ satisfies

$$(1.2) \qquad\qquad\qquad \rho = 1 - f(z),$$

and z is the solution in $[0, 1)$ of

$$(1.3) \qquad\qquad\qquad z = \frac{1}{\mu}f'(z),$$

where $f(s) = \sum_{i=0}^{\infty} \pi_i s^i$, $f'(s) = \sum_{i=1}^{\infty} \pi_i i s^{i-1}$ ($s \in [0, 1]$) and $\mu = \sum_{i=1}^{\infty} i\pi_i$. The equations for ρ given in (1.2) and (1.3) are based on those given in Newman, Strogatz and Watts (2001), where the deterministic degree sequence of Molloy and Reed (1995) is replaced by taking $D_1^n, D_2^n, \ldots, D_n^n$ to be i.i.d. (independent and identically distributed) according to a nonnegative integer-valued random variable, D, with $P(D = i) = \pi_i$ ($i = 0, 1, \ldots$). Note that using an i.i.d. degree sequence can result in an infeasible degree sequence, with $\sum_{i=1}^{n} D_i^n$ being odd. In such an event, the entire degree sequence may be resampled until a feasible degree sequence is obtained or alternatively the final half-edge in the construction of the random graph can simply be ignored. These equations are equivalent, but simpler, than those given in Molloy and Reed (1998) for ρ, and both the MR and NSW (Newman–Strogatz–Watts construction with i.i.d. degree sequence) random graphs have the same asymptotic proportion of vertices, ρ, in the giant component for a given $\boldsymbol{\pi}$. The above construction of the MR and NSW graphs can lead to a nonsimple graph. That is, the random graph contains *imperfections*, in that some individuals may be linked to themselves and there may be multiple edges between pairs of individuals. Provided that D has finite variance, such imperfections are sparse in the limit as $n \to \infty$, see Durrett (2007), Theorem 3.1.2. The results of this paper also hold if the graph is conditioned on being simple, that is, having no such imperfections [cf. Janson (2009b) and Britton, Janson and Martin-Löf (2007)].

The aim of the current work is to derive the (asymptotic) variance of the giant component for the MR random graph, $\text{var}(\sqrt{n}(R_n/n - \rho)) \to \sigma^2$ as $n \to \infty$ with an explicit, easy to compute formula for σ^2. This provides a key stepping stone for computing the asymptotic variance of the size of the giant component for more general random graphs such as the NSW random graph which we discuss briefly in Section 7. Variance calculations for random graphs are limited, although progress has been made in the near critical case, Riordan (2012). In Section 2, we define a sequence of MR random graphs $\{\mathcal{G}_n\}$ and state the key result Theorem 2.1, along with the weak conditions required on the degree sequence, which are similar to the conditions stated in Molloy and Reed (1995). Also in Section 2, we show that $\text{var}(\sqrt{n}(R_n/n - \rho))$ has the same asymptotic limit as $n \to \infty$ as $\text{var}(\tilde{U}_n/\sqrt{n})$, where $\tilde{U}_n$ is the total number of vertices which belong to components of size less than or equal to $[n^{\beta}]$ and β is any fixed real number satisfying $0 < \beta < 1/12$. In Section 3, we introduce a branching process approximation for the construction of components in the random graph $\mathcal{G}_n$ which assists with computing $\lim_{n\to\infty} \text{var}(\tilde{U}_n/\sqrt{n})$. The branching process approximation is valid for the initial growth of components and mimics similar branching process approximations used for epidemic models; see, for example, Whittle (1955) and Ball and Donnelly (1995). In Section 4, we compute $\lim_{n\to\infty} \text{var}(\tilde{U}_n/\sqrt{n})$ and this section contains most of the technical details of the proof. Then in Section 5, we show that the expression obtained for $\lim_{n\to\infty} \text{var}(\tilde{U}_n/\sqrt{n})$ in Section 4 is equal to σ^2. In Section 6, we consider the weakly supercritical case studied in Riordan (2012). In

this case, $\rho = 0$ and $\sigma^2 = \infty$ but by indexing ρ and σ^2 by n, we show that $\sigma_n^2 \sim \Sigma_n$ (i.e. $\sigma_n^2/\Sigma_n \to 1$ as $n \to \infty$), where Σ_n is defined in Riordan (2012), Theorem 1.1 (and Section 6) and shown to satisfy $\mathrm{var}(\sqrt{n}(R_n/n - \rho_n)) \sim \Sigma_n$. In Section 7, we compare the asymptotic calculations with simulation studies for graphs of size $n = 200$ and $n = 1000$ and conjecture the existence of a Gaussian central limit theorem for the size of the giant component. We also outline how to compute the asymptotic variance of the giant component of the NSW random graph, the details of which will be presented elsewhere. Finally, in the Appendix we present some useful results for Galton–Watson branching processes that are used in the proof of Theorem 2.1.

2. Conditions on the degree sequence and statement of theorem. For the asymptotic variance of the Molloy–Reed (MR) random graph, we need to impose conditions upon the sequence of degree distributions $\{\mathbf{D}^n\}$. Let $\bar{\mathbf{D}}^n = (\bar{D}_1^n, \bar{D}_2^n, \ldots, \bar{D}_n^n)$ denote the degrees of the vertices in $\mathcal{G}_n$, where $\{\mathcal{G}_n\}$ is a sequence of Molloy–Reed random graphs. However, we do not at this stage assign a particular degree to a given vertex but instead look to exploit the exchangeability of vertices with $\mathbf{D}^n$ simply being a random permutation of $\bar{\mathbf{D}}^n$ with D_i^n denoting the degree of vertex i. The following arguments are simpler using $\bar{\mathbf{D}}^n$ (and exchangeability) rather than using $\mathbf{D}^n$. Thus, throughout the paper we take $\mathcal{G}_n$ to be constructed using $\bar{\mathbf{D}}^n$ with all $n!$ vertex labellings equally likely. For $j = 0, 1, \ldots,$ let $\pi_j^n = \sum_{i=1}^{n} 1_{\{\bar{D}_i^n = j\}}/n$ with $\boldsymbol{\pi}^n = (\pi_0^n, \pi_1^n, \ldots)$ and let $\mu_n = \sum_{i=1}^{n} \bar{D}_i^n/n$. We assume that there exists a proper nonnegative, integer-valued random variable D with $P(D = j) = \pi_j$ $(j = 0, 1, \ldots)$ such that, for all $j = 0, 1, \ldots, \pi_j^n \to \pi_j$ as $n \to \infty$. Thus, the degree distributions are assumed to satisfy condition 1 of Molloy and Reed (1995), being *smooth* and *feasible*. In Molloy and Reed (1995, 1998), additional conditions are placed on the sequence of degree distributions for them to be *well behaved*. We place slightly stronger regularity conditions upon the sequence of degree distributions as follows.

For $k = 1, 2, 3$, let $L_k = \sum_{j=0}^{\infty} j^k \pi_j (= E[D^k])$, which we assume to be finite. (Note that $\mu = L_1$.) We require that $\nu = \sum_{j=0}^{\infty} j(j-1)\pi_j/\mu = (L_2 - L_1)/L_1 > 1$ and $\pi_1 > 0$. These ensure that the asymptotic degree sequence is supercritical (i.e., that a giant component exists in the sense defined in Section 1) and that $0 < z < 1$ [where z satisfies (1.3)], respectively.

We require also that $\{\bar{\mathbf{D}}^n\}$ satisfies the following conditions.

(a) For all $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$,

    (i) for all $j \geq 0$,

(2.1) $$j^2 |\pi_j^n - \pi_j| < \varepsilon;$$

    and

(ii) for $k = 1, 2, 3$,

$$
(2.2) \qquad \left| \sum_{j=0}^{\infty} j^k \pi_j^n - L_k \right| < \varepsilon.
$$

(b) There exists $\beta > 0$ such that for all $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$,

$$
(2.3) \qquad n^{2\beta} \sum_{j=0}^{\infty} j \left| \pi_j^n - \pi_j \right| < \varepsilon.
$$

(c) There exists $\delta > 0$ such that $n^{-(1/4-\delta)} \Delta_n \to 0$ as $n \to \infty$, where $\Delta_n = \max_{1 \leq i \leq n} \bar{D}_i^n$.

Note that condition (a)(i) implies Molloy and Reed (1995), condition 2. Throughout this paper, we take β to be an arbitrary, positive constant satisfying $0 < \beta < 1/12$ and condition (b). Immediate consequences of the above conditions include Molloy and Reed (1995), condition 3(a): for all $\varepsilon > 0$, there exists K and n_1 such that $|\sum_{j=1}^{K} j(j-2)\pi_j^n - \sum_{j=0}^{\infty} j(j-2)\pi_j| < \varepsilon$ for all $n \geq n_1$. Also $\mu_n = \sum_{i=1}^{n} \bar{D}_i^n / n \to \mu (= E[D])$, $\nu_n = \sum_{i=1}^{n} \bar{D}_i^n (\bar{D}_i^n - 1) / n \mu_n \to \nu$ and $\lambda_n = \sum_{i=1}^{n} \bar{D}_i^n (\bar{D}_i^n - 1)(\bar{D}_i^n - 2) / n \mu_n \to \lambda (= (L_3 - 3L_2 + 2L_1)/L_1) < \infty$ as $n \to \infty$.

THEOREM 2.1. *For a Molloy–Reed random graph satisfying conditions* (a)–(c), $\nu > 1$ *and* $\pi_1 > 0$,

$$
(2.4) \qquad \mathrm{var}\left(\sqrt{n} \left(\frac{R_n}{n} - \rho \right) \right) \to \sigma^2 \qquad as\ n \to \infty,
$$

with

$$
(2.5) \quad
\begin{aligned}
\sigma^2 = {}& 1 - \rho - f(z^2) + \frac{z^2}{1 - f''(z)/\mu} \{ (1 + z^2)\mu - 2f'(z^2) \} \\
& + \frac{z^2}{(1 - f''(z)/\mu)^2} \{ z^2 \mu + z^2 f''(z) - f'(z^2) - z^2 f''(z^2) \},
\end{aligned}
$$

where ρ satisfies (1.2).

A couple of remarks concerning Theorem 2.1 are as follows. First, provided a closed-form expression is available for the generating function f, the formula for σ^2 is easy to compute being a function of ρ, z, μ, ν, $f'(z^2)$, $f''(z^2)$ and $f''(z)$ only. Note that computing z usually requires numerical solution of a nonlinear equation. Second, since $f'(s)$ is a convex function in s, z is the solution in $[0, 1)$ of $z = f'(z)/\mu$ and $f'(0) \geq 0$, we have that $f''(z)/\mu < 1$.

The initial observations in proving Theorem 2.1 are as follows. First, let $U_n = n - R_n$ be the total number of vertices outside of the giant component and observe that to prove Theorem 2.1 it suffices to show that $\text{var}(\sqrt{n}(U_n/n - \omega)) = \text{var}(U_n/\sqrt{n}) \to \sigma^2$ as $n \to \infty$, where $\omega = 1 - \rho$. We introduce U_n as it is simpler to study small components than the giant component as we can utilise a branching process approximation (Section 3) for the initial growth of such components. However, working with U_n directly is difficult, so we introduce a new process $\tilde{U}_n$ defined below, which counts the total number of vertices in small components (less than $[n^\beta]$ vertices). For vertices i and j in the graph $\mathcal{G}_n$, let $i \leftrightarrow j$ denote that vertices i and j belong to the same connected component. For $n = 1, 2, \dots$ and $i = 1, 2, \dots, n$, let $\mathcal{C}_i^n = \{j : j \leftrightarrow i\}$ with $C_i^n = |\mathcal{C}_i^n|$. Note that either $\mathcal{C}_i^n = \mathcal{C}_j^n$ or $\mathcal{C}_i^n \cap \mathcal{C}_j^n = \varnothing$. Let $\chi_i^n = 1_{\{C_i^n \le [n^\beta]\}}$ and $\tilde{U}_n = \sum_{i=1}^n \chi_i^n$. Since the second largest component of a supercritical random graph almost surely consists of at most $\gamma \log n$ vertices for some $0 < \gamma < \infty$ [Molloy and Reed (1995), Lemma 11], the precise value of β is not important as the asymptotic behaviour of $\tilde{U}_n$ is the same for any value of $0 < \beta < 1/12$ and that is why reference to β in the notation of $\tilde{U}_n$ is omitted. (The above holds for any $0 < \beta < 1$.) Furthermore using the proof of Molloy and Reed (1995), Lemma 11, it is trivial to show that $n P(U_n \ne \tilde{U}_n) \to 0$, whence $E[(U_n - \tilde{U}_n)^2]/n \to 0$ as $n \to \infty$. Note that

$$
\begin{aligned}
(2.6) \quad & \text{var}(U_n/\sqrt{n}) - \text{var}(\tilde{U}_n/\sqrt{n}) \\
& = \text{var}((U_n - \tilde{U}_n)/\sqrt{n}) + 2\,\text{cov}((U_n - \tilde{U}_n)/\sqrt{n}, \tilde{U}_n/\sqrt{n}).
\end{aligned}
$$

Now $\text{var}((U_n - \tilde{U}_n)/\sqrt{n}) \le E[(U_n - \tilde{U}_n)^2]/n \to 0$ as $n \to \infty$. Suppose that $\text{var}(\tilde{U}_n/\sqrt{n}) \to \sigma^2$ as $n \to \infty$, where $\sigma^2 < \infty$. Then a simple argument using the Cauchy–Schwarz inequality shows that the right-hand side of (2.6) tends to 0 as $n \to \infty$, whence

$$
\big|\text{var}(U_n/\sqrt{n}) - \text{var}(\tilde{U}_n/\sqrt{n})\big| \to 0 \qquad \text{as } n \to \infty.
$$

Thus, we proceed by introducing useful branching process approximations for the construction of components (Section 3) before showing that $\text{var}(\tilde{U}_n/\sqrt{n}) \to \sigma^2$ as $n \to \infty$ (Section 4).

3. Branching process approximation. To study $\tilde{U}_n$, it is helpful to introduce approximating branching processes for the initial growth of components in the random graph $\mathcal{G}_n$. We begin by outlining the construction of a component in a random graph with a coupled branching process approximation. We also consider the limit, as $n \to \infty$, of the approximating branching processes. At the end of the section, we draw together the couplings of the size of a component in a random graph and the (total) size of the approximating branching processes by providing useful bounds on the probability that they are different.

Consider a subset of h of the n vertices in $\mathcal{G}_n$ and let $\mathbf{H}^n = (H_{k_1}^n, H_{k_2}^n, \dots, H_{k_h}^n)$ denote the degrees of the h vertices. Often, but not always, we will take $h = n$ and $\mathbf{H}^n = \bar{\mathbf{D}}^n$. Let $\theta_0^n, \theta_1^n, \dots$ be independent with θ_0^n drawn uniformly from

$\{k_1, k_2, \ldots, k_h\}$ and $\theta_1^n, \theta_2^n, \ldots$ distributed according to θ^n, where

$$(3.1) \qquad P(\theta^n = k_i) = \begin{cases} \dfrac{H_{k_i}^n}{h \mu^n(\mathbf{H}^n)}, & (i = 1, 2, \ldots, h), \\ 0, & \text{otherwise}, \end{cases}$$

with $\mu^n(\mathbf{H}^n) = h^{-1} \sum_{i=1}^{h} H_{k_i}^n$. Let $\mathcal{B}^n(\mathbf{H}^n)$ denote the (Galton–Watson) branching process with one initial ancestor constructed from $\mathbf{H}^n$ as follows. The initial ancestor in the branching process has $H_{\theta_0^n}^n$ offspring and the jth individual born in $\mathcal{B}^n(\mathbf{H}^n)$ has $H_{\theta_j^n}^n - 1$ offspring. Now a component, $\mathcal{C}^n(\mathbf{H}^n)$ can be constructed from the h vertices in $\mathbf{H}^n$ using $\theta_0^n, \theta_1^n, \theta_2^n, \ldots$ as follows.

Let $\mathcal{V}_0^n = \{\theta_0^n\}$, $\mathcal{E}_0^{n,U}$ be the set of the $H_{\theta_0^n}^n$ half-edges belonging to θ_0^n and $\mathcal{E}_0^{n,F}$ be the empty set. For $k = 0, 1, \ldots$, at stage k in the construction, $\mathcal{V}_k^n$, $\mathcal{E}_k^{n,U}$ and $\mathcal{E}_k^{n,F}$ are respectively the set of vertices currently in the component $\mathcal{C}^n(\mathbf{H}^n)$, the set of unattached half-edges and the set of formed complete edges. The construction now proceeds sequentially through stages $k = 1, 2, \ldots$. The construction stops at stage k if $\mathcal{E}_{k-1}^{n,U}$ is empty, in which case the component $\mathcal{C}^n(\mathbf{H}^n)$ is completely formed and $\mathcal{C}^n(\mathbf{H}^n) = \mathcal{V}_{k-1}^n$. Otherwise, a half-edge, E^* say, is chosen uniformly at random from $\mathcal{E}_{k-1}^{n,U}$ and is attached to a half-edge belonging to vertex θ_k^n, subject to the following conditions. If $\theta_k^n \notin \{\theta_0^n, \theta_1^n, \ldots, \theta_{k-1}^n\}$, then (i) $\mathcal{V}_k^n = \mathcal{V}_{k-1}^n \cup \{\theta_k^n\}$, (ii) the half-edge E^* is paired with a half-edge from θ_k^n to form a complete edge, which is added to the set $\mathcal{E}_{k-1}^{n,F}$ to give $\mathcal{E}_k^{n,F}$; and (iii) $\mathcal{E}_k^{n,U} = (\mathcal{E}_{k-1}^{n,U} \setminus \{E^*\}) \cup \mathcal{F}^*$, where $\mathcal{F}^*$ is the set of the other $H_{\theta_k^n}^n - 1$ half-edges that emanate from θ_k^n. If $\theta_k^n \in \{\theta_0^n, \theta_1^n, \ldots, \theta_{k-1}^n\}$, then an attempt is made to add the same vertex for a second (or higher order) time. In that case, choose a half-edge uniformly at random from the $H_{\theta_k^n}^n$ half-edges that emanated originally from θ_k^n. If the chosen half-edge belongs to $\mathcal{E}_{k-1}^{n,U} \setminus \{E^*\}$, then pair it with E^* to form a complete edge; this creates a cycle in $\mathcal{C}^n(\mathbf{H}^n)$, so $\mathcal{V}_k^n = \mathcal{V}_{k-1}^n$ with $\mathcal{E}_k^{n,F}$ and $\mathcal{E}_k^{n,U}$ being defined in the obvious fashion. Otherwise, a complete edge cannot be formed, in which case θ_k^n is not used in the construction of $\mathcal{C}^n(\mathbf{H}^n)$ and $(\mathcal{V}_k^n, \mathcal{E}_k^{n,U}, \mathcal{E}_k^{n,F}) = (\mathcal{V}_{k-1}^n, \mathcal{E}_{k-1}^{n,U}, \mathcal{E}_{k-1}^{n,F})$.

Let $B^n(\mathbf{H}^n)$ denote the total size, including the initial ancestor, of the branching process $\mathcal{B}^n(\mathbf{H}^n)$. Let $M_n^A(\mathbf{H}^n) = \min\{k : \theta_k^n \in \{\theta_0^n, \theta_1^n, \ldots, \theta_{k-1}^n\}\}$ and note that if $B^n(\mathbf{H}^n) \le M_n^A(\mathbf{H}^n)$, then $C^n(\mathbf{H}^n)(= |\mathcal{C}^n(\mathbf{H}^n)|) = B^n(\mathbf{H}^n)$. That is, the $B^n(\mathbf{H}^n)$ individuals in the branching process $\mathcal{B}^n(\mathbf{H}^n)$ correspond to distinct vertices in the graph $\mathcal{G}_n$. We make extensive use of the above coupling throughout the paper and this is the key motivation for focusing upon $\tilde{U}_n$, the total number of vertices in *small* components.

It is also helpful to introduce a branching process $\mathcal{B}$ which represents the limit as $n \to \infty$ of the branching processes $\mathcal{B}_1^n(= \mathcal{B}^n(\bar{\mathbf{D}}^n))$. Let $\tilde{D}$ be the random variable with probability mass function $P(\tilde{D} = k) = k\pi_k/\mu(= kP(D = k)/E[D])$ $(k = 1, 2, \ldots)$. Let B denote the total size, including the initial ancestor, of the branching (Galton–Watson) process $\mathcal{B}$, having offspring distribution D for the ini-

tial ancestor and $\tilde{D} - 1$ for all subsequent individuals. The initial ancestor in $\mathcal{B}_1^n$ has k offspring with probability π_k^n and subsequent individuals have $k - 1$ offspring with probability $k\pi_k^n/\mu_n$. Since by condition (a)(i), $\pi_k^n \to \pi_k$ as $n \to \infty$ for all $k \geq 0$, it follows that the offspring distribution of $\{\mathcal{B}_1^n\}$ converges in distribution to the offspring distribution of $\mathcal{B}$. Hence, $B_1^n (= B^n(\bar{\mathbf{D}}^n)) \xrightarrow{D} B$, the total size of the branching process $\mathcal{B}$; see, for example, Lefèvre and Utev (1999), Proposition 2.1. However, we shall require the following explicit coupling of $\mathcal{B}$ and $\mathcal{B}_1^n$ which goes beyond Lefèvre and Utev (1999), Proposition 2.1. Let $T_n(\bar{\mathbf{D}}^n) = \frac{1}{2}\sum_{j=0}^\infty |\pi_j^n - \pi_j|$, $\tilde{T}_n(\bar{\mathbf{D}}^n) = \frac{1}{2}\sum_{j=1}^\infty |\frac{j\pi_j^n}{\mu_n} - \frac{j\pi_j}{\mu}|$ and $\hat{T}_n(\bar{\mathbf{D}}^n) = \max\{T_n(\bar{\mathbf{D}}^n), \tilde{T}_n(\bar{\mathbf{D}}^n)\}$. Thus, $T_n(\bar{\mathbf{D}}^n)$ is the total variation distance between the offspring distributions of the initial ancestor in the branching processes $\mathcal{B}_1^n$ and $\mathcal{B}$, and $\tilde{T}_n(\bar{\mathbf{D}}^n)$ is the total variation distance between the offspring distributions of all subsequent individuals in these two branching processes. Using the triangle inequality, straightforward algebraic manipulation yields

$$(3.2) \quad \tilde{T}_n(\bar{\mathbf{D}}^n) \leq \frac{1}{2\mu}\sum_{j=1}^\infty j|\pi_j^n - \pi_j| + \frac{1}{2\mu}|\mu - \mu_n| \leq \frac{1}{\mu}\sum_{j=1}^\infty j|\pi_j^n - \pi_j|.$$

Hence,

$$(3.3) \qquad \hat{T}_n(\bar{\mathbf{D}}^n) \leq \frac{1}{\mu \wedge 2}\sum_{j=1}^\infty j|\pi_j^n - \pi_j|,$$

where $\mu \wedge 2 = \min\{\mu, 2\}$. By the existence of a maximal coupling [see, e.g., Barbour, Holst and Janson (1992), Appendix A1], random variables X_1^n and X_1 having distributions given by $P(X_1^n = k) = \pi_k^n$ and $P(X_1 = k) = \pi_k$ $(k = 0, 1, \ldots)$ can be defined on a common probability space so that $P(X_1^n \neq X_1) = T_n(\bar{\mathbf{D}}^n)$. Similarly, for $l = 2, 3, \ldots$, random variables X_l^n and X_l having distributions given by $P(X_l^n = k) = k\pi_k^n/\mu_n$ and $P(X_l = k) = k\pi_k/\mu$ $(k = 1, 2, \ldots)$ can be defined on a common probability space so that $P(X_l^n \neq X_l) = \tilde{T}_n(\bar{\mathbf{D}}^n)$. The branching processes $\mathcal{B}$ and $\mathcal{B}_1^n$ can be defined on a common probability space using independent realisations of (X_l^n, X_l) $(l = 1, 2, \ldots)$ in the obvious fashion. Let $M_n^D = \min\{l : X_l^n \neq X_l\}$ and note that M_n^D is stochastically larger than $\tilde{M}_n^D \sim$ Geom$(\hat{T}_n(\bar{\mathbf{D}}^n))$, a geometric distribution with support $\mathbb{N}$ and mean $\hat{T}_n(\bar{\mathbf{D}}^n)^{-1}$. [A random variable X is said to be stochastically larger than a random variable Y if, $P(X \leq x) \leq P(Y \leq x)$ for all $x \in \mathbb{R}$.] It follows that, if $B < M_n^D$, then $B_1^n = B$.

Let $\mathcal{C}_1^n = \mathcal{C}^n(\bar{\mathbf{D}}^n)$, corresponding to setting $D_1^n = \bar{D}_{\theta_0^n}^n$, let $C_1^n = |\mathcal{C}_1^n|$ and let $M_n^A = M_n^A(\bar{\mathbf{D}}^n)$. For $x \in \mathbb{R}$, let $[x]$ denote the greatest integer $\leq x$. The above couplings give that for any $1 \leq k \leq [n^\beta]$,

$$P(1_{\{C_1^n \leq k\}} \neq 1_{\{B_1^n \leq k\}}) \leq P(M_n^A \leq [n^\beta])$$

$$(3.4) \qquad\qquad = P\left(\bigcup_{0 \leq i < j \leq [n^\beta]} \{\theta_i^n = \theta_j^n\}\right)$$

$$\leq \frac{n^\beta (n^\beta + 1)}{2} \sum_{l=1}^{n} \left(\frac{\bar{D}_l^n}{n\mu_n} \right)^2$$

$$\to 0 \qquad \text{as } n \to \infty,$$

by condition (a)(ii), and

(3.5)
$$\begin{aligned}
P(1_{\{B_1^n \leq k\}} \neq 1_{\{B \leq k\}}) &\leq P(M_n^D \leq [n^\beta]) \\
&\leq P(\tilde{M}_n^D \leq [n^\beta]) \\
&\leq n^\beta \times \frac{1}{(\mu \wedge 2)} \sum_{j=1}^{\infty} j |\pi_j^n - \pi_j| \\
&\to 0 \qquad \text{as } n \to \infty,
\end{aligned}$$

by condition (b). [Note that in (3.4), $P(\theta_0^n = \theta_1^n) = \frac{1}{n}$, $P(\theta_1^n = \theta_2^n) = \sum_{l=1}^{n}(\bar{D}_l^n / n\mu_n)^2$ and application of the Cauchy–Schwarz inequality shows that $P(\theta_0^n = \theta_1^n) \leq P(\theta_1^n = \theta_2^n)$.] Equations (3.4) and (3.5) play a key role in the sequel and by the triangle inequality imply that for any $1 \leq k \leq [n^\beta]$, $P(1_{\{C_1^n \leq k\}} \neq 1_{\{B \leq k\}}) \to 0$ as $n \to \infty$.

4. Computing $\lim_{n \to \infty} \text{var}(\tilde{U}_n / \sqrt{n})$.

4.1. *Introduction.* We are now in position to start in earnest the proof of (2.5). Since all labellings of the vertices $\bar{\mathbf{D}}^n$ of the random graph $\mathcal{G}_n$ are exchangeable, we have that

(4.1)
$$\begin{aligned}
\text{var}(\tilde{U}_n / \sqrt{n}) &= \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \text{cov}(\chi_i^n, \chi_j^n) \\
&= \text{var}(\chi_1^n) + (n-1) \text{cov}(\chi_1^n, \chi_2^n).
\end{aligned}$$

Since the degree sequences $\bar{\mathbf{D}}^n$ are well behaved [Molloy and Reed (1995, 1998)],

(4.2) $\quad \text{var}(\chi_1^n) = E[\chi_1^n](1 - E[\chi_1^n]) \to \omega(1 - \omega)(= \rho(1 - \rho)) \qquad \text{as } n \to \infty.$

First, note that

(4.3)
$$\begin{aligned}
&(n-1)\text{cov}(\chi_1^n, \chi_2^n) \\
&= (n-1)E\big[1_{\{C_1^n \leq [n^\beta]\}}(1_{\{C_2^n \leq [n^\beta]\}} - E[1_{\{C_2^n \leq [n^\beta]\}}])\big] \\
&= (n-1)E\big[1_{\{2 \in C_1^n\}} 1_{\{C_1^n \leq [n^\beta]\}}(1_{\{C_2^n \leq [n^\beta]\}} - E[1_{\{C_2^n \leq [n^\beta]\}}])\big] \\
&\quad + (n-1)E\big[1_{\{2 \notin C_1^n\}} 1_{\{C_1^n \leq [n^\beta]\}}(1_{\{C_2^n \leq [n^\beta]\}} - E[1_{\{C_2^n \leq [n^\beta]\}}])\big].
\end{aligned}$$

The first term on the right-hand side of (4.3) is equal to

$$
(4.4) \quad \begin{aligned}
(n-1) & E\big[1_{\{2\in\mathcal{C}_1^n\}}1_{\{C_1^n\leq[n^\beta]\}}(1-E[1_{\{C_2^n\leq[n^\beta]\}}])\big] \\
&= (n-1)E[1_{\{2\in\mathcal{C}_1^n\}}1_{\{C_1^n\leq[n^\beta]\}}](1-E[\chi_2^n]).
\end{aligned}
$$

The following lemma provides the limit of the right-hand side of (4.4), and hence also of the first term on the right-hand side of (4.3), as $n\to\infty$. Note that $\{B<\infty\}$ is the event that the branching process $\mathcal{B}$ goes extinct.

LEMMA 4.1.

$$
(4.5) \quad (n-1)E[1_{\{2\in\mathcal{C}_1^n\}}1_{\{C_1^n\leq[n^\beta]\}}](1-E[\chi_2^n]) \to E[(B-1)1_{\{B<\infty\}}]\rho,
$$

as $n\to\infty$.

PROOF. Since $(1-E[\chi_2^n])\to\rho$ as $n\to\infty$, it suffices to show that

$$
(n-1)E[1_{\{2\in\mathcal{C}_1^n\}}1_{\{C_1^n\leq[n^\beta]\}}] \to E[(B-1)1_{\{B<\infty\}}] \qquad \text{as } n\to\infty.
$$

Note first that, by exchangeability, $P(2\in\mathcal{C}_1^n|C_1^n=k)=(k-1)/(n-1)$ ($k=1,2,\dots,n$). Hence,

$$
(4.6) \quad \begin{aligned}
(n-1)E[1_{\{2\in\mathcal{C}_1^n\}}1_{\{C_1^n\leq[n^\beta]\}}] &= (n-1)\sum_{k=1}^{[n^\beta]}P(2\in\mathcal{C}_1^n|C_1^n=k)P(C_1^n=k) \\
&= (n-1)\sum_{k=1}^{[n^\beta]}\frac{k-1}{n-1}P(C_1^n=k) \\
&= E\big[(C_1^n-1)1_{\{C_1^n\leq[n^\beta]\}}\big].
\end{aligned}
$$

Exploiting the couplings of $\mathcal{C}_1^n$, $\mathcal{B}_1^n$ and $\mathcal{B}$ as at the end of Section 3, we have that

$$
(4.7) \quad \begin{aligned}
\big|E\big[(C_1^n-1)&1_{\{C_1^n\leq[n^\beta]\}}\big] - E\big[(B-1)1_{\{B<[n^\beta]\}}\big]\big| \\
&\leq n^\beta P\big((C_1^n-1)1_{\{C_1^n\leq[n^\beta]\}}\neq(B-1)1_{\{B\leq[n^\beta]\}}\big) \\
&\leq n^\beta P\big(M_n^A\leq[n^\beta]\big)+n^\beta P\big(M_n^D\leq[n^\beta]\big) \\
&\leq n^\beta\Bigg(\frac{n^\beta(n^\beta+1)}{2}\sum_{i=1}^n\bigg(\frac{\bar{D}_i^n}{n\mu_n}\bigg)^2+n^\beta\times\frac{1}{(\mu\wedge2)}\sum_{j=1}^\infty j|\pi_j^n-\pi_j|\Bigg).
\end{aligned}
$$

By conditions (a)–(c), the right-hand side of (4.7) converges to 0 as $n\to\infty$.

Thus, the lemma follows by showing that

$$
\big|E\big[(B-1)1_{\{B\leq[n^\beta]\}}\big]-E\big[(B-1)1_{\{B<\infty\}}\big]\big|\to0 \qquad \text{as } n\to\infty.
$$

Note that

$$
\begin{aligned}
\big|E\big[(B-1)&1_{\{B\leq[n^\beta]\}}\big] - E\big[(B-1)1_{\{B<\infty\}}\big]\big| \\
&= E\big[(B-1)1_{\{[n^\beta]<B<\infty\}}\big] \\
&= \sum_{j=[n^\beta]+1}^{\infty}(j-1)P(B=j, B<\infty).
\end{aligned}
$$
(4.8)

By Markov's inequality $P(B=j, B<\infty) \leq E[B^3 1_{\{B<\infty\}}]/j^3$ and since the branching process $\mathcal{B}$ is super-critical, Lemma A.2 ensures that $E[B^3 1_{\{B<\infty\}}] < \infty$. Therefore, it follows that the right-hand side of (4.8) converges to 0 as $n \to \infty$ and the lemma is proved. $\square$

We turn to the second term on the right-hand side of (4.3). Let $\mathcal{H}^n$ denote a set of distinct vertices in $\mathcal{G}_n$. Then, conditioning on the cardinality C_1^n of $\mathcal{C}_1^n$ and exploiting exchangeability,

$$
(n-1)E\big[1_{\{2\notin\mathcal{C}_1^n\}}1_{\{C_1^n\leq[n^\beta]\}}\big(1_{\{C_2^n\leq[n^\beta]\}} - E[1_{\{C_2^n\leq[n^\beta]\}}]\big)\big]
$$

$$
= (n-1)\sum_{k=1}^{[n^\beta]}P(C_1^n=k)E\big[1_{\{2\notin\mathcal{C}_1^n\}}\big(1_{\{C_2^n\leq[n^\beta]\}} - E[1_{\{C_2^n\leq[n^\beta]\}}]\big)\big|C_1^n=k\big]
$$

$$
= (n-1)\sum_{k=1}^{[n^\beta]}P(C_1^n=k)\sum_{\mathcal{H}^n:|\mathcal{H}^n|=k}P(\mathcal{C}_1^n=\mathcal{H}^n|C_1^n=k)
$$

(4.9)
$$
\times E\big[1_{\{2\notin\mathcal{C}_1^n\}}\big(1_{\{C_2^n\leq[n^\beta]\}} - E[1_{\{C_2^n\leq[n^\beta]\}}]\big)\big|\mathcal{C}_1^n=\mathcal{H}^n\big]
$$

$$
= (n-1)\sum_{k=1}^{[n^\beta]}P(C_1^n=k)\sum_{\mathcal{H}^n:|\mathcal{H}^n|=k}P(\mathcal{C}_1^n=\mathcal{H}^n|C_1^n=k)
$$

$$
\times \big\{E\big[1_{\{2\notin\mathcal{C}_1^n\}}1_{\{C_2^n\leq[n^\beta]\}}\big|\mathcal{C}_1^n=\mathcal{H}^n\big] - E\big[1_{\{2\notin\mathcal{C}_1^n\}}\big|\mathcal{C}_1^n=\mathcal{H}^n\big]E[1_{\{C_2^n\leq[n^\beta]\}}]\big\}.
$$

We note that $|\mathcal{H}^n|=k$ implies that

$$
E\big[1_{\{2\notin\mathcal{C}_1^n\}}\big|\mathcal{C}_1^n=\mathcal{H}^n\big] = \frac{n-k}{n-1},
$$

which in turn gives

$$
E\big[1_{\{2\notin\mathcal{C}_1^n\}}1_{\{C_2^n\leq[n^\beta]\}}\big|\mathcal{C}_1^n=\mathcal{H}^n\big]
$$

(4.10)
$$
= P\big(C_2^n\leq[n^\beta]\big|2\notin\mathcal{C}_1^n, \mathcal{C}_1^n=\mathcal{H}^n\big) \times P\big(2\notin\mathcal{C}_1^n\big|\mathcal{C}_1^n=\mathcal{H}^n\big)
$$

$$
= \frac{n-k}{n-1}P\big(C_2^n\leq[n^\beta]\big|2\notin\mathcal{C}_1^n, \mathcal{C}_1^n=\mathcal{H}^n\big).
$$

Therefore, it follows from (4.9) and (4.10) that

$$(n-1)E\big[1_{\{2\notin\mathcal{C}_1^n\}}1_{\{C_1^n\leq[n^\beta]\}}\big(1_{\{C_2^n\leq[n^\beta]\}}-E[1_{\{C_2^n\leq[n^\beta]\}}]\big)\big]$$

$$(4.11)\qquad =\sum_{k=1}^{[n^\beta]}(n-k)P\big(C_1^n=k\big)\sum_{\mathcal{H}^n:|\mathcal{H}^n|=k}P\big(\mathcal{C}_1^n=\mathcal{H}^n|C_1^n=k\big)$$

$$\times\big\{P\big(C_2^n\leq[n^\beta]|2\notin\mathcal{C}_1^n,\mathcal{C}_1^n=\mathcal{H}^n\big)-P\big(C_2^n\leq[n^\beta]\big)\big\}.$$

The final term on the right-hand side of (4.11) is difficult to analyse directly. Therefore, we exploit further couplings of the component construction to the branching process approximation to derive the limiting behaviour of (4.11). In Section 4.2, we show that, as $n\to\infty$, the limit of (4.11) is the same as the limit of n times the difference between the extinction probabilities of two sequences of branching processes, the latter being far more amenable to analysis. Then in Section 4.3, we obtain the limit of the difference between the extinction probabilities.

4.2. *Recasting the limit of* (4.11). Let $\bar{\mathbf{D}}^n_{-\mathcal{H}^n}$ denote the set of vertices in $\mathcal{G}_n$ excluding those vertices in a given set of vertices, $\mathcal{H}^n$. Suppose that $2\notin\mathcal{H}^n$ and let $\check{\mathcal{C}}_2^n$ and $\check{\mathcal{B}}_2^n$ denote respectively the component containing vertex 2 and the associated branching process constructed using vertices $\bar{\mathbf{D}}^n_{-\mathcal{H}^n}$. For the present, $\mathcal{H}^n$ is an arbitrary but specified set of vertices that satisfies $|\mathcal{H}^n|\leq[n^\beta]$. Throughout the remainder of this section, $\mathcal{H}^n$ will be equated with $\mathcal{C}_1^n$ and we will construct $\check{\mathcal{C}}_2^n$ and the associated branching process $\check{\mathcal{B}}_2^n$ using the vertices not in $\mathcal{C}_1^n$. Let $\mathcal{C}_2^{n,I}\stackrel{D}{=}\mathcal{C}_2^n$ and $\mathcal{B}_2^{n,I}\stackrel{D}{=}\mathcal{B}_2^n$ denote the component containing vertex 2 and the associated branching process constructed using vertices $\bar{\mathbf{D}}^n$, independently of the construction of $\mathcal{C}_1^n$. Let $\check{C}_2^n$, $\check{B}_2^n$, $C_2^{n,I}$ and $B_2^{n,I}$ denote the cardinalities of the appropriate sets. We outline how the above four processes can be usefully coupled on a common probability space. Let $h_n=n-|\mathcal{H}^n|$, $\mu_n^H=\{n\mu_n-\sum_{i\in\mathcal{H}^n}\bar{D}_i^n\}/h_n$ and

$$(4.12)\qquad P\big(\check{\theta}^n=i\big)=\begin{cases}\dfrac{\bar{D}_i^n}{h_n\mu_n^H}&(i\in\{1,2,\ldots,n\}\setminus\mathcal{H}^n),\\0,&\text{otherwise.}\end{cases}$$

Let $\check{\theta}_0^n,\check{\theta}_1^n,\ldots$ be independent with $\check{\theta}_0^n$ distributed uniformly on $\{1,2,\ldots,n\}\setminus\mathcal{H}^n$ and $\check{\theta}_k^n\stackrel{D}{=}\check{\theta}^n$ for $k\geq1$. Then $\check{\mathcal{C}}_2^n$ and $\check{\mathcal{B}}_2^n$ can be constructed using $\bar{\mathbf{D}}^n_{-\mathcal{H}^n}$ and $\check{\theta}_0^n,\check{\theta}_1^n,\ldots$ in an analagous fashion to that described in Section 3. Let $Q_n^A(k)=\sum_{i=0}^{k-1}\sum_{j=i+1}^k 1_{\{\check{\theta}_i^n=\check{\theta}_j^n\}}$ be the total number of matches in $\{\check{\theta}_0^n,\check{\theta}_1^n,\ldots,\check{\theta}_k^n\}$. Then if $\check{B}_2^n=k$ and $Q_n^A(k-1)=0$, we have that $\check{C}_2^n=\check{B}_2^n$, that is, the addition of each new individual in the branching process corresponds to the addition of a new vertex to the component containing vertex 2. (Recall that $\check{B}_2^n$ includes

the initial ancestor.) A key observation going forward is that given a set $\mathcal{H}^n$, $\check{\mathcal{C}}_2^n \overset{D}{=} \mathcal{C}_2^n | \mathcal{C}_1^n = \mathcal{H}^n, 2 \notin \mathcal{C}_1^n$.

To couple $(\mathcal{B}_2^{n,I}, \mathcal{C}_2^{n,I})$ with $(\check{\mathcal{B}}_2^n, \check{\mathcal{C}}_2^n)$, we use independent (given $\mathcal{H}^n$) Bernoulli random variables $U_0^n, U_1^n, \ldots, U_{[n^\beta]+1}^n$. Let $P(U_0^n = 1 | \mathcal{H}^n) = |\mathcal{H}^n|/n$, the probability that a randomly chosen vertex belongs to $\mathcal{H}^n$ and for $k \geq 1$, let $P(U_k^n = 1 | \mathcal{C}_1^n) = \sum_{j \in \mathcal{H}^n} \bar{D}_j^n / (n \mu_n)$, the probability that a randomly chosen edge belongs to a vertex in $\mathcal{H}^n$. For $k \geq 0$, set $\theta_k^{n,I} = \check{\theta}_k^n$ if $U_k^n = 0$. If $U_0^n = 1$, draw $\theta_0^{n,I}$ uniformly at random from $\mathcal{H}^n$ and for $k \geq 1$, if $U_k^n = 1$, set $\theta_k^{n,I} = j$ with probability $\bar{D}_j^n / \sum_{i \in \mathcal{H}^n} \bar{D}_i^n$ ($j \in \mathcal{H}^n$). It is straightforward to show that $\theta_0^{n,I} \overset{D}{=} \theta_0^n$ and for $k \geq 1$, $\theta_k^{n,I} \overset{D}{=} \theta^n$. Then construct $(\mathcal{B}_2^{n,I}, \mathcal{C}_2^{n,I})$ as before using $\theta_0^{n,I}, \theta_1^{n,I}, \ldots$. Let $Q_n^B(k) = \sum_{i=0}^k U_i^n$, the total number of times in the first $k + 1$ chosen vertices, a vertex from $\mathcal{H}^n$ appears in the construction of $\mathcal{B}_2^{n,I}$. Note that if $\check{B}_2^n = k$ and $Q_n^B(k-1) = 0$ then $B_2^{n,I} = \check{B}_2^n$. Also if $\check{C}_2^n = k$, $Q_n^A(k) \leq 1$ and $Q_n^B(k) = 0$, then $C_2^{n,I} = \check{C}_2^n$, since we need to take into account that there may exist $0 \leq i < j \leq k$ such that $\check{\theta}_i^n = \check{\theta}_j^n$. [If $\check{C}_2^n = k$ and there is at most one match among $\{\check{\theta}_0^n, \check{\theta}_1^n, \ldots, \check{\theta}_k^n\}$ then $\{\check{\theta}_0^n, \check{\theta}_1^n, \ldots, \check{\theta}_k^n\}$ determines $\check{\mathcal{C}}_2^n$. It then follows that $\mathcal{C}_2^{n,I} = \check{\mathcal{C}}_2^n$ (whence $C_2^{n,I} = \check{C}_2^n$), since $Q_n^B(k) = 0$ implies $\theta_j^{n,I} = \check{\theta}_j^n$ ($j = 0, 1, \ldots, k$).] Finally, if $B_2^{n,I} = k$ with $Q_n^A(k-1) = 0$ and $Q_n^B(k-1) \leq 1$ then $C_2^{n,I} = B_2^{n,I}$, since $Q_n^A(k-1) = 0$ and $Q_n^B(k-1) \leq 1$ ensure that all the individuals in $\mathcal{B}_2^{n,I}$ correspond to distinct vertices in $\mathcal{C}_2^{n,I}$. [If $Q_n^A(k-1) = Q_n^B(k-1) = 0$ then $\theta_0^{n,I}, \theta_1^{n,I}, \ldots, \theta_{k-1}^{n,I}$ are clearly distinct. If $Q_n^A(k-1) = 0$ and $Q_n^B(k-1) = 1$, then there exists $0 \leq j_0 \leq k-1$ such that $\theta_{j_0}^{n,I} \in \mathcal{H}^n$ and $\theta_j^{n,I} = \check{\theta}_j^n$ for $j = 0, 1, \ldots, k-1$ ($j \neq j_0$). Now $\check{\theta}_j^n \notin \mathcal{H}^n$ ($j = 0, 1, \ldots, k-1$), so since $\check{\theta}_0^n, \check{\theta}_1^n, \ldots, \check{\theta}_{k-1}^n$ are distinct then so are $\theta_0^{n,I}, \theta_1^{n,I}, \ldots, \theta_{k-1}^{n,I}$.]

LEMMA 4.2.

$$\lim_{n \to \infty} (n-1) E\big[1_{\{2 \notin \mathcal{C}_1^n\}} 1_{\{C_1^n \leq [n^\beta]\}} \big(1_{\{C_2^n \leq [n^\beta]\}} - E[1_{\{C_2^n \leq [n^\beta]\}}]\big)\big]$$

$$= \lim_{n \to \infty} (n-1) E\big[1_{\{2 \notin \mathcal{C}_1^n\}} 1_{\{C_1^n \leq [n^\beta]\}} \big(1_{\{C_2^n \leq [n^\beta]\}} - E[1_{\{C_2^{n,I} \leq [n^\beta]\}}]\big)\big]$$

(4.13)

$$= \lim_{n \to \infty} E\big[1_{\{C_1^n \leq [n^\beta]\}} 1_{\{2 \notin \mathcal{C}_1^n\}}$$

$$\times (n-1)\{E[1_{\{\check{B}_2^n \leq [n^\beta]\}} | \mathcal{H}^n = \mathcal{C}_1^n] - E[1_{\{B_2^{n,I} \leq [n^\beta]\}}]\}\big],$$

provided that the final limit exists and is finite.

PROOF. The first equality follows from $E[1_{\{C_2^{n,I} \leq [n^\beta]\}}] = E[1_{\{C_2^n \leq [n^\beta]\}}]$.

Using (4.11) and a similar decomposition for the right-hand side of (4.13), we have that

$$\left|(n-1)E\left[1_{\{2\notin\mathcal{C}_1^n\}}1_{\{C_1^n\leq[n^\beta]\}}(1_{\{C_2^n\leq[n^\beta]\}}-E[1_{\{C_2^{n,I}\leq[n^\beta]\}}])\right]\right.$$

$$-(n-1)E\left[1_{\{C_1^n\leq[n^\beta]\}}1_{\{2\notin\mathcal{C}_1^n\}}\right.$$

$$\times\{E[1_{\{\check{B}_2^n\leq[n^\beta]\}}|\mathcal{H}^n=\mathcal{C}_1^n]-E[1_{\{B_2^{n,I}\leq[n^\beta]\}}]\}]\Big|$$

(4.14)
$$=\left|\sum_{k=1}^{[n^\beta]}(n-k)P(C_1^n=k)\sum_{\mathcal{H}^n:|\mathcal{H}^n|=k}P(\mathcal{C}_1^n=\mathcal{H}^n|C_1^n=k)\right.$$

$$\times(\{P(C_2^n\leq[n^\beta]|2\notin\mathcal{C}_1^n,\mathcal{C}_1^n=\mathcal{H}^n)-P(C_2^{n,I}\leq[n^\beta])\}$$

$$\left.-\{P(\check{B}_2^n\leq[n^\beta]|\mathcal{H}^n)-P(B_2^{n,I}\leq[n^\beta])\})\right|.$$

Note that $\{\check{C}_2^n|\mathcal{H}^n=\mathcal{C}_1^n\}\stackrel{D}{=}\{C_2^n|2\notin\mathcal{C}_1^n,\mathcal{C}_1^n=\mathcal{H}^n\}$, from the construction of $\check{C}_2^n$, so

$$P(C_2^n|2\notin\mathcal{C}_1^n,\mathcal{C}_1^n=\mathcal{H}^n)=P(\check{C}_2^n|\mathcal{H}^n=\mathcal{C}_1^n).$$

Therefore, it follows that the right-hand side of (4.14) is less than or equal to

(4.15)
$$\max_{\{\mathcal{H}^n;|\mathcal{H}^n|\leq n^\beta\}}(n-1)|\{P(\check{C}_2^n\leq[n^\beta]|\mathcal{H}^n)-P(C_2^{n,I}\leq[n^\beta])\}$$
$$-\{P(\check{B}_2^n\leq[n^\beta]|\mathcal{H}^n)-P(B_2^{n,I}\leq[n^\beta])\}|.$$

We start by considering fixed $\mathcal{H}^n$ with $|\mathcal{H}^n|\leq[n^\beta]$.

Using the above couplings, we have that, for $k=1,2,\ldots,[n^\beta]$,

$$P(\check{C}_2^n=k|\mathcal{H}^n)=P(\check{B}_2^n=k,Q_n^A([n^\beta])=0,Q_n^B([n^\beta])\leq1|\mathcal{H}^n)$$
$$+P(C_2^{n,I}=k,Q_n^A([n^\beta])=1,Q_n^B([n^\beta])=0|\mathcal{H}^n)$$
$$+P(\check{C}_2^n=k,Q_n^A([n^\beta])+Q_n^B([n^\beta])\geq2|\mathcal{H}^n).$$

Since the distribution of $C_2^{n,I}$ is independent of $\mathcal{H}^n$, we have that, for $k=1,2,\ldots,[n^\beta]$,

$$P(C_2^{n,I}=k)=P(C_2^{n,I}=k|\mathcal{H}^n)$$
$$=P(B_2^{n,I}=k,Q_n^A([n^\beta])=0,Q_n^B([n^\beta])\leq1|\mathcal{H}^n)$$
$$+P(C_2^{n,I}=k,Q_n^A([n^\beta])=1,Q_n^B([n^\beta])=0|\mathcal{H}^n)$$
$$+P(C_2^{n,I}=k,Q_n^A([n^\beta])+Q_n^B([n^\beta])\geq2|\mathcal{H}^n).$$

Also, for $k = 1, 2, \ldots, [n^\beta]$,

$$
\begin{aligned}
P(\check{B}_2^n = k | \mathcal{H}^n) = {} & P(\check{B}_2^n = k, Q_n^A([n^\beta]) = 0, Q_n^B([n^\beta]) \le 1 | \mathcal{H}^n) \\
& + P(B_2^{n,I} = k, Q_n^A([n^\beta]) = 1, Q_n^B([n^\beta]) = 0 | \mathcal{H}^n) \\
& + P(\check{B}_2^n = k, Q_n^A([n^\beta]) + Q_n^B([n^\beta]) \ge 2 | \mathcal{H}^n).
\end{aligned}
$$

Therefore, using a similar decomposition for $P(B_2^{n,I} = k)$, it is straightforward to show that

$$
\begin{aligned}
(4.16) \quad & (n-1) | \{ P(\check{C}_2^n \le [n^\beta] | \mathcal{H}^n) - P(C_2^{n,I} \le [n^\beta]) \} \\
& \quad - \{ P(\check{B}_2^n \le [n^\beta] | \mathcal{H}^n) - P(B_2^{n,I} \le [n^\beta]) \} | \\
& \quad \le 2(n-1) P(Q_n^A([n^\beta]) + Q_n^B([n^\beta]) \ge 2 | \mathcal{H}^n).
\end{aligned}
$$

We study the right-hand side of (4.16) in order to bound the right-hand side of (4.14).

Given $\mathcal{H}^n$, we have that $Q_n^A([n^\beta])$ and $Q_n^B([n^\beta])$ are independent. Therefore, the right-hand side of (4.16) is bounded above by

$$
\begin{aligned}
(4.17) \quad & 2\{ (n-1) P(Q_n^A([n^\beta]) \ge 2 | \mathcal{H}^n) + (n-1) P(Q_n^B([n^\beta]) \ge 2 | \mathcal{H}^n) \\
& \quad + (n-1) P(Q_n^A([n^\beta]) = 1 | \mathcal{H}^n) P(Q_n^A([n^\beta]) = 1 | \mathcal{H}^n) \}.
\end{aligned}
$$

We are interested in the case $\mathcal{H}^n = \mathcal{C}_1^n$, where $|C_1^n| \le [n^\beta]$. Under condition (c), for all sufficiently large n, $\Delta_n = \max_{\{1 \le i \le n\}} \bar{D}_i^n \le n^{\frac{1}{4}}$, and consequently, we have that, for $k \ge 0$, $|\sum_{i \in \mathcal{H}^n} (\bar{D}_i^n)^k| \le n^\beta n^{\frac{k}{4}}$. Since $\beta < 1/12$, it follows from condition (a)(ii) that, for $k = 1, 2, 3$, $n^{-1} \sum_{i \notin \mathcal{H}^n} (\bar{D}_i^n)^k > \frac{1}{2} L_k$ and $n^{-1} \sum_{i=1}^n (\bar{D}_i^n)^k \le 2L_k$, for all sufficiently large n.

For $Q_n^A([n^\beta]) \ge 2$, we require that either there exist distinct $i, j, k \in \{0, 1, \ldots, [n^\beta]\}$ such that $\check{\theta}_i^n = \check{\theta}_j^n = \check{\theta}_k^n$ or there exist distinct $i, j, k, l \in \{0, 1, \ldots, [n^\beta]\}$ such that $\check{\theta}_i^n = \check{\theta}_j^n$ and $\check{\theta}_k^n = \check{\theta}_l^n$. Without loss of generality, we may assume that $i < j, k$ in the first case and $i < j, k, l$ in the second case. There are therefore two cases to consider $i = 0$ and $i \ge 1$. For $i = 0$,

$$
\begin{aligned}
P(\check{\theta}_0^n = \check{\theta}_j^n = \check{\theta}_k^n | \mathcal{H}^n) &= \sum_{l \notin \mathcal{H}^n} \left\{ \frac{1}{n - |\mathcal{H}^n|} \times \left(\frac{\bar{D}_l^n}{\sum_{m \notin \mathcal{H}^n} \bar{D}_m^n} \right)^2 \right\} \\
&\le \frac{8}{L_1^2 n^3} \sum_{l=1}^n (\bar{D}_l^n)^2 \le \frac{16 L_2}{L_1^2 n^2},
\end{aligned}
$$

for all sufficiently large n. For $i \ge 1$,

$$
\begin{aligned}
P(\check{\theta}_i^n = \check{\theta}_j^n = \check{\theta}_k^n | \mathcal{H}^n) &= \sum_{l \notin \mathcal{H}^n} \left(\frac{\bar{D}_l^n}{\sum_{m \notin \mathcal{H}^n} \bar{D}_m^n} \right)^3 \\
&\le \frac{8}{L_1^3 n^3} \sum_{l=1}^n (\bar{D}_l^n)^3 \le \frac{16 L_3}{L_1^3 n^2},
\end{aligned}
$$

for all sufficiently large n. Using the independence of $\check{\theta}_i^n, \check{\theta}_j^n, \check{\theta}_k^n, \check{\theta}_l^n$, we have that for $i \geq 1$,

$$P(\check{\theta}_i^n = \check{\theta}_j^n, \check{\theta}_k^n = \check{\theta}_l^n | \mathcal{H}^n) = P(\check{\theta}_i^n = \check{\theta}_j^n | \mathcal{H}^n) P(\check{\theta}_k^n = \check{\theta}_l^n | \mathcal{H}^n)$$

$$= \left(\sum_{r \notin \mathcal{H}^n} \left(\frac{\bar{D}_r^n}{\sum_{m \notin \mathcal{H}^n} \bar{D}_m^n} \right)^2 \right)^2$$

$$\leq \frac{16}{L_1^4 n^4} \left(\sum_{r=1}^n (\bar{D}_r^n)^2 \right)^2 \leq \frac{64 L_2^2}{L_1^4 n^2},$$

for all sufficiently large n, and by similar arguments for $i = 0$,

$$P(\check{\theta}_0^n = \check{\theta}_j^n, \check{\theta}_k^n = \check{\theta}_l^n | \mathcal{H}^n) \leq \frac{64 L_2}{L_1^2 n^2},$$

for all sufficiently large n. Since $\beta < 1/12$ and the bound for $i \geq 1$ is the larger in both cases,

$$(n-1) P(Q_n^A([n^\beta]) \geq 2 | \mathcal{H}^n)$$

(4.18)
$$\leq \frac{n-1}{n^2} \left\{ ([n^\beta] + 1)^3 \frac{16 L_3}{L_1^3} + ([n^\beta] + 1)^4 \frac{64 L_2^2}{L_1^4} \right\}$$

$$\rightarrow 0 \qquad \text{as } n \rightarrow \infty.$$

Similarly, for all sufficiently large n,

$$\sqrt{n-1} P(Q_n^A([n^\beta]) = 1 | \mathcal{H}^n) \leq \frac{\sqrt{n-1}([n^\beta] + 1)^2}{n} \frac{8 L_2}{L_1^2}$$

(4.19)
$$\rightarrow 0 \qquad \text{as } n \rightarrow \infty.$$

Now turning to $Q_n^B([n^\beta])$, we have that

$$(n-1) P(Q_n^B([n^\beta]) \geq 2 | \mathcal{H}^n)$$

$$\leq \sum_{i=0}^{[n^\beta - 1]} \sum_{j=i+1}^{[n^\beta]} (n-1) P(U_i^n = 1, U_j^n = 1 | \mathcal{H}^n)$$

$$= \sum_{i=0}^{[n^\beta - 1]} \sum_{j=i+1}^{[n^\beta]} (n-1) P(U_i^n = 1 | \mathcal{H}^n) P(U_j^n = 1 | \mathcal{H}^n)$$

(4.20)
$$\leq (n-1) \left\{ [n^\beta] \frac{|\mathcal{H}^n|}{n} \frac{\sum_{j \in \mathcal{H}^n} \bar{D}_j^n}{n \mu_n} + [n^\beta]^2 \left(\frac{\sum_{j \in \mathcal{H}^n} \bar{D}_j^n}{n \mu_n} \right)^2 \right\}$$

$$\leq (n-1) \left\{ \frac{n^{3\beta} \Delta_n}{n^2 \mu / 2} + n^{2\beta} \frac{(n^\beta \Delta_n)^2}{n^2 \mu^2 / 4} \right\} \qquad \text{(for all sufficiently large } n)$$

$$\leq \frac{n^{4\beta + \frac{1}{2}}}{n} \left(\frac{2}{\mu} + \frac{4}{\mu^2} \right) \rightarrow 0 \qquad \text{as } n \rightarrow \infty.$$

Similarly, for all sufficiently large n, we have that

$$\sqrt{n-1}\, P\big(Q_n^B([n^\beta])=1||\mathcal{H}^n|\leq n^\beta\big) \leq \sqrt{n-1}\Big(\frac{|\mathcal{H}^n|}{n}+[n^\beta]\sum_{j\in\mathcal{H}^n}\bar{D}_j^n/n\mu_n\Big)$$

(4.21)
$$\leq \sqrt{n-1}\Big(\frac{n^\beta}{n}+\frac{2n^{2\beta+\frac{1}{4}}}{n\mu}\Big)$$

$$\to 0 \qquad \text{as } n\to\infty.$$

Therefore, it follows from (4.18)–(4.21) that for any $\mathcal{H}^n$ satisfying $|\mathcal{H}^n|\leq [n^\beta]$, (4.17) converges to 0 as $n\to\infty$; moreover, this convergence is uniform over such $\mathcal{H}^n$. Hence, the right-hand side of (4.14) converges to 0 as $n\to\infty$ and the lemma follows. $\square$

For a branching process, it is far simpler to study its extinction probability than the probability that its total size is less than $[n^\beta]$. The following lemma provides useful bounds between these two probabilities as $n\to\infty$.

LEMMA 4.3.

(4.22)
$$(n-1)E\big[1_{\{C_1^n\leq[n^\beta]\}}1_{\{2\notin\mathcal{C}_1^n\}}\big(E\big[1_{\{\check{B}_2^n<\infty\}}|\mathcal{H}^n=\mathcal{C}_1^n\big]$$
$$- E\big[1_{\{\check{B}_2^n\leq[n^\beta]\}}|\mathcal{H}^n=\mathcal{C}_1^n\big]\big)\big]\to 0$$

and

(4.23) $\quad (n-1)E\big[1_{\{C_1^n\leq[n^\beta]\}}1_{\{2\notin\mathcal{C}_1^n\}}\big(E\big[1_{\{B_2^{n,I}<\infty\}}\big]-E\big[1_{\{B_2^{n,I}\leq[n^\beta]\}}\big]\big)\big]\to 0$

as $n\to\infty$.

PROOF. We prove (4.22) with (4.23) following by similar but simpler arguments.

Given that $C_1^n\leq[n^\beta]$, it follows from conditions (a)–(c) that there exists $\delta>0$ such that for all sufficiently large n, regardless of the set $\mathcal{C}_1^n$, $\check{\nu}_n=\sum_{i\notin\mathcal{C}_1^n}\bar{D}_i^n(\bar{D}_i^n-1)/\sum_{i\notin\mathcal{C}_1^n}\bar{D}_i^n>1+\delta$ and $\check{\lambda}_n=\sum_{i\notin\mathcal{C}_1^n}\bar{D}_i^n(\bar{D}_i^n-1)(\bar{D}_i^n-2)/\sum_{i\notin\mathcal{C}_1^n}\bar{D}_i^n<\lambda+1<\infty$. Therefore, it follows by Corollary A.4 that for any $l\in\mathbb{N}$, there exists a finite constant $A_l^*=A_{l,\lambda+1,\delta}$ such that for sufficiently large n, $E[(\check{B}_2^n)^l1_{\{\check{B}_2^n<\infty\}}]\leq A_l^*$.

By Markov's inequality for any $l\in\mathbb{N}$ and for all sufficiently large n,

$$E\big[1_{\{\check{B}_2^n<\infty\}}|\mathcal{H}^n=\mathcal{C}_1^n\big]-E\big[1_{\{\check{B}_2^n\leq[n^\beta]\}}|\mathcal{H}^n=\mathcal{C}_1^n\big]$$

$$= P\big(\check{B}_2^n<\infty\big)-P\big(\check{B}_2^n\leq[n^\beta],\check{B}_2^n<\infty\big)$$

(4.24)
$$= P\big(\check{B}_2^n>[n^\beta]|\check{B}_2^n<\infty\big)P\big(\check{B}_2^n<\infty\big)$$

$$\leq [n^\beta]^{-l}E\big[(\check{B}_2^n)^l|\check{B}_2^n<\infty\big]P\big(\check{B}_2^n<\infty\big)$$

$$\leq [n^\beta]^{-l}A_l^*.$$

By choosing any $l > 1/\beta$, we can bound from above the left-hand side of (4.22) by $(n - 1)[n^\beta]^{-l} A_l^* \to 0$ as $n \to \infty$. $\square$

By Lemma 4.3 and the triangle inequality, we have that

$$
\begin{aligned}
(4.25) \quad & \lim_{n \to \infty} E\big[1_{\{C_1^n \leq [n^\beta]\}} 1_{\{2 \notin C_1^n\}} \\
& \quad \times (n - 1)\{E[1_{\{\check{B}_2^n \leq [n^\beta]\}} | \mathcal{H}^n = \mathcal{C}_1^n] - E[1_{\{B_2^{n,I} \leq [n^\beta]\}}]\}\big] \\
& = \lim_{n \to \infty} E\big[1_{\{C_1^n \leq [n^\beta]\}} 1_{\{2 \notin C_1^n\}} \\
& \quad \times (n - 1)\{E[1_{\{\check{B}_2^n < \infty\}} | \mathcal{H}^n = \mathcal{C}_1^n] - E[1_{\{B_2^{n,I} < \infty\}}]\}\big],
\end{aligned}
$$

should the latter limit exist. Thus, we have reformulated the limit, as $n \to \infty$, of (4.11) in terms of the limiting behaviour of the extinction probabilities of two sequences of branching processes. Moreover, there is a straightforward coupling between the two branching processes which is exploited in Section 4.3.

4.3. *Comparing extinction probabilities of coupled branching processes.* Let $\check{y}_n = P(\check{B}_2^n < \infty | \mathcal{H}^n = \mathcal{C}_1^n, C_1^n \leq [n^\beta])$ and $y_n = P(B_2^{n,I} < \infty)$ denote the extinction probabilities of the branching processes $\check{\mathcal{B}}_2^n$ and $\mathcal{B}_2^{n,I}$, respectively. Then $\check{y}_n = \check{f}_n(\check{z}_n)$ and $y_n = f_n(z_n)$, where for $s \geq 0$, $\check{f}_n(s) = \frac{1}{n - C_1^n} \sum_{j \notin \mathcal{C}_1^n} s^{\bar{D}_j^n}$, $f_n(s) = \frac{1}{n} \sum_{j=1}^n s^{\bar{D}_j^n} = \sum_{k=0}^\infty \pi_k^n s^k$, and $\check{z}_n$ and z_n satisfy

$$
\begin{aligned}
(4.26) \quad & \check{z}_n = \frac{1}{\check{\mu}_n} \check{f}_n'(\check{z}_n), \\
& z_n = \frac{1}{\mu_n} f_n'(z_n)
\end{aligned}
$$

with $\check{\mu}_n = \frac{1}{n - C_1^n} \sum_{j \notin \mathcal{C}_1^n} \bar{D}_j^n$, $\check{f}_n'(s) = \frac{d}{ds} \check{f}_n(s)$ and $f_n'(s) = \frac{d}{ds} f_n(s)$. Note that for all $0 \leq s \leq 1$, $\check{f}_n(s) | \{C_1^n \leq [n^\beta]\}, f_n(s) \xrightarrow{P} f(s)$, $\check{f}_n'(s) | \{C_1^n \leq [n^\beta]\}, f_n'(s) \xrightarrow{P} f'(s)$, $\check{\mu}_n | \{C_1^n \leq [n^\beta]\}, \mu_n \xrightarrow{P} \mu$ and $\check{v}_n | \{C_1^n \leq [n^\beta]\}, v_n \xrightarrow{P} v$ as $n \to \infty$, where $\check{v}_n = \frac{1}{(n - C_1^n)\check{\mu}_n} \sum_{j \notin \mathcal{C}_1^n} \bar{D}_j^n(\bar{D}_j^n - 1)$. Then, using Britton, Janson and Martin-Löf (2007), Lemma 4.1, it is straightforward to show that $\check{z}_n | \{C_1^n \leq [n^\beta]\}, z_n \xrightarrow{P} z$ as $n \to \infty$.

For $k = 0, 1, \ldots$, let b_k^n denote the total number of vertices in $\mathcal{C}_1^n$ with degree k. Let $H_1^n = \sum_{k=1}^\infty k b_k^n$, the sum of the degrees of the members of $\mathcal{C}_1^n$. If $\mathcal{C}_1^n$ contains no cycles, then $H_1^n = 2(C_1^n - 1)$. Using M_n^A and $M_n^{A,2} = \min\{k > M_n^A; \theta_k^n \in \{\theta_0^n, \theta_1^n, \ldots, \theta_{k-1}^n\}\}$, it is straightforward to show that $P(H_1^n = 2(C_1^n - 1) | C_1^n \leq [n^\beta]) \to 1$ and there exists $\varepsilon > 0$ such that $n^{1+\varepsilon} P(H_1^n > 2C_1^n | C_1^n \leq [n^\beta]) \to 0$ as $n \to \infty$ [cf. (3.4) and (4.18), resp.]. In words, if $\mathcal{C}_1^n$ is small ($\leq [n^\beta]$) then with high

probability it contains no cycles and the probability of containing 2 or more cycles is very small. For $k = 0, 1, \ldots$, let $\check{\pi}_k^n = \frac{1}{n - C_1^n} \sum_{j \notin C_1^n} 1_{\{\check{D}_j^n = k\}}$, so for $0 \le s \le 1$, $\check{f}_n(s) = \sum_{k=0}^{\infty} \check{\pi}_k^n s^k$.

The limit, as $n \to \infty$, of the right-hand side of (4.25) is computed in Lemma 4.6 following a preliminary lemma and corollary. A key observation in obtaining the limit is that, using the coupling of C_1^n, B_1^n and $\mathcal{B}$, it is easily shown that, for $k = 0, 1, \ldots$,

$$(4.27) \qquad b_k^n | \{C_1^n \le [n^\beta]\} \xrightarrow{D} b_k | \{B < \infty\} \qquad \text{as } n \to \infty,$$

where b_k denotes the total number of individuals having degree k in the branching process $\mathcal{B}$, where an individual has degree k if it has $k - 1$ offspring (k offspring in the case of the initial ancestor).

LEMMA 4.4.

$$n(\check{z}_n - z_n) | \{C_1^n \le [n^\beta]\} \xrightarrow{D} \frac{1}{1 - f''(z)/\mu} \times \frac{1}{\mu} \sum_{k=0}^{\infty} k b_k (z - z^{k-1}) \Big| \{B < \infty\}$$

as $n \to \infty$.

PROOF. First, note that

$$n(\check{z}_n - z_n) = n \left\{ \frac{1}{\check{\mu}_n} \sum_{k=1}^{\infty} k \check{\pi}_k^n \check{z}_n^{k-1} - \frac{1}{\mu_n} \sum_{k=1}^{\infty} k \pi_k^n z_n^{k-1} \right\}$$

$$(4.28) \qquad = \frac{n}{\check{\mu}_n \mu_n} (\mu_n - \check{\mu}_n) \sum_{k=1}^{\infty} k \check{\pi}_k^n \check{z}_n^{k-1} + \frac{n}{\mu_n} \sum_{k=1}^{\infty} k (\check{\pi}_k^n - \pi_k^n) \check{z}_n^{k-1}$$

$$+ \frac{n}{\mu_n} \sum_{k=1}^{\infty} k \pi_k^n (\check{z}_n^{k-1} - z_n^{k-1}).$$

Since $C_1^n | \{C_1^n \le [n^\beta]\} \xrightarrow{D} B | \{B < \infty\}$ and $H_1^n | \{C_1^n \le [n^\beta]\} \xrightarrow{D} 2(B - 1) | \{B < \infty\}$ as $n \to \infty$, we have that

$$n(\mu_n - \check{\mu}_n) | \{C_1^n \le n^\beta\} = \left(-\frac{C_1^n}{n - C_1^n} \right) \sum_{i=1}^{n} \bar{D}_i^n + \frac{n}{n - C_1^n} H_1^n \Big| \{C_1^n \le n^\beta\}$$

$$\xrightarrow{D} -B\mu + 2(B - 1) | \{B < \infty\} \qquad \text{as } n \to \infty.$$

Then, since $\check{\mu}_n | \{C_1^n \le n^\beta\}$, $\mu_n \xrightarrow{P} \mu$, $\check{z}_n | \{C_1^n \le n^\beta\} \xrightarrow{P} z$ and $\check{f}_n'(\check{z}_n) | \{C_1^n \le n^\beta\} \xrightarrow{P} f'(z)$ as $n \to \infty$, where z and $f'(z)$ are defined in (1.3), we have that

$$\frac{n}{\check{\mu}_n \mu_n} (\mu_n - \check{\mu}_n) \sum_{k=1}^{\infty} k \check{\pi}_k^n \check{z}_n^{k-1} \Big| \{C_1^n \le n^\beta\}$$

$$(4.29)$$

$$\xrightarrow{D} \frac{1}{\mu^2} \{2(B - 1) - B\mu\} f'(z) | \{B < \infty\} \qquad \text{as } n \to \infty.$$

Similarly,

$$\frac{n}{\mu_n}\sum_{k=1}^{\infty}k(\check{\pi}_k^n-\pi_k^n)\check{z}_n^{k-1}\Big|\{C_1^n\leq n^{\beta}\}$$

$$=\frac{n}{\mu_n}\sum_{k=1}^{\infty}k\Bigg(\frac{\sum_{j\notin\mathcal{C}_1^n}1_{\{\bar{D}_j^n=k\}}}{n-C_1^n}-\frac{\sum_{j=1}^{n}1_{\{\bar{D}_j^n=k\}}}{n}\Bigg)\check{z}_n^{k-1}\Big|\{C_1^n\leq n^{\beta}\}$$

$$=\frac{1}{\mu_n}\sum_{k=1}^{\infty}k\Bigg(\sum_{j\notin\mathcal{C}_1^n}1_{\{\bar{D}_j^n=k\}}\bigg(\frac{n}{n-C_1^n}-\frac{n}{n}\bigg)-\sum_{j\in\mathcal{C}_1^n}1_{\{\bar{D}_j^n=k\}}\Bigg)\check{z}_n^{k-1}\Big|\{C_1^n\leq n^{\beta}\}$$

(4.30)

$$=\frac{1}{\mu_n}\sum_{k=1}^{\infty}k\Bigg\{\bigg(\frac{1}{n-C_1^n}\sum_{j\notin\mathcal{C}_1^n}1_{\{\bar{D}_j^n=k\}}\bigg)C_1^n-b_k^n\Bigg\}\check{z}_n^{k-1}\Big|\{C_1^n\leq n^{\beta}\}$$

$$\xrightarrow{D}\frac{1}{\mu}\sum_{k=1}^{\infty}k(\pi_k B-b_k)z^{k-1}\Big|\{B<\infty\}$$

$$=\frac{Bf'(z)}{\mu}-\frac{1}{\mu}\sum_{k=1}^{\infty}kb_k z^{k-1}\qquad\text{as }n\to\infty.$$

(Convergence in distribution of the infinite sum can be justified by exploiting that, for any $z_0\in(0,1)$, $\sum_{k=k_0}^{\infty}kz^{k-1}\to 0$ as $k_0\to\infty$ uniformly in $z\in[0,z_0]$.)

Turning to the third term on the right-hand side of (4.28), by the mean value theorem, for fixed n, there exists φ_n lying between z_n and $\check{z}_n$ such that

$$(4.31)\qquad\frac{n}{\mu_n}\sum_{k=1}^{\infty}k\pi_k^n(\check{z}_n^{k-1}-z_n^{k-1})=(\check{z}_n-z_n)\frac{n}{\mu_n}\sum_{k=1}^{\infty}k(k-1)\pi_k^n\varphi_n^{k-2}.$$

Now $\check{z}_n|\{C_1^n\leq n^{\beta}\}$ and z_n both converge in probability to z as $n\to\infty$, hence so does φ_n. It then follows that $\frac{1}{\mu_n}\sum_{k=1}^{\infty}k(k-1)\pi_k^n\varphi_n^{k-2}|\{C_1^n\leq n^{\beta}\}\xrightarrow{P}\frac{1}{\mu}f''(z)$ as $n\to\infty$. As noted in the paragraph following Theorem 2.1, $f''(z)/\mu<1$. Therefore combining (4.29), (4.30) and (4.31) with (4.28), we have that

$$n(\check{z}_n-z_n)|\{C_1^n\leq n^{\beta}\}$$

$$\xrightarrow{D}\frac{1}{1-f''(z)/\mu}\Bigg\{\frac{f'(z)}{\mu^2}\{2(B-1)\}-\frac{1}{\mu}\sum_{k=1}^{\infty}kb_k z^{k-1}\Bigg\}\Big|\{B<\infty\}$$

$$=\frac{1}{1-f''(z)/\mu}\times\frac{1}{\mu}\sum_{k=1}^{\infty}kb_k(z-z^{k-1})\Big|\{B<\infty\}\qquad\text{as }n\to\infty,$$

since $z=f'(z)/\mu$ and $2(B-1)=\sum_{k=1}^{\infty}kb_k$. $\square$

COROLLARY 4.5.

$$n(\check{y}_n - y_n)|\{\mathcal{H}^n = \mathcal{C}_1^n, C_1^n \le [n^\beta]\}$$

$$(4.32) \quad \xrightarrow{D} \left\{ Bf(z) - \sum_{k=1}^{\infty} b_k z^k + \frac{z}{1 - f''(z)/\mu} \times \sum_{k=1}^{\infty} k b_k (z - z^{k-1}) \right\} \Big|\{B < \infty\}$$

$$as\ n \to \infty.$$

PROOF. Note that

$$(4.33) \quad n(\check{y}_n - y_n) = n(\check{f}_n(\check{z}_n) - f_n(\check{z}_n)) + n(f_n(\check{z}_n) - f_n(z_n)).$$

It is straightforward using a similar argument to (4.30) to show that

$$n(\check{f}_n(\check{z}_n) - f_n(\check{z}_n))|\{C_1^n \le [n^\beta]\}$$

$$(4.34) \qquad = n \sum_{k=1}^{\infty} (\check{\pi}_k^n - \pi_k^n) \check{z}_n^k \Big|\{C_1^n \le [n^\beta]\}$$

$$\xrightarrow{D} \sum_{k=1}^{\infty} (B\pi_k - b_k) z^k \Big|\{B < \infty\} \qquad as\ n \to \infty.$$

By the mean value theorem, there exists $\varphi_n^2 \xrightarrow{P} z$ as $n \to \infty$, such that

$$(4.35) \qquad n(f_n(\check{z}_n) - f_n(z_n)) = f_n'(\varphi_n^2) n(\check{z}_n - z_n).$$

The corollary then follows by substituting (4.34) and (4.35) into (4.33), and then using Lemma 4.4 and $z = f'(z)/\mu$. [It is easily shown that $f_n'(\varphi_n^2) \xrightarrow{P} f'(z)$ as $n \to \infty$.] $\square$

LEMMA 4.6.

$$E\big[1_{\{C_1^n \le [n^\beta]\}} 1_{\{2 \notin \mathcal{C}_1^n\}} \times (n-1)\{E[1_{\{\check{B}_2^n < \infty\}}|\mathcal{H}^n = \mathcal{C}_1^n] - E[1_{\{B_2^{n,I} < \infty\}}]\}\big]$$

$$(4.36) \quad \to E\left[\left\{ Bf(z) - \sum_{k=1}^{\infty} b_k z^k + \frac{z}{1 - f''(z)/\mu} \times \sum_{k=1}^{\infty} k b_k (z - z^{k-1}) \right\} 1_{\{B < \infty\}} \right]$$

$$as\ n \to \infty.$$

PROOF. The lemma follows from Corollary 4.5 by showing that

$$(4.37) \quad \begin{aligned} &\{1_{\{C_1^n \le [n^\beta]\}} 1_{\{2 \notin \mathcal{C}_1^n\}} (n-1)\{E[1_{\{\check{B}_2^n < \infty\}}|\mathcal{H}^n = \mathcal{C}_1^n] - E[1_{\{B_2^{n,I} < \infty\}}]\}\} \\ &\qquad = \frac{n-1}{n} \times \{1_{\{C_1^n \le [n^\beta]\}} 1_{\{2 \notin \mathcal{C}_1^n\}} n(\check{y}_n - y_n)\} \end{aligned}$$

is uniformly integrable.

For $k = 1, 2, \ldots$, let $\phi_k^n = \frac{k}{n\mu_n} \sum_{j \notin \mathcal{C}_1^n} 1_{\{\bar{D}_j^n = k\}}$ and let $\varpi_k^n = \frac{1}{n} \sum_{j \notin \mathcal{C}_1^n} 1_{\{\bar{D}_j^n = k\}}$. Let $\hat{\phi}^n = \frac{H_1^n}{n\mu_n}$ and $\hat{\varpi}^n = \frac{C_1^n}{n}$. For $0 \leq s \leq 1$, let $g_n(s) = \sum_{k=1}^\infty \phi_k^n s^{k-1}$ and let $h_n(s) = \sum_{k=0}^\infty \varpi_k^n s^k$. Let $\mathcal{B}^{n,A}$ and $\mathcal{B}^{n,B}$ be branching processes coupled to $\check{\mathcal{B}}_2^n$ and $\mathcal{B}_2^{n,I}$ such that if $\theta_k^n \notin \mathcal{C}_1^n$, the kth individual in all four branching processes has $\bar{D}_{\theta_k^n}^n - 1$ ($k \geq 1$) offspring. (The 0th individual is the initial ancestor and has $\bar{D}_{\theta_0^n}^n$ offspring.) If $\theta_k^n \in \mathcal{C}_1^n$, let the kth individual in $\mathcal{B}^{n,A}$ have no offspring, whilst the kth individual in $\mathcal{B}^{n,B}$ has infinitely many offspring. Let $\hat{y}_n$ and $\bar{y}_n$ denote the extinction probabilities of $\mathcal{B}^{n,A}$ and $\mathcal{B}^{n,B}$, respectively. Then $\hat{y}_n$ and $\bar{y}_n$ are given by

$$\hat{y}_n = h_n(\hat{z}_n) + \hat{\varpi}^n,$$

$$\bar{y}_n = h_n(\bar{z}_n),$$

where $\hat{z}_n$ and $\bar{z}_n$ are the smallest solutions in $[0, 1]$ of

$$\hat{z}_n = g_n(\hat{z}_n) + \hat{\phi}^n,$$

$$\bar{z}_n = g_n(\bar{z}_n).$$

Then $\hat{y}_n \geq \bar{y}_n$, $\hat{z}_n \geq \bar{z}_n$ and $|n(\check{y}_n - y_n)| \leq n(\hat{y}_n - \bar{y}_n)$, since $\hat{y}_n$ and $\bar{y}_n$ form upper and lower bounds for the extinction probabilities of the branching processes $\check{\mathcal{B}}_2^n$ and $\mathcal{B}_2^{n,I}$. Noting that $\frac{d}{ds} g_n(s)$ is increasing in s, and simple algebraic manipulation gives

$$n(\hat{z}_n - \bar{z}_n) \leq \frac{H_1^n}{\mu_n(1 - g_n'(\hat{z}_n))}.$$

Thus,

$$(4.38) \quad |n(\check{y}_n - y_n)| \leq C_1^n + h_n'(1)\frac{H_1^n}{\mu_n(1 - g_n'(\hat{z}_n))} \leq C_1^n + \frac{H_1^n}{1 - g_n'(\hat{z}_n)},$$

since $h_n'(1) \leq \mu_n$.

Fix $0 < \varepsilon < \min\{1, \nu - 1\}$ and $K = 2f'''(1)/L_1 = 2\lambda < \infty$. It is straightforward using Lemma A.3 to show that for all sufficiently large n, $1 - g_n'(\hat{z}_n) > \varepsilon^2/(4K)$. Therefore, since $n^{1+\varepsilon} P(H_1^n > 2C_1^n) \to 0$ as $n \to \infty$, we have that for all sufficiently large n,

$$(4.39) \quad \begin{aligned} &E\big[|n(\check{y}_n - y_n)|^{1+\varepsilon} 1_{\{|\mathcal{C}_1^n| \leq [n^\beta]\}} 1_{\{2 \notin \mathcal{C}_1^n\}}\big] \\ &\quad \leq 1 + E\big[(C_1^n)^{1+\varepsilon} 1_{\{|\mathcal{C}_1^n| \leq [n^\beta]\}}\big]\Big(1 + \frac{8K\mu}{\varepsilon^2}\Big)^{1+\varepsilon}. \end{aligned}$$

By (3.4) and (3.5) at the end of Section 3, $n^{2\beta} P(1_{\{C_1^n \leq [n^\beta]\}} \neq 1_{\{B \leq [n^\beta]\}}) \to 0$ as $n \to \infty$. The lemma follows, since Lemma A.2 proves that $E[B^2 1_{\{B < \infty\}}] < \infty$. □

5. Computing σ^2. From (4.1) and (4.3), we have that

$$\text{var}(\tilde{U}_n/\sqrt{n})$$

$$= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\text{cov}(\chi_i^n, \chi_j^n)$$

(5.1)

$$= \text{var}(\chi_1^n) + (n-1)E\big[1_{\{2\in C_1^n\}}1_{\{C_1^n\leq[n^\beta]\}}\big(1_{\{C_2^n\leq[n^\beta]\}} - E[1_{\{C_2^n\leq[n^\beta]\}}]\big)\big]$$

$$+ (n-1)E\big[1_{\{2\notin C_1^n\}}1_{\{C_1^n\leq[n^\beta]\}}\big(1_{\{C_2^n\leq[n^\beta]\}} - E[1_{\{C_2^n\leq[n^\beta]\}}]\big)\big].$$

Therefore, it follows from Section 4 [specifically (4.1), (4.2) and Lemmas 4.1, 4.2, 4.3 and 4.6] that $\lim_{n\to\infty}\text{var}(\tilde{U}_n/\sqrt{n})$ is equal to

$$\rho(1-\rho) + E\big[(B-1)1_{\{B<\infty\}}\big]\rho$$

(5.2)

$$+ E\left[\left\{Bf(z) - \sum_{k=1}^{\infty}b_k z^k + \frac{z}{1-f''(z)/\mu} \times \sum_{k=1}^{\infty}kb_k(z - z^{k-1})\right\}1_{\{B<\infty\}}\right].$$

We proceed by showing that (5.2) is equal to σ^2, which completes the proof of Theorem 2.1. In order to complete this, we need three preliminary lemmas for the branching process $\mathcal{B}$.

Let $\tilde{\mathcal{B}}$ denote a single-type branching process with one initial ancestor and off-spring distribution $\tilde{D} - 1$, where $P(\tilde{D} = k) = k\pi_k/\mu$ ($k = 1, 2, \ldots$). Let Y and X be nonnegative integer valued random variables with probability mass functions $P(Y = l) = z^l\pi_l/\omega$ and $P(X = l) = z^{l-2}l\pi_l/\mu$ ($l = 0, 1, \ldots$), where ω and z are the extinction probabilities of the branching processes $\mathcal{B}$ and $\tilde{\mathcal{B}}$, respectively. In the branching process $\mathcal{B}$, the offspring of the initial ancestor and subsequent individuals, conditional upon extinction of $\mathcal{B}$, are distributed according to Y and $X - 1$, respectively. [Note that $P(X = 0) = 0$.] We start by deriving the expected value of b_k [the total number of individuals having degree k in the branching process $\mathcal{B}$; see (4.27)] conditional upon extinction of the branching process.

LEMMA 5.1. *For $k = 0, 1, \ldots,$*

(5.3) $\quad E[b_k 1_{\{B<\infty\}}] = P(B < \infty)P(Y = k) + E\big[(B-1)1_{\{B<\infty\}}\big]P(X = k).$

PROOF. Let $Z_0, Z_1, \ldots,$ be independent with $Z_0 \overset{D}{=} Y$ and $Z_i \overset{D}{=} X - 1$ ($i = 1, 2, \ldots$). Let $T = \min\{n : Z_0 + Z_1 + \cdots + Z_n = n\}$. Then T is the total size (not including the initial ancestor) of the branching process $\mathcal{B}$ conditioned on extinction. Fix $k \in \mathbb{Z}_+$. Let $V_0 = 1_{\{Z_0=k\}} - P(Y = k)$ and $V_n = 1_{\{Z_0=k\}} - P(Y = k) + \sum_{i=1}^{n}(1_{\{Z_i=k-1\}} - P(X = k))$ ($n = 1, 2, \ldots$). Then $\{V_0, V_1, \ldots\}$ is a martingale with respect to $\{Z_0, Z_1, \ldots\}$. Also, T is a stopping time with respect to $\{Z_0, Z_1, \ldots\}$, $E[T] < \infty$ as the branching process is subcritical and $E[|V_{n+1} -$

$V_n || Z_0, Z_1, \ldots, Z_n] \leq 1$ for all n. Thus, by Corollary 3.1 on page 260 of Karlin and Taylor (1975),

$$E[V_T] = E[V_0] = 0.$$

Now,

$$V_T = b_k - P(Y = k) - TP(X = k),$$

so

(5.4) $$E[b_k | B < \infty] = P(Y = k) + P(X = k)E[T].$$

Equation (5.3) follows immediately from (5.4), since $E[T] = E[(B - 1)|B < \infty]$ and $E[b_k 1_{\{B<\infty\}}] = E[b_k | B < \infty]P(B < \infty)$. □

LEMMA 5.2.

$$E[(B - 1)1_{\{B<\infty\}}] = \frac{z^2 \mu}{1 - f''(z)/\mu}.$$

PROOF.   Let $\tilde{B}$ denote the total size of the branching process $\tilde{\mathcal{B}}$. Then

$$E[\tilde{B}|\tilde{B} < \infty] = 1 + E\left[\sum_{i=1}^{X-1} \tilde{B}_i \Big| \tilde{B} < \infty\right],$$

where $\tilde{B}_1, \tilde{B}_2, \ldots$ are i.i.d. according to $\tilde{B}$ (and independent of X) and the sum is vacuous if $X = 1$. Thus,

$$E[\tilde{B}|\tilde{B} < \infty] = 1 + E[X - 1]E[\tilde{B}|\tilde{B} < \infty],$$

whence, since $E[X - 1] = f''(z)/\mu < 1$ (see remarks after Theorem 2.1)

$$E[\tilde{B}|\tilde{B} < \infty] = \frac{1}{1 - E[X - 1]} = \frac{1}{1 - f''(z)/\mu}.$$

Now

$$E[(B - 1)1_{\{B<\infty\}}] = E[(B - 1)|B < \infty]P(B < \infty)$$

(5.5) $$= E\left[\sum_{i=1}^{Y} \tilde{B}_i \Big| B < \infty\right]P(B < \infty)$$

$$= E[Y]E[\tilde{B}|\tilde{B} < \infty]P(B < \infty),$$

where

$$E[Y] = \sum_{l=0}^{\infty} l \frac{z^l \pi_l}{\omega} = \frac{z f'(z)}{\omega} = \frac{z^2 \mu}{\omega},$$

by (1.3). The lemma follows since $P(B < \infty) = \omega$. □

LEMMA 5.3. *For any $0 \le s \le 1$,*

$$\sum_{k=0}^{\infty} s^k E[b_k 1_{\{B<\infty\}}] = f(sz) + \frac{szf'(sz)}{1 - f''(z)/\mu}$$

and

$$(5.6) \quad \sum_{k=1}^{\infty} ks^{k-1} E[b_k 1_{\{B<\infty\}}] = zf'(sz) + \frac{zf'(sz)}{1 - f''(z)/\mu} + \frac{sz^2 f''(sz)}{1 - f''(z)/\mu}.$$

PROOF. Using (5.3) and Lemma 5.2,

$$\sum_{k=0}^{\infty} s^k E[b_k 1_{\{B<\infty\}}] = \omega \sum_{k=0}^{\infty} s^k \frac{z^k \pi_k}{\omega} + \frac{z^2 \mu}{1 - f''(z)/\mu} \sum_{k=1}^{\infty} s^k \frac{z^{k-2} k \pi_k}{\mu}$$

$$(5.7) \qquad\qquad = f(sz) + \frac{sz}{1 - f''(z)/\mu} \sum_{k=1}^{\infty} k(sz)^{k-1} \pi_k$$

$$\qquad\qquad = f(sz) + \frac{sz}{1 - f''(z)/\mu} f'(sz).$$

The proof of the lemma is completed by noting that (5.6) follows by differentiating (5.7) with respect to s. □

We complete the proof of Theorem 2.1 by noting that $\omega = f(z)$ and, by (5.2) and Lemmas 5.2 and 5.3, we have that

$$\lim_{n\to\infty} \mathrm{var}\left(\frac{\tilde{U}_n}{\sqrt{n}}\right)$$

$$= \rho(1-\rho) + E[(B-1)1_{\{B<\infty\}}]\rho$$

$$+ E\left[\left\{Bf(z) - \sum_{k=1}^{\infty} b_k z^k + \frac{z}{1 - f''(z)/\mu} \times \sum_{k=1}^{\infty} kb_k(z - z^{k-1})\right\} 1_{\{B<\infty\}}\right]$$

$$(5.8) \; = \rho(1-\rho) + \frac{\rho z^2 \mu}{1 - f''(z)/\mu}$$

$$+ \left\{\omega + \frac{z^2 \mu}{1 - f''(z)/\mu}\right\}\omega - f(z^2) - \frac{z^2 f'(z^2)}{1 - f''(z)/\mu}$$

$$+ \frac{z}{1 - f''(z)/\mu}\left\{z\left(zf'(z) + \frac{zf'(z)}{1 - f''(z)/\mu} + \frac{z^2 f''(z)}{1 - f''(z)/\mu}\right)\right.$$

$$\left. - \left(zf'(z^2) + \frac{zf'(z^2)}{1 - f''(z)/\mu} + \frac{z^3 f''(z^2)}{1 - f''(z)/\mu}\right)\right\}.$$

Straightforward algebraic manipulation, using $\omega + \rho = 1$, $\omega = f(z)$ and $f'(z) = z\mu$, gives

$$
\begin{aligned}
\lim_{n\to\infty} \mathrm{var}\left(\frac{\tilde{U}_n}{\sqrt{n}}\right) &= \rho(1-\rho) + (1-\rho)^2 + \frac{z^2\mu}{1-f''(z)/\mu} - f(z^2) \\
&\quad + \frac{z^2}{1-f''(z)/\mu}(-f'(z^2) + zf'(z) - f'(z^2)) \\
&\quad + \frac{z^2}{(1-f''(z)/\mu)^2}\{z^2\mu + z^2 f''(z) - f'(z^2) - z^2 f''(z^2)\} \\
&= \sigma^2
\end{aligned}
$$

(5.9)

as required.

6. The weakly supercritical case. The weakly supercritical case, studied in Riordan (2012), corresponds to $\nu_n \downarrow 1$ as $n \to \infty$, sufficiently slowly such that $n^{1/3}(\nu_n - 1) \to \infty$ as $n \to \infty$. In Riordan (2012), Theorem 1.1, it is shown that $\sqrt{n}(R_n/n - \rho_n)$ is asymptotically normal with mean 0 and $\mathrm{var}(\sqrt{n}(R_n/n - \rho_n)) \sim 2\mu_n/(\nu_n - 1)$, where ρ_n and z_n denote the solutions to (1.2) and (1.3), respectively, with $\{\pi_i\}$ replaced by $\{\pi_i^n\}$, μ replaced by μ_n, and $f(s)$ and its derivatives replaced by $f_n(s) = \sum_{i=0}^{\infty} \pi_i^n s^i$ and its derivatives [cf. (4.26)]. Similarly, let σ_n^2 be given by (2.5), with z, ρ, μ and f replaced by z_n, ρ_n, μ_n and f_n, respectively. We proceed by showing that

(6.1)
$$
\sigma_n^2 \sim \frac{2\mu_n}{\nu_n - 1},
$$

whence from Riordan (2012), Theorem 1.1, it follows that $\mathrm{var}(\sqrt{n}(R_n/n - \rho_n)) \sim \sigma_n^2$. We assume that conditions (a)–(c) in Section 2 hold, and also that $\sum_{j=0}^{\infty} j^4 \pi_j < \infty$ and $\sum_{j=0}^{\infty} j^4 \pi_j^n \to \sum_{j=0}^{\infty} j^4 \pi_j$ as $n \to \infty$.

Let $\delta_n = 1 - z_n$, then $\delta_n \to 0$ ($z_n \to 1$) as $n \to \infty$. Using (1.3), we have that $z_n = \frac{1}{\mu_n} f_n'(z_n)$, whence, using Maclaurin's theorem,

$$
\begin{aligned}
1 - \delta_n &= \frac{1}{\mu_n} \sum_{i=1}^{\infty} i \pi_i^n (1 - \delta_n)^{i-1} \\
&= \frac{1}{\mu_n} \sum_{i=1}^{\infty} i \pi_i^n \left(1 - (i-1)\delta_n + \frac{(i-1)(i-2)}{2}\delta_n^2\right) + o(\delta_n^2) \\
&= 1 - \nu_n \delta_n + \frac{\lambda_n}{2}\delta_n^2 + o(\delta_n^2).
\end{aligned}
$$

(6.2)

Recall that z_n is the smallest solution in $[0, 1]$ of $s = \frac{1}{\mu_n} f'_n(s)$. Thus, (6.2) implies that $\delta_n = 2(\nu_n - 1)/\lambda_n + o(\delta_n)$. Therefore, it follows from (1.2) that

(6.3)
$$\rho_n = 1 - \sum_{i=0}^{\infty}(1 - \delta_n)^i \pi_i^n = 1 - \sum_{i=0}^{\infty}(1 - i\delta_n)\pi_i^n + o(\delta_n)$$

$$= \mu_n \delta_n + o(\delta_n) = \frac{2\mu_n}{\lambda_n}(\nu_n - 1) + o(\delta_n).$$

Thus, $\rho_n \sim 2(\nu_n - 1)\mu_n/\lambda_n$, in agreement with Riordan (2012) equation (1.11). It is straightforward using a binomial expansion to show that, for $k = 1, 2$,

(6.4)
$$f_n(z_n^k) = \sum_{i=0}^{\infty}\pi_i^n(1 - \delta_n)^{ki}$$

$$= \sum_{i=0}^{\infty}\pi_i^n(1 - ki\delta_n) + o(\delta_n) = 1 - k\delta_n\mu_n + o(\delta_n),$$

and, similarly, that

(6.5)
$$f'_n(z_n^k) = \mu_n(1 - k\delta_n\nu_n) + o(\delta_n),$$

(6.6)
$$f''_n(z_n^k) = \mu_n(\nu_n - k\delta_n\lambda_n) + o(\delta_n).$$

Since $\delta_n\lambda_n \sim 2(\nu_n - 1)$, it follows from (6.6) that

(6.7)
$$1 - f''_n(z_n)/\mu_n = 1 - \{\nu_n - 2(\nu_n - 1) + o(\delta_n)\}$$

$$= \nu_n - 1 + o(\delta_n).$$

Therefore, turning to σ_n^2 [cf. (2.5)], it follows from (6.3) and (6.4) that

(6.8)
$$1 - \rho_n - f_n(z_n^2) = \mu_n\delta_n + o(\delta_n);$$

using (6.5) and (6.7) that

(6.9)
$$\frac{z_n^2}{1 - f''_n(z_n)/\mu_n}\{(1 + z_n^2)\mu_n - 2f'_n(z_n^2)\} = \frac{2\delta_n\mu_n(2\nu_n - 1)}{\nu_n - 1} + o(1);$$

and, using (6.5)–(6.7), that

(6.10)
$$\frac{z_n^2}{(1 - f''_n(z_n)/\mu_n)^2}\{z_n^2\mu_n + z_n^2 f''_n(z_n) - f'_n(z_n^2) - z_n^2 f''_n(z_n^2)\}$$

$$= \frac{2\delta_n\mu_n(\nu_n - 1) + \delta_n\mu_n\lambda_n + o(\delta_n)}{(\nu_n - 1)^2}$$

$$= \frac{2\mu_n(1 + \delta_n) + o(\delta_n)}{\nu_n - 1}.$$

Substituting (6.8)–(6.10) into the equation for σ_n^2 [cf. (2.5)] yields

$$(\nu_n - 1)\sigma_n^2 = \delta_n \mu_n(\nu_n - 1) + 2\delta_n \mu_n(2\nu_n - 1) + 2\mu_n(1 + \delta_n) + \frac{o(\delta_n)}{\nu_n - 1}$$

$$= 2\mu_n + \frac{o(\delta_n)}{\delta_n},$$

since $\nu_n - 1 \sim \delta_n$, and (6.1) follows.

7. Concluding remarks. The main interest in the variance of the giant component of a configuration model random graph is for a given degree sequence $\bar{\mathbf{D}}^n$, where n is finite. Therefore, we briefly illustrate that the asymptotic expressions are applicable for moderate n by considering $n = 200$ and $n = 1000$. For four different distributions, D, given below, we generate $\bar{\mathbf{D}}^n$ mimicking D. That is, we ensure that $\bar{\mathbf{D}}^n$ satisfies $[n P(D = k)] \leq \sum_{i=1}^{n} 1_{\{\bar{D}_i^n = k\}} < [n P(D = k)] + 1$ for $k = 0, 1, \ldots$. Then using $\bar{\mathbf{D}}^n$, we simulated 10,000 random graphs and recorded the size of the giant component in each graph. For each $\bar{\mathbf{D}}^n$, the mean proportion (ρ_n) and scaled variance (σ_n^2) of the giant component from the simulated graphs are compared with ρ and σ^2, the formulae for the mean and variance, respectively, given by the asymptotic results. The results are presented in Table 1 for the four degree distributions:

(1) $P(D = 1) = P(D = 3) = 1/2$;
(2) $D \sim \text{Geom}(q)$ with $q = 0.5$ and support $\mathbb{N}$, i.e. $P(D = k) = 0.5^k$ ($k = 1, 2, \ldots$);
(3) $D \sim \text{Po}(\mu)$ with $\mu = 2$, i.e. $P(D = k) = e^{-2}2^k/k!$ ($k = 0, 1, \ldots$);
(4) $P(D = k) \propto k^{-m}$ ($k = 2, 3, \ldots, n$) and $P(D = 1) = 2P(D = 3)$ with $m = 4$.

Table 1 shows that there is good agreement for $n = 1000$ between the theoretical calculations of the mean and variance of the giant component and the results obtained via simulation. Although larger discrepancies are observed for $n = 200$, the theoretical results are still useful in this case.

TABLE 1
Simulation results against theoretical (*asymptotic*) *calculations*

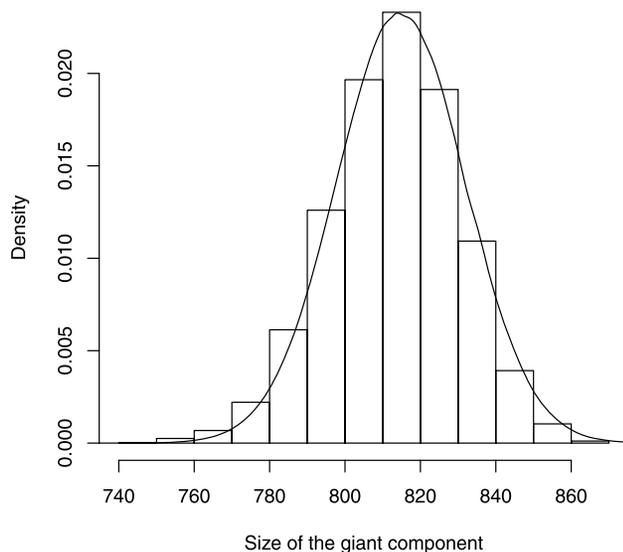|	Theoretical		$n = 200$		$n = 1000$	
D	ρ	σ^2	$\lvert\rho_n - \rho\rvert$	$\lvert\sigma_n^2 - \sigma^2\rvert$	$\lvert\rho_n - \rho\rvert$	$\lvert\sigma_n^2 - \sigma^2\rvert$
1	0.8148	0.2936	0.0053	0.0456	0.0008	0.0051
2	0.7639	0.3416	0.0017	0.0152	0.0002	0.0050
3	0.7968	0.1365	0.0017	0.0025	0.0002	0.0016
4	0.8906	0.3530	0.0057	0.0346	0.0035	0.0167

FIG. 1.   *Histogram of R_n with $n = 1000$, based upon a sample of size* 10,000, *together with density plot of* $N(n\rho, n\sigma^2)$, *both with degree distribution* 1.

A natural extension of Theorem 2.1 is to seek a central limit theorem for the size of the giant component R_n [cf. (1.1)], that is, $\sqrt{n}(R_n/n - \rho) \xrightarrow{D} N(0, \sigma^2)$ as $n \to \infty$. This is supported by the simulation study as illustrated by Figure 1 where, for degree distribution 1, a histogram of the simulated R_n is plotted, together with the density of the $N(n\mu, n\sigma^2)$ distribution. Similar plots were observed for the other three degree distributions mentioned above. Further support for the existence of a central limit theorem is given in Figure 2 as follows. For each of the four degree distributions and for $n = 200, 300, \ldots, 1000$, 10,000 random graphs were simulated. For $n = 200, 300, \ldots, 1000$ and $i = 1, 2, \ldots, 10,000$, the size R_n^i of the giant component and the normalised squared difference $N_n^i = (R_n^i - n\rho)^2/(n\sigma^2)$ were recorded. The latter were grouped into batches of 10 with, for $j = 1, 2, \ldots, 1000$, $S_n^j = \sum_{i=10j-9}^{10j} N_n^i$. If a central limit theorem exists then S_n^1 will converge in distribution to a χ_{10}^2 distribution as $n \to \infty$. In Figure 2, the empirical 5%, 50% and 95% quantiles for $(S_n^1, S_n^2, \ldots, S_n^{1000})$ ($n = 200, 300, \ldots, 1000$) are plotted for each degree distribution, with the three horizontal lines denoting the corresponding quantiles of the χ_{10}^2 distribution. Figure 2 shows convergence of the S_n quantiles towards the χ_{10}^2 quantiles as n increases.

A key extension of the results of this paper is the asymptotic variance of the giant component of a Newman–Strogatz–Watts (NSW) random graph model, where the deterministic sequence $\bar{\mathbf{D}}^n$ is replaced by taking $\bar{D}_1, \bar{D}_2, \ldots, \bar{D}_n$ to be i.i.d. according to a nonnegative integer-valued random variable, D. As noted in Section 1, the degree sequence might be infeasible with $\sum_{i=1}^n \bar{D}_i^n$ being odd, in which case,
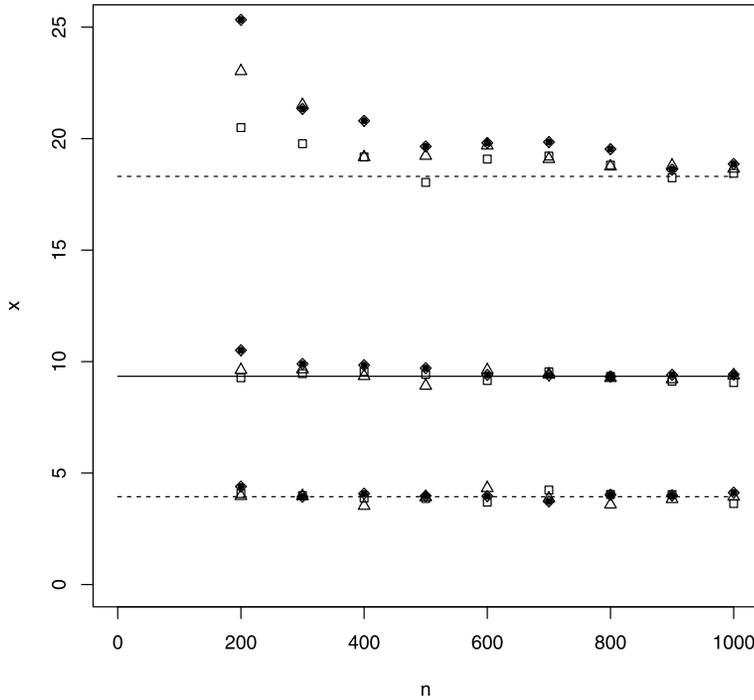
FIG. 2. *Empirical 5%, 50% and 95% quantiles for each degree distribution (1—dot; 2—squares; 3—diamonds; 4—triangles) and corresponding χ_{10}^2 quantiles (lines).*

we simply ignore the last half-edge in the construction of the random graph. For the NSW random graph, we can again show that $\mathrm{var}(\sqrt{n}(R_n/n - \rho))$ and $\mathrm{var}(\tilde{U}_n/\sqrt{n})$ have the same asymptotic limit (should one exist) as $n \to \infty$. However, in computing $\mathrm{var}(\tilde{U}_n/\sqrt{n}) = \mathrm{var}(\chi_1^n) + (n-1)\mathrm{cov}(\chi_1^n, \chi_2^n)$, we need to take into account variability in $\bar{\mathbf{D}}^n$. As above, it is straightforward to show that $\mathrm{var}(\chi_1^n) \to \rho(1-\rho)$ as $n \to \infty$. However, for $(n-1)\mathrm{cov}(\chi_1^n, \chi_2^n)$, we use the law of total covariance, writing

$$
(n-1)\mathrm{cov}(\chi_1^n, \chi_2^n)
$$

$$
(7.1) \qquad = (n-1)E[\mathrm{cov}(\chi_1^n, \chi_2^n | \bar{\mathbf{D}}^n)] + (n-1)\mathrm{cov}(E[\chi_1^n | \bar{\mathbf{D}}^n], E[\chi_2^n | \bar{\mathbf{D}}^n])
$$

$$
= E[(n-1)\mathrm{cov}(\chi_1^n, \chi_2^n | \bar{\mathbf{D}}^n)] + \mathrm{var}(\sqrt{n-1}E[\chi_1^n | \bar{\mathbf{D}}^n]),
$$

since, by exchangeability, $E[\chi_2^n | \bar{\mathbf{D}}^n] = E[\chi_1^n | \bar{\mathbf{D}}^n]$. That is, the covariance comprises two parts, variation in the construction, and hence in the size of the giant component of $\mathcal{G}^n$ given $\bar{\mathbf{D}}^n$ (this is the variation which is observed for the MR random graph) and variance in the (mean) size of the giant component of $\mathcal{G}^n$ due to variability in $\bar{\mathbf{D}}^n$. For $\varepsilon, \delta > 0$ and $n = 1, 2, \ldots$, let $K_{\varepsilon,\delta}^n$ denote the event that (i) $n^{-(1/4-\delta)}\Delta_n < \varepsilon$, (ii) (2.1) holds for all $j \geq 0$ and (iii) (2.3) holds for

$k = 1, 2, 3$, with the components of $\bar{\mathbf{D}}^n$ drawn independently according to D. Note that the conditions $\nu > 1$ and $\pi_1 > 0$ are conditions upon D, which are satisfied if $E[D(D-1)] > E[D]$ and $P(D = 1) > 0$. It is relatively straightforward, but lengthy, to show that provided that there exists $\gamma > 0$, such that $E[D^{8+\gamma}] < \infty$, then there exists $\delta > 0$, such that for all $\varepsilon > 0$, $nP((K_{\varepsilon,\delta}^n)^C) \to 0$ as $n \to \infty$, from which it is straightforward to show that the first term on the right-hand side of (7.1) converges to $\sigma_1^2 = \sigma^2 - \rho(1-\rho)$ as $n \to \infty$. It can also be shown that

$$\text{var}\big(\sqrt{n-1}\,E\big[\chi_1^n|\bar{\mathbf{D}}^n\big]\big)$$

(7.2)
$$\to f(z^2) - f(z)^2 + \left(\frac{z}{1 - f''(z)/\mu}\right)^2$$
$$\times \big\{z^2(f''(z^2) + f''(1) - 3\mu) + f'(z^2)(3 - 2f''(z)/\mu)\big\}$$

as $n \to \infty$. This gives for the NSW random graph that $\text{var}(\sqrt{n}(R_n/n - \rho)) \to \sigma_{\text{NSW}}^2$, where

(7.3)
$$\sigma_{\text{NSW}}^2 = \rho(1-\rho) + \frac{z^2}{1 - f''(z)/\mu}\mu(1 + z^2)$$
$$+ \frac{z^4}{(1 - f''(z)/\mu)^2}(\mu(\nu - 2) + f''(z)).$$

The details of the proof of (7.3) will be presented elsewhere along with other extensions such as the variance of the size of a Reed–Frost epidemic which becomes established on a (MR or NSW) random graph, Britton, Janson and Martin-Löf (2007), Section 2.3, and the size of the giant component of a percolation model on a random graph, Janson (2009a).

APPENDIX: RESULTS FOR GALTON–WATSON BRANCHING PROCESSES

In this section, we present various useful results for Galton–Watson branching processes utilised in the proof of Theorem 2.1.

Let $\mathcal{B}$ be a Galton–Watson branching process where the initial ancestor has offspring distribution Y and all subsequent individuals have offspring distribution X. For $0 \le s \le 1$, let $f(s) = E[s^Y]$ and $g(s) = E[s^X]$. Let B denote the total size, including the initial ancestor, of the branching process $\mathcal{B}$. Then $B = 1 + \sum_{i=1}^{Y} \tilde{B}_i$ where $\tilde{B}_1, \tilde{B}_2, \ldots$ are i.i.d. copies of $\tilde{B}$, the total size, including the initial ancestor, of the branching process $\tilde{B}$ in which all individuals have offspring distribution X. Further, $\tilde{B}_1, \tilde{B}_2, \ldots$ are independent of Y. Let p be the probability that the branching process $\mathcal{B}$ goes extinct. Then $p = f(z)$ where z is the smallest solution in $[0, 1]$ of $s = g(s)$. We assume throughout that $P(X = 0) > 0$ implying that $z > 0$.

We present two lemmas concerning the moments of the total size of branching processes. Lemma A.1 is for subcritical branching processes. Lemma A.2 is for supercritical branching processes conditioned upon extinction and follows straightforwardly from Lemma A.1.

LEMMA A.1. *Suppose that there exists $k \in \mathbb{N}$, such that $E[X^k] < \infty$ and $E[X] < 1$. Then $E[\tilde{B}^k] < \infty$.*
If in addition $E[Y^k] < \infty$, then $E[B^k] < \infty$.

PROOF.   Note that, for $E[X] < 1$,

$$E[\tilde{B}] = 1 + E[X]E[\tilde{B}]$$

$$= \frac{1}{1 - E[X]}.$$

We now prove that $E[\tilde{B}^k] < \infty$ by induction on k.
Let $\tilde{B}_1, \tilde{B}_2, \ldots$ be i.i.d. according to $\tilde{B}$. Then

$$E[\tilde{B}^k] = E\left[\left(1 + \sum_{i=1}^{X} \tilde{B}_i\right)^k\right]$$

(A.1)
$$= E\left[\sum_{j=0}^{k} \binom{k}{j}\left(\sum_{i=1}^{X} \tilde{B}_i\right)^j\right]$$

$$= \sum_{j=0}^{k} \binom{k}{j} E\left[E\left[\left(\sum_{i=1}^{X} \tilde{B}_i\right)^j \Big| X\right]\right].$$

Since for all $j \in \mathbb{N}$, $(\sum_{i=1}^{X} \tilde{B}_i)^j \le X^{j-1} \sum_{i=1}^{X} \tilde{B}_i^j$ [as $f(x) = x^j$ is convex], it follows that

(A.2)
$$E\left[\left(\sum_{i=1}^{X} \tilde{B}_i\right)^j\right] = E\left[E\left[\left(\sum_{i=1}^{X} \tilde{B}_i\right)^j \Big| X\right]\right]$$

$$\le E\left[E\left[X^{j-1} \sum_{i=1}^{X} \tilde{B}_i^j \Big| X\right]\right]$$

$$= E\left[X^{j-1} \sum_{i=1}^{X} E[\tilde{B}_i^j]\right]$$

(A.3)
$$= E[X^j]E[\tilde{B}^j].$$

Therefore,

(A.4)
$$E[\tilde{B}^k] \le \sum_{j=0}^{k-1} \binom{k}{j} E[X^j]E[\tilde{B}^j] + E\left[\left(\sum_{i=1}^{X} \tilde{B}_i\right)^k\right].$$

Now

$$E\left[\left(\sum_{i=1}^{X} \tilde{B}_i\right)^k\right] = E\left[E\left[\left(\sum_{i=1}^{X} \tilde{B}_i\right)^k \Big| X\right]\right],$$

where the inner expectation consists of X^k terms of which X are $E[\tilde{B}^k]$. The remaining $X^k - X$ terms are of the form $\prod_{j=1}^{M} E[\tilde{B}_{i_j}^{l_j}]$, where $M \geq 2$ and $l_1, l_2, \ldots, l_M \geq 1$ with $l_1 + l_2 + \cdots + l_M = k$. By Jensen's inequality, for $1 \leq l < k$, $E[\tilde{B}^l] = E[(\tilde{B}^{k-1})^{l/(k-1)}] \leq E[\tilde{B}^{k-1}]^{l/(k-1)}$, so

$$E\left[\left(\sum_{i=1}^{X} \tilde{B}_i\right)^k\right] \leq E[X]E[\tilde{B}^k] + E[X^k - X]E[\tilde{B}^{k-1}]^{\frac{k}{k-1}}.$$

Hence, since $E[X] < 1$,

$$E[\tilde{B}^k] \leq \frac{1}{1 - E[X]}\left\{\sum_{j=0}^{k-1}\binom{k}{j}E[X^j]E[\tilde{B}^j] + E[X^k - X]E[\tilde{B}^{k-1}]^{\frac{k}{k-1}}\right\},$$

which is finite by induction.

A similar argument based upon $B \overset{D}{=} 1 + \sum_{i=1}^{Y} \tilde{B}_i$ completes the proof of the lemma. $\square$

LEMMA A.2.   *Suppose that $E[X] > 1$. Then for any $k \in \mathbb{N}$,*

$$E[B^k 1_{\{B<\infty\}}] < \infty.$$

Consequently, for any $\alpha, \beta > 0$,

(A.5) $n^\alpha(P(B < \infty) - P(B \leq [n^\beta])) \to 0$ *as $n \to \infty$.*

PROOF.   Fix $k \in \mathbb{N}$. Let $\hat{Y}$ and $\hat{X}$ be integer-valued random variables with $P(\hat{Y} = l) = z^l P(Y = l)/p$ and $P(\hat{X} = l) = z^{l-1} P(X = l)$ $(l = 0, 1, \ldots)$. Then $E[\hat{Y}^k] < \infty$ and $E[\hat{X}^k] < \infty$. Let $\hat{\mathcal{B}}$ be a branching process where the initial ancestor has $\hat{Y}$ offspring and all subsequent individuals offspring distribution $\hat{X}$. Then $E[\hat{X}] = g'(z) < 1$ implying that $\hat{\mathcal{B}}$ is a subcritical branching process, and furthermore, letting $\hat{B}$ denote the total size of $\hat{\mathcal{B}}$, $B|\{B < \infty\} \overset{D}{=} \hat{B}$; see, for example, Waugh (1958), Section 6. Thus,

(A.6) $E[B^k 1_{\{B<\infty\}}] = E[B^k | B < \infty]P(B < \infty) = E[\hat{B}^k]P(B < \infty),$

and the right-hand side of (A.6) is finite by Lemma A.1.

Fix $\alpha, \beta > 0$ and take $l \in \mathbb{N}$ such that $l > \alpha/\beta$. Then note that

$$P(B < \infty) = P(B \leq [n^\beta]) + P([n^\beta] < B < \infty)$$
$$= P(B \leq [n^\beta]) + P(B > [n^\beta]|B < \infty)P(B < \infty),$$

since $\{B \leq [n^\beta]\} \subset \{B < \infty\}$. However, by (A.6) and Markov's inequality, $P(B > [n^\beta]|B < \infty)P(B < \infty) \leq ([n^\beta])^{-l}P(B < \infty)E[\hat{B}^l]$ and (A.5) follows. $\square$

It is well known that for supercritical branching processes $z < 1$, and moreover, $g'(z) < 1$. However, $g'(z)$ can be arbitrarily close to 1. Lemma A.3 provides a useful upper bound for $g'(z)$ in terms of the first and second moment of X.

LEMMA A.3. *Suppose that there exists $\varepsilon > 0$ such that $g'(1)(= E[X]) \geq 1 + \varepsilon$. Then if $g''(1)(= E[X(X-1)]) < \infty$, we have that $z \leq 1 - \frac{\varepsilon}{g''(1)}$ and*

$$g'(z) \leq 1 - \frac{\varepsilon^2}{4g''(1)}.$$

PROOF. Since $g''(s)$ is increasing in s, we have by the mean value theorem that, for all $w > 1 - \varepsilon/g''(1)$, $g'(w) > 1$. Hence, for all $w > 1 - \varepsilon/g''(1)$, $g(w) < w$, implying that $z \leq 1 - \varepsilon/g''(1)$.

Let $y = 1 - \frac{\varepsilon}{2g''(1)}$. Then similarly, $g'(y) \geq g'(1) - (1 - y)g''(1)$ giving $g'(y) \geq 1 + \varepsilon/2$. Since $g'(s)$ is increasing in s, by the mean value theorem,

$$g'(z)(y - z) \leq g(y) - g(z) = g(y) - z,$$

and

$$1 - g(y) = g(1) - g(y) \geq g'(y)(1 - y)$$

(A.7)

$$\geq \left(1 + \frac{\varepsilon}{2}\right)(1 - y).$$

Thus, $y - g(y) \geq (1 - y)\varepsilon/2$, giving (since $y - z \leq 1$)

$$g'(z) \leq \frac{g(y) - z}{y - z} = 1 - \frac{y - g(y)}{y - z}$$

$$\leq 1 - \frac{\varepsilon}{2}(1 - y) = 1 - \frac{\varepsilon^2}{4g''(1)},$$

as required. □

Corollary A.4 follows from Lemma A.3 and is required for the proof of Lemma 4.3.

COROLLARY A.4. *For any $k \in \mathbb{N}$ and any $0 < \varepsilon < L < \infty$, there exists a constant $A_{k,L,\varepsilon} < \infty$ such that for any supercritical branching process $\mathcal{B}$ with $E[X] \geq 1 + \varepsilon$ and $E[X(X-1)] \leq L$,*

$$E\left[B^k 1_{\{B < \infty\}}\right] \leq A_{k,L,\varepsilon}.$$

PROOF. For $k \in \mathbb{N}$ and $0 < \varepsilon < L < \infty$, let $C_{k,L,\varepsilon} = \sum_{i=0}^{\infty} i^k (1 - \varepsilon/L)^i$ and $D_{k,L,\varepsilon} = \sum_{i=0}^{\infty} i^k (1 - \varepsilon/L)^{i-1}$. Then $C_{k,L,\varepsilon}$ and $D_{k,L,\varepsilon}$ are both finite. Therefore, by Lemma A.3,

(A.8)

$$E[\hat{X}] = \sum_{i=0}^{\infty} i P(X = i) z^{i-1} = g'(z) \leq 1 - \frac{\varepsilon^2}{4L},$$

$$(A.9) \qquad E\big[\hat{X}^k\big] \le \sum_{i=0}^{\infty} i^k z^{i-1} \le \sum_{i=0}^{\infty} i^k (1-\varepsilon/L)^{i-1} = D_{k,L,\varepsilon},$$

$$(A.10) \qquad E\big[\hat{Y}^k\big] \le \frac{1}{p}\sum_{i=0}^{\infty} i^k z^i \le \frac{1}{p}\sum_{i=0}^{\infty} i^k (1-\varepsilon/L)^i = C_{k,L,\varepsilon}/p,$$

where $\hat{X}$ and $\hat{Y}$ are defined as in Lemma A.2.

Recall the branching process $\tilde{\mathcal{B}}$, defined at the start of the Appendix. Let $\check{B}$ denote the total size of $\tilde{\mathcal{B}}$, conditional upon $\tilde{\mathcal{B}}$ going extinct. Then it is straightforward, using the proof of Lemma A.1, (A.8) and (A.9), to show that there exists a constant $\tilde{A}_{k,L,\varepsilon}$, such that

$$(A.11) \qquad E\big[\check{B}^k\big] \le \tilde{A}_{k,L,\varepsilon}.$$

Following (A.4), we have that

$$(A.12) \qquad E\big[\hat{B}^k\big] \le \sum_{l=0}^{k} \binom{k}{l} E\big[\hat{Y}^l\big] E\big[\check{B}^l\big].$$

From (A.10), (A.11) and (A.12), there exists a constant $A_{k,L,\varepsilon}$, such that $E[\hat{B}^k] \le A_{k,L,\varepsilon}/p$. Finally, note that

$$E\big[B^k 1_{\{B<\infty\}}\big] = P(B < \infty)E\big[\hat{B}^k\big] \le p\frac{A_{k,L,\varepsilon}}{p} = A_{k,L,\varepsilon},$$

as required. □

REFERENCES

BALL, F. and DONNELLY, P. (1995). Strong approximations for epidemic models. *Stochastic Process. Appl.* **55** 1–21. MR1312145

BARBOUR, A. D., HOLST, L. and JANSON, S. (1992). *Poisson Approximation. Oxford Studies in Probability* **2**. Oxford Univ. Press, New York. MR1163825

BARBOUR, A. D. and MOLLISON, D. (1990). Epidemics and random graphs. In *Stochastic Processes in Epidemic Theory. Lecture Notes in Biomathematics* **86** 86–89. Springer, Berlin.

BOLLOBÁS, B. (1980). A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.* **1** 311–316. MR0595929

BOLLOBÁS, B. (2001). *Random Graphs*, 2nd ed. *Cambridge Studies in Advanced Mathematics* **73**. Cambridge Univ. Press, Cambridge. MR1864966

BRITTON, T., DEIJFEN, M. and MARTIN-LÖF, A. (2006). Generating simple random graphs with prescribed degree distribution. *J. Stat. Phys.* **124** 1377–1397. MR2266448

BRITTON, T., JANSON, S. and MARTIN-LÖF, A. (2007). Graphs with specified degree distributions, simple epidemics, and local vaccination strategies. *Adv. in Appl. Probab.* **39** 922–948. MR2381582

CHUNG, F. and LU, L. (2002). Connected components in random graphs with given expected degree sequences. *Ann. Comb.* **6** 125–145. MR1955514

DURRETT, R. (2007). *Random Graph Dynamics. Cambridge Series in Statistical and Probabilistic Mathematics.* Cambridge Univ. Press, Cambridge. MR2271734

ERDŐS, P. and RÉNYI, A. (1959). On random graphs. I. *Publ. Math. Debrecen* **6** 290–297. MR0120167

JANSON, S. (2009a). On percolation in random graphs with given vertex degrees. *Electron. J. Probab.* **14** 87–118. MR2471661

JANSON, S. (2009b). The probability that a random multigraph is simple. *Combin. Probab. Comput.* **18** 205–225. MR2497380

KARLIN, S. and TAYLOR, H. M. (1975). *A First Course in Stochastic Processes*, 2nd ed. Academic Press, New York. MR0356197

LEFÈVRE, C. and UTEV, S. (1999). Branching approximation for the collective epidemic model. *Methodol. Comput. Appl. Probab.* **1** 211–228. MR1728893

MOLLOY, M. and REED, B. (1995). A critical point for random graphs with a given degree sequence. *Random Structures Algorithms* **6** 161–179. MR1370952

MOLLOY, M. and REED, B. (1998). The size of the giant component of a random graph with a given degree sequence. *Combin. Probab. Comput.* **7** 295–305. MR1664335

NEAL, P. (2007). Coupling of two SIR epidemic models with variable susceptibilities and infectivities. *J. Appl. Probab.* **44** 41–57. MR2312985

NEWMAN, M., STROGATZ, S. and WATTS, D. (2001). Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E* **64** 026118.

RIORDAN, O. (2012). The phase transition in the configuration model. *Combin. Probab. Comput.* **21** 265–299. MR2900063

WAUGH, W. A. O'N. (1958). Conditioned Markov processes. *Biometrika* **45** 241–249. MR0092271

WHITTLE, P. (1955). The outcome of a stochastic epidemic—a note on Bailey's paper. *Biometrika* **42** 116–122. MR0070099

SCHOOL OF MATHEMATICAL SCIENCES
UNIVERSITY OF NOTTINGHAM
THE MATHEMATICAL SCIENCES BUILDING
UNIVERSITY PARK
NOTTINGHAM, NG7 2RD
UNITED KINGDOM
E-MAIL: Frank.Ball@nottingham.ac.uk

DEPARTMENT OF MATHEMATICS
    AND STATISTICS
FYLDE COLLEGE
LANCASTER UNIVERSITY
LANCASTER, LA1 4YF
UNITED KINGDOM
E-MAIL: p.neal@lancaster.ac.uk