Combinatorics, Probability and Computing (2012) 21, 265–299.
doi:10.1017/S0963548311000666

The Phase Transition in the Configuration Model

OLIVER RIORDAN

Mathematical Institute, University of Oxford, 24–29 St Giles', Oxford OX1 3LB, UK
and
Department of Mathematical Sciences, University of Memphis, TN 38152, USA
(e–mail: riordan@maths.ox.ac.uk)

Received 5 April 2011; revised 4 November 2011; first published online 2 February 2012

Let $G = G(\mathbf{d})$ be a random graph with a given degree sequence $\mathbf{d}$, such as a random r-regular graph where $r \geqslant 3$ is fixed and $n = |G| \to \infty$. We study the percolation phase transition on such graphs G, *i.e.*, the emergence as p increases of a unique giant component in the random subgraph $G[p]$ obtained by keeping edges independently with probability p. More generally, we study the emergence of a giant component in $G(\mathbf{d})$ itself as $\mathbf{d}$ varies. We show that a single method can be used to prove very precise results below, inside and above the 'scaling window' of the phase transition, matching many of the known results for the much simpler model $G(n, p)$. This method is a natural extension of that used by Bollobás and the author to study $G(n, p)$, itself based on work of Aldous and of Nachmias and Peres; the calculations are significantly more involved in the present setting.

1. Introduction and results

In 1997, Aldous showed that inside the 'scaling window' of the phase transition, *i.e.*, when $p = (1 + \alpha n^{-1/3})/n$ with $\alpha = O(1)$, the rescaled sizes of the largest components of $G(n, p)$ converge to a certain distribution related to Brownian motion. His proof was based on a natural exploration process, introduced in the context of random graphs by Karp [16] and considered in a closely related context a little earlier by Martin-Löf [18], but with a twist: after finishing exploring a component one starts exploring the next component in such a way that a certain quantity related to the exploration behaves very much like a random walk with independent increments.

Recently, Nachmias and Peres [22] used the same process (with the same 'restarts') to study $G(n, p)$ *outside* the scaling window, giving a simpler proof of somewhat weaker forms of known results for this case; in [6] Bollobás and the author showed that with a little more work, much stronger results could be proved by analysing the same exploration process in a new way.

Nachmias and Peres [23] adapted their approach to study the phase transition in random r-regular graphs, *i.e.*, the emergence of a giant component in the subgraph of a random r-regular graph obtained by selecting edges with probability p; they showed in particular that the 'window' is when $p = (1 + \Theta(n^{-1/3}))/(r-1)$. Independently, Janson and Luczak [14] used a different approach to prove the supercritical part of this result in a more general context.

In this paper we shall extend and generalize the results mentioned in the previous paragraph. Firstly, rather than a random subgraph of a random regular graph, we study the 'configuration model' of Bollobás [3], giving a random (simple or multi-)graph with a given (here bounded) degree sequence; it is easy to see that random subgraphs of random regular graphs can be viewed in this way (see Fountoulakis [11]). Secondly, we prove much more precise results, obtaining essentially the full strength of the corresponding results for $G(n,p)$. In particular, we show that above the window the size of the giant component is asymptotically normally distributed, corresponding to the result of Pittel and Wormald [26] for $G(n,p)$, and prove results equivalent to those of Aldous [1] in the critical case. The approach used here is very much that of Bollobás and the author in [6], adapted to the configuration model.

Throughout we assume that we are 'near' the phase transition, in the range corresponding to $np \to 1$ in $G(n,p)$. Perhaps surprisingly, the proofs are easier the closer the graphs are to critical. The method used here could probably be extended to the more strongly supercritical case (covered for $G(n,p)$ in [6]), but further away from the phase transition other approaches (based on studying small components) are likely to be simpler. We assume here that the maximum degree remains bounded; this assumption can doubtless be weakened.

Before turning to the details, let us comment briefly on the history of this problem. The existence and size of the giant component in the configuration model were first studied by Molloy and Reed [20, 21], who found the size of the largest component up to a $o(n)$ error (not only near the phase transition, of course). It is easy to check that a random *subgraph* of a random graph generated by the configuration model is again an instance of the configuration model. Hence, studying the percolation phase transition in the (random) environment of the configuration model reduces to studying the transition in the configuration model itself as its parameters are varied. (This is spelled out in detail by Fountoulakis [11]; see also Janson [13].) Nevertheless, percolation on random r-regular graphs has received separate attention; for many proof techniques this special case is much easier to handle. In this special case, the critical point of the phase transition was established explicitly by Goerdt [12], and Benjamini raised the question of finding the 'window' of the phase transition (see [25, 23]). Results establishing the approximate width (either for this special case or more generally for the configuration model itself) were given recently by Kang and Seierstad [15], Pittel [25] and Janson and Luczak [14], in all cases with logarithmic gaps in the bounds. The exact width of the window, and the asymptotic size of the largest component in all ranges, was found by Nachmias and Peres [23], but only for the case of random subgraphs of random r-regular graphs. As mentioned above, here we not only extend this result to the configuration model, but greatly improve the precision, establishing not only the asymptotic size of the largest component above and below the window, but also the scale and limiting distribution of its fluctuations.

Let $\mathbf{d} = \mathbf{d}_n = (d_1^{(n)}, \ldots, d_n^{(n)})$ be a *degree sequence*, *i.e.*, a sequence of non-negative integers with even sum. For the moment we assume only that all degrees are at most some constant $d_{\max} \geqslant 2$. In the following results the sequence of course depends on n, but often we suppress this in the notation, writing d_i for the degree of vertex i, for example.

Let $G_{\mathbf{d}}$ be the *configuration multigraph* with degree sequence $\mathbf{d}$, introduced by Bollobás [3], defined as follows. Let $S_1, \ldots, S_n$ be disjoint sets with $|S_i| = d_i$; we call the elements of the S_i *stubs*. Let $\mathcal{P}$ be a pairing of $\bigcup S_i$, *i.e.*, a partition of $\bigcup S_i$ into parts of size 2, chosen uniformly at random from all such pairings. Form the multigraph $G_{\mathbf{d}}$ from $\mathcal{P}$ by replacing each pair $\{s, t\}$ with $s \in S_i$ and $t \in S_j$ by an edge with endvertices i and j.

Let us write $L_i(G)$ for the number of vertices in the ith largest (when sorted by number of vertices) component of a graph G, noting that the definition is unambiguous even if there are ties in the component sizes. Our main aim is to study the distribution of $L_1(G_{\mathbf{d}})$. We shall also consider random *simple* graphs briefly, in Theorem 1.5; when our random graphs are required to be simple, we indicate this explicitly in the notation.

When it comes to asymptotic results, we shall always make the following two assumptions. Firstly, there is a constant $d_{\max}$ such that all degrees satisfy

$$d_i^{(n)} \leqslant d_{\max}. \tag{1.1}$$

Secondly, there is a constant $c_0 > 0$ such that

$$|\{i : d_i^{(n)} \notin \{0, 2\}\}| \geqslant c_0 n \tag{1.2}$$

for all large enough n. There are two aspects to this second condition. Firstly, degree-0 vertices play no role in the construction, so we could just as well rule them out. However, in applications we shall consider sequences containing some zeros, and to avoid an extra rescaling step (where n is replaced by the number of non-zero degrees), it is convenient to allow them.

The situation with degree-2 vertices is somewhat similar: apart from the possibility of cycle components, the multigraph $G_{\mathbf{d}}$ is a random subdivision of the multigraph $G_{\mathbf{d}'}$, where $\mathbf{d}'$ is obtained by deleting all degree-2 vertices from $\mathbf{d}$. For the scaling behaviour, it is the number of (non-isolated) vertices in $\mathbf{d}'$ that matters, not the number in $\mathbf{d}$. The condition (1.2) ensures that n is (up to a constant factor) the correct scaling parameter in conditions such as $\varepsilon^3 n \to \infty$ below.

Given a degree sequence $\mathbf{d}$, let

$$\mu_r = \mu_r(\mathbf{d}) = n^{-1} \sum_{i=1}^{n} (d_i)_r \tag{1.3}$$

denote the rth *factorial moment* of $\mathbf{d}$, where $(x)_r = x(x-1)\cdots(x-r+1)$. Let

$$\lambda = \lambda(\mathbf{d}) = \frac{\sum d_i(d_i - 1)}{\sum d_i} = \frac{\mu_2}{\mu_1}. \tag{1.4}$$

As is well known (see, *e.g.*, [20, 5]), the quantity λ corresponds to the average 'branching factor' in $G_{\mathbf{d}}$. Our main interest is the 'weakly supercritical' case where $\lambda \to 1$ from above,

https://doi.org/10.1017/S0963548311000666 Published online by Cambridge University Press

but $(\lambda - 1)n^{1/3} \to \infty$, so we are outside the 'scaling window' of the phase transition. However, we shall also prove results for the critical and weakly subcritical cases.

We usually write $\lambda = \lambda(n)$ as $1 + \varepsilon(n)$, and assume throughout that

$$\lambda = \frac{\mu_2}{\mu_1} \to 1 \tag{1.5}$$

as $n \to \infty$, *i.e.*, that $\varepsilon \to 0$.

Let $\eta = \eta(n)$ denote the random variable obtained by choosing an element of $\mathbf{d}$ at random, with each element chosen with probability proportional to its value. Thus $\lambda = \mathbb{E}(\eta - 1)$ and $\varepsilon = \mathbb{E}(\eta - 2)$. The quantity

$$v_0 = \mathrm{Var}(\eta) = \mathrm{Var}(\eta - 2) \tag{1.6}$$

will play an important role in our results. An elementary calculation shows that

$$v_0 = \frac{\mu_3\mu_1 + \mu_2\mu_1 - \mu_2^2}{\mu_1^2}.$$

Under our assumptions we have $\mathbb{E}(\eta - 2) \to 0$ and, from (1.2), for large n the probability that $\eta = 2$ is bounded away from 1. Hence $\mathrm{Var}(\eta - 2) = \Theta(1)$. Since $\mu_2/\mu_1 \to 1$, it follows that

$$v_0 \sim \frac{\mu_3}{\mu_1} = \Theta(1) \tag{1.7}$$

as $n \to \infty$.

Writing $p_d = p_d(n) = n^{-1}|\{i : d_i = d\}|$ for the fraction of vertices with degree d, for $\lambda > 1$, let z be the smallest solution in $[0, 1]$ to the equation

$$z = \sum_d z^{d-1} \frac{dp_d}{\mu_1}, \tag{1.8}$$

and set

$$\rho = \rho(\mathbf{d}) = 1 - \sum_d z^d p_d. \tag{1.9}$$

This quantity is most naturally seen as the survival probability of a certain branching process (see [5]; z is essentially the probability that when we follow a random edge, we end up in a small component). What we call ρ here is exactly the quantity $\varepsilon_{\mathcal{D}}$ appearing in the result of Molloy and Reed [21]. Let

$$\rho^* = \rho^*(\mathbf{d}) = \sum_d z^d p_d - \frac{\mu_1 z^2}{2} - 1 + \frac{\mu_1}{2}; \tag{1.10}$$

it will turn out that $\rho^* n$ will give asymptotically the *nullity* (or excess) of the giant component $\mathcal{C}_1$, *i.e.*, the difference $e(\mathcal{C}_1) - (|\mathcal{C}_1| - 1)$ between the number of edges and the number of edges of a tree with the same order. Although we shall not prove this, this is also asymptotically the number of vertices in the 'kernel', *i.e.*, the multi-graph formed from the 2-core of giant component by contracting vertices with degree 2.

We shall show later (see Lemma 3.4(v) and (5.12)) that

$$\rho \sim \frac{2\mu_1^2}{\mu_3}\varepsilon \quad \text{and} \quad \rho^* \sim \frac{2\mu_1^3}{3\mu_3^2}\varepsilon^3. \tag{1.11}$$

Molloy and Reed [21] showed, under different conditions, that $L_1(G_{\mathbf{d}}) = \rho(\mathbf{d})n + o_{\mathrm{p}}(n)$. (Their result allowed much larger degrees, and they did not assume that $\varepsilon \to 0$; indeed, their result does not 'bite' in this case.) Our main result concerns the fluctuations of $L_1(G_{\mathbf{d}})$ around $\rho(\mathbf{d})n$ in the weakly supercritical case.

Theorem 1.1. *Let $d_{\max} \geqslant 2$ and $c_0 > 0$ be fixed. For each n let $\mathbf{d} = \mathbf{d}_n$ be a degree sequence satisfying* (1.1) *and* (1.2). *Define μ_i, λ, ρ and ρ^* as above, noting that these quantities depend on n. Setting $\varepsilon = \lambda - 1$, suppose that $\varepsilon \to 0$ and $\varepsilon^3 n \to \infty$. Let L_1 and N_1 denote the order and nullity of the largest component of $G_{\mathbf{d}}$. Then $L_1' = L_1 - \rho n$ and $N_1' = N_1 - \rho^* n$ are asymptotically jointly normally distributed with mean* 0,

$$\operatorname{Var}(L_1') \sim 2\mu_1\varepsilon^{-1}n, \quad \operatorname{Var}(N_1') \sim 5\rho^* n \sim \frac{10\mu_1^3}{3\mu_3^2}\varepsilon^3 n, \quad \textit{and} \quad \operatorname{Cov}(L_1', N_1') \sim \frac{2\mu_1^2}{\mu_3}\varepsilon n.$$

Furthermore,

$$L_2(G_{\mathbf{d}}) = O_{\mathrm{p}}\big(\varepsilon^{-2}\log(\varepsilon^3 n)\big). \tag{1.12}$$

Here, as usual, given a sequence (Z_n) of random variables and a deterministic function $f(n)$, $Z_n = O_{\mathrm{p}}(f(n))$ means that $Z_n/f(n)$ is bounded in probability, *i.e.*, for any $\delta > 0$ there exists C such that $\mathbb{P}(|Z_n| \leqslant Cf(n)) \geqslant 1 - \delta$ for all (large enough) n. We say that an event (formally a sequence of events) holds *with high probability*, or w.h.p., if its probability tends to 1 as $n \to \infty$. We write $Z_n = o_{\mathrm{p}}(f(n))$ if $Z_n/f(n)$ converges to 0 in probability, *i.e.*, if for any $\delta > 0$ we have $|Z_n| \leqslant \delta f(n)$ w.h.p.

We assume a bounded maximum degree for simplicity. The proof extends to the case where the maximum degree grows reasonably slowly; we have not investigated this further.

The asymptotic correlation coefficient $\operatorname{Cov}(L_1', N_1')/\sqrt{\operatorname{Var}(L_1')\operatorname{Var}(N_1')}$ given by Theorem 1.1 is simply $\sqrt{3/5}$. Although this case is not covered by our result, if we take the degree distribution to be Poisson as in $G(n,p)$, then when $p \sim 1/n$ we have $\mu_i \sim 1$ for all i, and the variance and covariance formulae above are consistent with those given by Pittel and Wormald [26, Note 4] for $G(n,p)$.

The proof of Theorem 1.1 will show that for each d the number $L_1(d)$ of degree-d vertices in the largest component satisfies

$$L_1(d) = n_d(1 - z^d) + O_{\mathrm{p}}(\sqrt{n/\varepsilon}), \tag{1.13}$$

where $n_d = np_d$ is the total number of degree-d vertices and z is defined by (1.8). The parameter z corresponds to ξ in [14], so (1.13) refines the results there. Our method would allow us to establish joint normality of these numbers with variances and covariances of order n/ε, but we shall not give the details.

We next consider the subcritical case. Writing, as before, $p_d = p_d(n)$ for the proportion of vertices with degree d, let $q_d = dp_d/\mu_1 = \mathbb{P}(\eta = d)$ be the corresponding size-biased distribution. Let a_n be the (unique – see Section 7) solution to $\sum(d-2)q_d e^{a_n(d-2)} = 0$, and define

$$\delta_n = -\log\left(\sum_d q_d e^{a_n(d-2)}\right). \tag{1.14}$$

Theorem 1.2. *Let $d_{\max} \geqslant 2$ and $c_0 > 0$ be fixed. For each n let $\mathbf{d} = \mathbf{d}_n$ be a degree sequence satisfying* (1.1) *and* (1.2). *Define μ_i and λ as above, noting that these quantities depend on n. Setting $\varepsilon = 1-\lambda$, suppose that $\varepsilon \to 0$ and $\varepsilon^3 n \to \infty$. Then, for all $x \in \mathbb{R}$ we have*

$$\mathbb{P}\big(L_1(G_{\mathbf{d}}) \leqslant \delta_n^{-1}(\log\Lambda - \tfrac{5}{2}\log\log\Lambda + x)\big) = \exp\big(-ce^{-x}\big) + o(1), \tag{1.15}$$

where $\Lambda = \varepsilon^3 n$, $c = c(\mathbf{d}) = \Theta(1)$ is given by (7.7), *and δ_n, defined in* (1.14), *satisfies*

$$\delta_n \sim \frac{\varepsilon^2}{2v_0} \sim \frac{\varepsilon^2\mu_1}{2\mu_3}. \tag{1.16}$$

In particular,

$$L_1(G_{\mathbf{d}}) = \delta_n^{-1}\big(\log\Lambda - \tfrac{5}{2}\log\log\Lambda + O_{\mathrm{p}}(1)\big), \tag{1.17}$$

and $L_1(G_{\mathbf{d}}) = (2\mu_3/\mu_1 + o_{\mathrm{p}}(1))\varepsilon^{-2}\log\Lambda$.

The bound (1.12) in Theorem 1.1 is proved by applying Theorem 1.2 to what remains of the supercritical graph after deleting the largest component, so in Theorem 1.1 we in fact obtain bounds of the type (1.15), (1.17) but with c, δ_n and Λ defined for the 'dual' distribution with $n_d z^d$ vertices of each degree d; see (1.13) and the remark at the start of Section 1.1.

Theorem 1.2 is the equivalent of (the corrected form of – see [5]) Łuczak's extension [17] of Bollobás's result [4] for $G(n,p)$ in the subcritical case.

Finally, in the critical case, we obtain an analogue of the results of Aldous [1] for $G(n,p)$. The statement requires a few definitions, analogous to those in [1].

Let $(W(s))_{0\leqslant s<\infty}$ be a standard Brownian motion. Given real numbers α_0, α_1 and α_2 with $\alpha_0 > 0$, let

$$W_{\alpha_0,\alpha_1,\alpha_2}(s) = \alpha_0^{1/2}W(s) + \alpha_1 s - \frac{\alpha_2 s^2}{2} \tag{1.18}$$

be a rescaled Brownian motion with drift $\alpha_1 - \alpha_2 s$ at time s, and set

$$B(s) = B_{\alpha_0,\alpha_1,\alpha_2}(s) = W_{\alpha_0,\alpha_1,\alpha_2}(s) - \min_{0\leqslant s'\leqslant s} W_{\alpha_0,\alpha_1,\alpha_2}(s'). \tag{1.19}$$

Also, define a process $N(s)$ of 'marks' so that, given $B(s)$, $N(s)$ is a Poisson process with intensity $\beta B(s)\mathrm{d}s$, where $\beta > 0$ is constant. (For the formal details, see [1].) Finally, order the excursions γ_j of $B(s)$, *i.e.*, the maximum intervals on which $B(s)$ is strictly positive, in decreasing order of their lengths $|\gamma_j|$. Writing $N(\gamma_j)$ for the number of marks in γ_j, this defines a joint distribution

$$\big(|\gamma_j|, N(\gamma_j)\big)_{j\geqslant 1} \tag{1.20}$$

that depends on the parameters α_0, α_1, α_2 and β.

Theorem 1.3. *Suppose that* $\mathbf{d} = \mathbf{d}_n$ *satisfies assumptions* (1.1) *and* (1.2) *above. Suppose also that* $n^{1/3}(\lambda - 1)$ *converges to some* $\alpha_1 \in \mathbb{R}$, *that* $\mu_3/\mu_1 \to \alpha_0$, *that* $\mu_3/\mu_1^2 \to \alpha_2$, *and that* $1/\mu_1 \to \beta$. *Let* $\mathcal{C}_1$, $\mathcal{C}_2, \ldots$ *be the components of* $G_{\mathbf{d}}$ *ordered* (*largest first*) *by number of vertices. Then, for any fixed* r *the sequence* $(n^{-2/3}|\mathcal{C}_j|, n(\mathcal{C}_j))_{j=1}^r$ *converges in distribution to the first* r *terms of the sequence* $(|\gamma_j|, N(\gamma_j))$ *defined above, where* $n(\mathcal{C}_j)$ *is the nullity of* $\mathcal{C}_j$.

Note that in the light of (1.7), assuming $\mu_3/\mu_1 \to \alpha_0$ is equivalent to assuming $v_0 \to \alpha_0$.

We have written Theorem 1.3 with the scaling that arises most naturally in the proof. From the scaling properties of Brownian motion, one can check that $B_{\alpha_0,\alpha_1,\alpha_2}(s)$ is equal in distribution (as a process) to $\alpha' B_{1,\alpha_1',1}(\alpha_0' s)$, where $\alpha' = \alpha_0^{2/3}\alpha_2^{-1/3}$, $\alpha_0' = \alpha_2^{2/3}\alpha_0^{-1/3}$ and $\alpha_1' = \alpha_1\alpha_0^{-1/3}\alpha_2^{-1/3}$. Hence, if we consider only the component sizes, there is a single-parameter family of limiting processes, characterized by α_1', the limiting value of $n^{1/3}(\lambda - 1)\mu_1\mu_3^{-2/3}$. Noting that the excursion lengths are scaled by α_0', Theorem 1.3 shows that if $n^{1/3}(\lambda - 1)\mu_1\mu_3^{-2/3} \to \alpha$, then $(n^{-2/3}\mu_3^{1/3}\mu_1^{-1}L_i)_{i=1}^r$ converges to the first r sorted excursion lengths of $B_{1,\alpha,1}$. These excursion lengths are exactly the rescaled component sizes appearing in Aldous's result for $G(n, 1/n + \alpha n^{-4/3})$.

If we also consider the nullities, or mark counts, then there is a two-parameter family of possible limits.

1.1. Applications and extensions

Let $\mathbf{d}$ satisfy the assumptions above (in particular, (1.1), (1.2), (1.5) and $\Lambda = \varepsilon^3 n \to \infty$). Observing at all times the restriction that all degrees are at most $d_{\max}$, changing m entries of $\mathbf{d}$ changes the proportion p_d of degree-d vertices by $O(m/n)$, and hence changes the quantities μ_1, μ_2 and μ_3, which are all of order 1, by $O(m/n)$. It follows that $\varepsilon = \lambda - 1 = \mu_2/\mu_1 - 1$ changes by an absolute amount that is $O(m/n)$, so if $m = o(\varepsilon n)$ the relative change in ε is $O(m/(\varepsilon n))$. It is easy to see, and will follow from the results later in the paper, that the relative change in ρ, ρ^*, $\Lambda = \varepsilon^3 n$ or δ_n is of this same order $O(m/(\varepsilon n))$, as one would expect from the formulae (1.11) and (1.16).

When $m = o(\sqrt{n/\varepsilon})$, the changes in ρn and $\rho^* n$ are small compared to the relevant standard deviations, and it follows that Theorem 1.1 applies just as well to $G_{\mathbf{d}'}$ for any $\mathbf{d}_n'$ obtained from $\mathbf{d}_n$ by changing $o(\sqrt{n/\varepsilon})$ entries, even when all quantities in the conclusion of the theorem are calculated for $\mathbf{d}$ rather than for $\mathbf{d}'$. Similarly, if $m = O(\sqrt{n/\varepsilon})$ then the relative change in δ_n is $O(1/\sqrt{\Lambda}) = o(1/\log \Lambda)$, and it is not hard to check that Theorem 1.2 applies to $G_{\mathbf{d}'}$ in this case; for Theorem 1.3, the corresponding condition is simply $m = o(n^{2/3})$, so that the limit of $n^{1/3}(\lambda - 1)$ is unchanged. Note that $m = O(\sqrt{n})$ satisfies the bound on m in all three cases.

One application of these observations concerns random subgraphs of random r-regular graphs, or indeed random subgraphs of configuration (multi-)graphs.

Let $\mathbf{d}$ be the (random) degree sequence of the graph $G_r[p]$ obtained by starting with an r-regular configuration multigraph, selecting edges independently with probability p, and retaining the selected edges. Conditional on $\mathbf{d}$, the distribution of $G_r[p]$ is simply that of $G_{\mathbf{d}}$. Hence, as noted by Fountoulakis [11], one can study $G_r[p]$ by studying $G_{\mathbf{d}}$.

It is very easy to check that for $0 \leqslant d \leqslant r$, the proportion p_d of degree-d vertices in $\mathbf{d}$ satisfies

$$p_d = p_d^0 + O_{\mathrm{p}}(n^{-1/2}) \quad \text{where} \quad p_d^0 = \binom{r}{d} p^d (1-p)^{r-d}.$$

It follows that the quantities μ_i defined above satisfy

$$\mu_i = \mu_i^0 + O_{\mathrm{p}}(n^{-1/2}) \quad \text{where} \quad \mu_i^0 = p^i (r)_i = p^i r(r-1)\cdots(r-i+1).$$

Note that $\lambda^0 = \mu_2^0/\mu_1^0 = p(r-1)$, and when $\lambda^0 \to 1$, *i.e.*, $p \sim 1/(r-1)$, then $\mu_1^0 \sim r/(r-1)$ and $\mu_3^0 \sim r(r-2)/(r-1)^2$. From the remarks above, even though the actual number of vertices of each degree is random, $G_{\mathbf{d}}$ will satisfy the conclusions of Theorems 1.1–1.3 for the idealized sequence with p_d replaced by p_d^0. Hence, defining ρ^0 by (1.8) and (1.9) with p_d replaced by p_d^0, Theorems 1.1–1.3 have the following consequence. (We omit the nullity result and the analogue of (1.15) for simplicity; these also carry over.)

Corollary 1.4. *Let $r \geqslant 3$ be fixed and let $p = (1+\varepsilon)/(r-1)$ where $\varepsilon = \varepsilon(n) \to 0$. Let G be the random subgraph of the random r-regular configuration multigraph on n vertices obtained by selecting edges independently with probability p.*

If $\varepsilon > 0$ and $\varepsilon^3 n \to \infty$ then

$$\frac{L_1(G) - \rho^0 n}{\sigma\sqrt{n}} \xrightarrow{\mathrm{d}} N(0,1),$$

where ρ^0 (defined above) satisfies $\rho^0 \sim 2\varepsilon r/(r-2)$, and $\sigma^2 = 2\varepsilon^{-1} r/(r-1)$.

If $\varepsilon < 0$ and $|\varepsilon|^3 n \to \infty$ then

$$L_2(G) = \delta_n^{-1}\big(\log\Lambda - \tfrac{5}{2}\log\log\Lambda + O_{\mathrm{p}}(1)\big),$$

where $\Lambda = |\varepsilon|^3 n$ and δ_n, defined by (1.14) with p_d^0 in place of p_d, satisfies $\delta_n \sim \frac{r-1}{2(r-2)}\varepsilon^2$.

Finally, if $n^{1/3}\varepsilon \to \alpha_1 \in \mathbb{R}$ then the sizes and nullities of the components sorted in decreasing order of size converge in distribution to the distribution described in (1.20) with $\alpha_0 = (r-2)/(r-1)$, $\alpha_2 = (r-2)/r$ and $\beta = (r-1)/r$. □

Note that this result is consistent with, but much sharper than, the results of Nachmias and Peres [23]. Of course, one can formulate a similar corollary concerning the random subgraph $G_{\mathbf{d}}[p]$ of $G_{\mathbf{d}}$, for any degree sequence $\mathbf{d}$ satisfying (1.1) and (1.2); passing to the subgraph multiplies μ_i by p^i, just as in the r-regular case, and p must be chosen so that the value of the 'branching factor' λ in the subgraph (p times that in the original) is of the form $1+\varepsilon$ with $\varepsilon \to 0$.

As shown by Bollobás [3] (see also [2]), when the maximum degree is bounded, the probability that the configuration multigraph $G_{\mathbf{d}}$ is simple is bounded away from 0, and conditional on this event, $G_{\mathbf{d}}$ has the distribution of $G_{\mathbf{d}}^{\mathrm{s}}$, a uniformly random simple graph with degree sequence $\mathbf{d}$. It follows that any 'w.h.p.' results for $G_{\mathbf{d}}$ transfer to $G_{\mathbf{d}}^{\mathrm{s}}$. This applies to Theorem 1.2, in the weaker form (1.17), but not to the other results above. More precisely, as shown in [3] (or, after translating from the enumerative to probabilistic viewpoint, [2]), when the maximum degree is bounded, the probability p that $G_{\mathbf{d}}$ is simple satisfies $p \sim \exp(-\theta-\theta^2)$, where $\theta = \sum_i \binom{d_i}{2} / \sum_i d_i$, which in our notation is $\mu_2/(2\mu_1)$.

The proofs of Theorems 1.1–1.3 involve 'exploring' part of the graph. We can end these explorations when certain entire components have been revealed, comprising in total $o(n)$ vertices. It is easy to check that the probability of encountering a loop or multiple edge in the exploration is $o(1)$. Moreover, the unexplored part of the graph may be seen as a configuration multigraph G'. Since only $o(n)$ vertices have been explored, the (conditional) probability that G' is simple is $p + o(1)$: the corresponding θ is within $o(1)$ of the original θ. It follows that if $\mathcal{E}$ is some not too unlikely event defined in terms of our exploration, then the probability that $\mathcal{E}$ holds and $G_{\mathbf{d}}$ is simple is $(\mathbb{P}(\mathcal{E}) + o(1))(p + o(1)) \sim p\mathbb{P}(\mathcal{E})$. Thus, conditioning on $G_{\mathbf{d}}$ being simple hardly changes the probability of $\mathcal{E}$. Using this observation it is easy to transfer the results above to random simple graphs; we omit the details.

Theorem 1.5. *All the results above apply unchanged if $G_{\mathbf{d}}$ is replaced by the random simple graph $G^{\mathrm{s}}_{\mathbf{d}}$.* □

Note that in the analogue of Corollary 1.4 this means that we consider a random subgraph of a random r-regular simple graph. Here one must be slightly careful with the argument: we need to explore the subgraph, but then check whether the *original* graph is simple.

The rest of the paper is organized as follows. In the next section we define the exploration process that we study, and two corresponding random walks (X_t) and (Y_t). In Section 3 we establish some key properties of (X_t), including the 'idealized trajectory' that we expect it to remain close to. Using these properties, and assuming Theorem 1.2 for the moment, we prove Theorem 1.3 in Section 4 and Theorem 1.1 in Section 5. In Section 6 we prove a local limit theorem (Lemma 6.3) for certain sums of independent random variables. Finally, in Section 7 we prove Theorem 1.2. The proof (which uses the local limit theorem) is rather different from that of the other main results: we use domination arguments to study the initial behaviour of (X_t), rather than following its evolution as in the main part of the paper. This is the reason for postponing the proof, even though (a weak form of) the result is needed in the proofs of Theorems 1.1 and 1.3, to rule out 'other' large components.

2. The exploration process

Consider the following exploration process for uncovering the components of the configuration multigraph $G = G_{\mathbf{d}}$. This is a slight variant of the standard process, in that we check for 'back-edges' forming cycles as we go. A form of this variant was used by Ding, Kim, Lubetzky and Peres [9] in their study of the diameter of $G(n, p)$; we could in fact use a more standard exploration here, but the variant results in slightly cleaner calculations.

We shall define an exploration so that after t steps of the process, t vertices have been 'reached'; the other $n - t$ are 'unreached'. Furthermore, a certain random number of stubs will have been paired with each other, and each unpaired stub will be either 'active' or 'unreached'; the active stubs are attached to reached vertices, the unreached ones to unreached vertices. We write A_t for the number of active stubs, and U_t for the

https://doi.org/10.1017/S0963548311000666 Published online by Cambridge University Press

number of unreached stubs, noting that $A_0 = 0$. The process we define will be such that in the complete pairing $\mathcal{P}$, the active stubs are paired to a subset of the unreached stubs, and the remaining unreached stubs are paired with each other. Moreover, the conditional distribution of the pairing $\mathcal{P}$ given the first t steps of the process is such that all pairings of the active and unreached stubs satisfying this condition are equally likely.

At step $t+1$ of the process, if $A_t > 0$ then we pick an active stub a_{t+1}, for example the first in some order fixed in advance. Then we reveal its partner u_{t+1} in the random pairing $\mathcal{P}$, which is necessarily unreached. Let v_{t+1} be the corresponding unreached vertex. The stubs a_{t+1} and u_{t+1} are now paired (so in particular a_{t+1} is no longer active), and the remaining stubs $u_{t+1,1}, \ldots, u_{t+1,r_{t+1}}$ attached to v_{t+1} are provisionally declared active. (Here r_{t+1}, the number of these remaining stubs, is equal to $d(v_{t+1}) - 1$.) But now:

(i) we check whether any other (previously) active stubs are paired to any of the $u_{t+1,i}$, and then
(ii) we check whether any of the remaining $u_{t+1,i}$ are paired to each other.

We declare any pairs found in (i) or (ii) 'paired', and continue. We call edges corresponding to these pairs 'back-edges' since they go 'back' to already-reached vertices, though in case (ii) the edges are loops.

If $A_t = 0$ then we simply pick v_{t+1} to be a random unreached vertex, chosen with probability proportional to degree, provisionally declare all $d(v_{t+1})$ of its stubs active, and then perform the second check (ii) above on these stubs. In this case we say that we 'start a new component' at step $t+1$.

For (partial) compatibility with the notation in [6], let η_{t+1} denote the degree of v_{t+1}. We write θ_{t+1} for the number of back-edges found during step $t+1$, and $Y_t = \sum_{i \leqslant t} \theta_i$ for the total number of back-edges found during the first t steps.

Let C_t be the number of components that we have started exploring within the first t steps, and set

$$X_t = A_t - 2C_t. \tag{2.1}$$

Considering separately the cases $A_t > 0$ and $A_t = 0$, and noting that finding a back-edge pairs off two stubs, we see that

$$X_{t+1} - X_t = \eta_{t+1} - 2 - 2\theta_{t+1}, \tag{2.2}$$

while by the definition of Y_t,

$$Y_{t+1} - Y_t = \theta_{t+1}. \tag{2.3}$$

Let $\mathcal{F}_t$ denote the (finite, of course) sigma-field generated by the information revealed by step t. Let $U_{d,t}$ denote the number of unreached *vertices* of degree d after t steps, so, recalling that U_t denotes the (total) number of unreached *stubs*, we have

$$U_t = \sum_d dU_{d,t}. \tag{2.4}$$

In both cases above, the vertex v_{t+1} is chosen from the unreached vertices and, given $\mathcal{F}_t$, the probability that any given vertex is chosen is proportional to its degree. Hence

$$\mathbb{P}(\eta_{t+1} = d \mid \mathcal{F}_t) = dU_{d,t}/U_t. \tag{2.5}$$

In particular,

$$\mathbb{E}(\eta_{t+1} - 1 \mid \mathcal{F}_t) = U_t^{-1} \sum_d d(d-1)U_{d,t}.$$

In the analysis that follows, we shall impose the assumption

$$t \leqslant c_0 n/2, \tag{2.6}$$

where c_0 is the constant in (1.2). Since at least $c_0 n$ vertices have degree at least 1, (2.6) implies that

$$U_t \geqslant c_0 n/2, \tag{2.7}$$

and in particular that $U_t = \Theta(n)$. This will simplify some formulae in the calculations.

Suppose that $A_t > 0$. Then, given $\mathcal{F}_t$ and v_{t+1}, the expected number of pairs discovered during check (i) above is exactly

$$(A_t - 1)\frac{\eta_{t+1} - 1}{U_t - 1} = (\eta_{t+1} - 1)A_t/U_t + O(1/n),$$

using $U_t = \Theta(n)$ for the approximation. Indeed, each of the $A_t - 1$ other stubs that were active before this step is equally likely to be paired to any of the $U_t - 1$ remaining unreached stubs. We could write an exact formula for the expected number of pairs found during check (ii), but there is no need: instead we simply note that the expectation is $O(1/n)$. It follows that

$$\mathbb{E}(\theta_{t+1} \mid \mathcal{F}_t, \eta_{t+1}) = (\eta_{t+1} - 1)A_t/U_t + O(1/n), \tag{2.8}$$

and hence that

$$\mathbb{E}(\theta_{t+1} \mid \mathcal{F}_t) = \mathbb{E}(\eta_{t+1} - 1 \mid \mathcal{F}_t)A_t/U_t + O(1/n). \tag{2.9}$$

The last two formulae are also valid when $A_t = 0$: this time there is no check (i), and the expected number of back-edges found during check (ii) is $O(1/n)$. From (2.2) and (2.1) it follows that

$$\mathbb{E}(X_{t+1} - X_t \mid \mathcal{F}_t) = -1 + \mathbb{E}(\eta_{t+1} - 1 \mid \mathcal{F}_t)(1 - 2X_t/U_t) + O(C_t/n). \tag{2.10}$$

We use $C_t \geqslant 1$ for $t \geqslant 1$ to avoid writing a separate $O(1/n)$ error term, not that it matters.

Simply knowing the expected changes at each step is good enough to allow us to deduce fairly tight bounds on the size of the giant component. But for asymptotic normality we need a bound on the variance. We could give a formula that is useful when A_t is comparable with U_t, but we shall not need this. Instead, we simply note that the number θ_{t+1} of pairs found during our checks (i) and (ii) is bounded, and for $t = o(n)$, which implies $A_t = o(U_t)$, we have $\theta_{t+1} = 0$ w.h.p. Since η_{t+1} is bounded, it follows easily that for $t = o(n)$ we have

$$\mathrm{Var}(X_{t+1} - X_t \mid \mathcal{F}_t) = \mathrm{Var}(\eta_{t+1} \mid \mathcal{F}_t) + o(1) = v_0 + o(1),$$

where v_0 is defined in (1.6), and the second equality follows from the fact that only $o(n)$ vertices have been 'used up' by time t.

Since $A_t \geqslant 0$, with equality only when we have just finished exploring a component, the times $t_1, \ldots, t_{c(G)}$ at which we finish exploring components are given by

$$t_i = \min\{t : X_t = -2i\}. \tag{2.11}$$

Since exactly one vertex is revealed at each stage, if $\mathcal{C}_i$ denotes the ith component explored, then

$$|\mathcal{C}_i| = t_i - t_{i-1},$$

where we set $t_0 = 0$.

Recall that Y_t denotes the number of back-edges found within the first t steps. Then $Y_{t_i} - Y_{t_{i-1}}$ is simply the nullity of $\mathcal{C}_i$:

$$n(\mathcal{C}_i) = Y_{t_i} - Y_{t_{i-1}}. \tag{2.12}$$

In the rest of the paper we shall study the behaviour of the random walks (X_t) and (Y_t), and use this to prove our main results.

3. Trajectory and deviations

As in [6], we shall write the difference $X_{t+1} - X_t$ as $D_{t+1} + \Delta_{t+1}$, where $D_{t+1} = \mathbb{E}(X_{t+1} - X_t \mid \mathcal{F}_t)$, so the Δ_t may be regarded as a sequence of martingale differences. It is more or less automatic that $\sum_{i \leqslant t} \Delta_i$ is asymptotically normally distributed, so we need to understand the sum of the D_t. Each D_t depends on the earlier X_i, but it turns out that the dependence is not very strong. So if we can find an 'ideal' trajectory (corresponding to all Δ_t being equal to zero), then bounding the deviations of (X_t) from this trajectory will not be too difficult.

The first problem is that the term U_t appearing in (2.5) is not so simple. We start with some lemmas. The first is a standard result about order statistics.

Lemma 3.1. *Let $Z_1, \ldots, Z_n$ be i.i.d. samples from a distribution Z on $[0,1]$ with distribution function $G(x) = \mathbb{P}(Z \leqslant x)$, and let $N_n(x) = |\{i : Z_i \leqslant x\}|$. Then, for any (deterministic) function $y = y(n)$ we have*

$$\sup_{0 \leqslant x \leqslant y} |N_n(x) - nG(x)| = O_{\mathrm{p}}(\sqrt{nG(y)}). \tag{3.1}$$

Proof. Let A have a Poisson distribution with mean n, and given A, let $Z'_1, \ldots, Z'_A$ be i.i.d. with distribution Z, so the set $\{Z'_i\}$ forms a Poisson process on $[0,1]$. (If Z has a density function $g(x) = G'(x)$, then the intensity measure of the Poisson process is $ng(x)\mathrm{d}x$.) Writing $N'(x)$ for the number of Z'_i in $[0,x]$, consider the random function $F(x) = N'(x) - nG(x)$. From basic properties of Poisson processes, this function is a continuous-time martingale on $[0,1]$ with independent increments. Hence Doob's maximal

https://doi.org/10.1017/S0963548311000666 Published online by Cambridge University Press

inequality [10, Ch. III, Theorem 2.1] gives

$$\mathbb{P}\Big(\sup_{x\leqslant y}|F(x)|\geqslant t\Big)\leqslant\frac{\mathrm{Var}(F(y))}{t^2}=\frac{nG(y)}{t^2},$$

since $\mathrm{Var}(F(y))=\mathrm{Var}(N'(y))$, and up to an additive constant, $N'(y)$ is simply Poisson with mean $nG(y)$. It follows that

$$\sup_{x\leqslant y}|F(x)|=O_{\mathrm{p}}(\sqrt{nG(y)}). \tag{3.2}$$

The expected value of $|A-n|$ is $O(\sqrt{n})$. When $A>n$, delete a random subset of the points $\{Z_i'\}$ of size $A-n$. When $A<n$, add $n-A$ i.i.d. new points to $\{Z_i'\}$ with distribution Z. In this way we obtain a set of n i.i.d. samples from Z. Given A, the added/deleted points have distribution Z, so the expected number in $[0,y]$ is $|A-n|G(y)$. Hence the unconditional expectation of the number of points added or deleted in $[0,y]$ is $O(\sqrt{n}G(y))=O(\sqrt{nG(y)})$. Hence (3.1) follows from (3.2). □

In our exploration process, the next vertex v_{t+1} is always chosen with probability proportional to its degree. Hence the (distribution of) the random sequence $\sigma=(v_1,\dots,v_n)$ has the following alternative description: first assign a random order to all stubs. Then sort the vertices so that v comes before w if and only if v's earliest stub comes before w's earliest stub. In turn, we may realize the random order on the stubs by assigning i.i.d. $U[0,1]$ variables to the stubs; we shall call these variables *stub values.*

For each t, there is a random 'cut-off' $Z_t\in[0,1]$ so that a vertex v is among $v_1,\dots,v_t$ if and only if its smallest stub value is at most $1-Z_t$. Fixing a cut-off value z, the expected number of vertices with all stub values at least $1-z$ is exactly $\sum_d n_d z^d$, so we expect to have

$$n-t\sim\sum_d n_d Z_t^d.$$

Let $f(z)=f_n(z)$ denote the probability generating function of the degree distribution $\mathbf{d}$, so

$$f(z)=\sum_{d=1}^{d_{\max}}\frac{n_d}{n}z^d. \tag{3.3}$$

Recall that $U_{d,i}$ denotes the number of unreached degree-d vertices after i steps, *i.e.*, the number of degree-d vertices *not* among the first i elements of our random order on the vertices.

Theorem 3.2. *Let $\mathbf{d}=\mathbf{d}_n$ be any degree sequence of length n with all degrees between 1 and some constant $d_{\max}$, and let $(v_1,\dots,v_n)$ be the random order on the vertices defined above. Define a function $\tau\mapsto z(\tau)$ from $[0,1]$ to $[0,1]$ by $f(z(\tau))+\tau=1$, where f is the probability generating function of $\mathbf{d}$. Then, for any $t=t(n)\geqslant 1$ we have*

$$\max_{1\leqslant d\leqslant d_{\max}}\max_{1\leqslant i\leqslant t}|U_{d,i}-n_d z(i/n)^d|=O_{\mathrm{p}}(\sqrt{t}),$$

where n_d is the number of degree-d vertices in $\mathbf{d}$.

Proof. Construct the sequence $v_1,\ldots,v_n$ from stub values as above. Let $Z^{(d)}$ denote the distribution obtained by taking the minimum of d independent $U[0,1]$ random variables. Note that $Z^{(d)}$ has distribution function $G_d(x) = 1-(1-x)^d$.

For each d, let $Z_{d,i}$ denote the minimum stub value of the ith degree-d vertex. Then the random variables $Z_{d,1},\ldots,Z_{d,n_d}$ are i.i.d. with distribution $Z^{(d)}$. Let $m_d(x)$ denote the number of degree-d vertices whose smallest stub value is at most x. Then by Lemma 3.1, for any (deterministic) $y = y(n)$ and any d we have

$$E_{d,y} = \sup_{0\leqslant x\leqslant y} |m_d(x) - n_d G_d(x)| = O_{\mathrm{p}}(\sqrt{n_d G_d(y)}).$$

Summing over d, and using $\sum_{i=1}^k \sqrt{a_i} \leqslant \sqrt{k}\sqrt{\sum a_i}$, we have

$$E_y = \sum_d E_{d,y} = O_{\mathrm{p}}(\sqrt{nG(y)}),$$

where

$$G(x) = \sum_{d=1}^{d_{\max}} \frac{n_d}{n} G_d(x) = 1 - f(1-x). \tag{3.4}$$

Let $m(x) = \sum_d m_d(x)$. Then by the triangle inequality we have

$$\sup_{0\leqslant x\leqslant y} |m(x) - nG(x)| \leqslant E_y. \tag{3.5}$$

Define y by $nG(y) = t$, so $E_y = O_{\mathrm{p}}(\sqrt{t})$. For $1 \leqslant i \leqslant t$ set $x_i = 1 - z(i/n)$. From (3.4) we have $G(x_i) = 1 - f(z(i/n))$, so by the definition of the function z, we have $G(x_i) = i/n$, *i.e.*, $nG(x_i) = i$. Note that $x_1 < x_2 < \cdots < x_t = y$.

From (3.5), for every $i \leqslant t$ we have $|m(x_i) - i| \leqslant E_y$. Furthermore, the number of degree-d vertices among the first $m(x_i)$ vertices, namely $m_d(x_i)$, differs from $n_d G_d(x_i)$ by at most $E_{d,y} \leqslant E_y$. It follows that the number of degree-d vertices among the first i vertices differs from $n_d G_d(x_i)$ by at most $2E_y$, *i.e.*,

$$|(n_d - U_{d,i}) - n_d G_d(x_i)| \leqslant 2E_y.$$

The difference on the left is exactly $|U_{d,i} - n_d z(i/n)^d|$, so the result follows. □

Returning to our process, recall that the vertices $v_1, v_2, \ldots$ are chosen according to the random distribution considered above. Recall also that when (2.6) holds, then $U_t = \Theta(n)$.

Corollary 3.3. *Define $z(\tau)$ by $f(z(\tau)) + \tau = 1$. Then, for any $t \leqslant c_0 n/2$, writing z for $z(i/n)$, we have*

$$\sup_{i\leqslant t} \left| \frac{U_i}{n} - zf'(z) \right| = O_{\mathrm{p}}(\sqrt{t}/n) \tag{3.6}$$

https://doi.org/10.1017/S0963548311000666 Published online by Cambridge University Press

and

$$\sup_{i\leqslant t}\left|\mathbb{E}\big(\eta_{i+1}-1 \mid \mathcal{F}_i\big)-\frac{zf''(z)}{f'(z)}\right| = O_{\mathrm{p}}(\sqrt{t}/n). \tag{3.7}$$

Furthermore, if $t=o(n)$*, then*

$$\sup_{i\leqslant t}\left|\mathrm{Var}\big(\eta_{i+1}-1 \mid \mathcal{F}_i\big)-v_0\right| = O(t/n)=o(1),$$

where v_0 *is defined by* (1.6).

Proof. By Theorem 3.2 there is some random K_t such that $K_t = O_{\mathrm{p}}(\sqrt{t})$ and, for all $i\leqslant t$ and all $d\leqslant d_{\max}$,

$$|U_{d,i}-n_d z(i/n)^d| \leqslant K_t. \tag{3.8}$$

Fix i and write z for $z(i/n)$. Noting that $zf'(z)=n^{-1}\sum_d dn_d z^d$, (2.4) and (3.8) give

$$|U_i/n - zf'(z)| = \frac{1}{n}\left|\sum_{d=1}^{d_{\max}} dU_{d,i} - \sum_d dn_d z^d\right| \leqslant d_{\max}^2 K_t/n.$$

Since $K_t = O_{\mathrm{p}}(\sqrt{t})$, this proves (3.6).

For (3.7), note that $\mathbb{P}(\eta_{i+1}=d \mid \mathcal{F}_i)$ is by definition equal to $dU_{d,i}/U_i$ (see (2.5)), so when (3.8) holds, $\mathbb{P}(\eta_{i+1}=d \mid \mathcal{F}_i)$ is within $O(K_t/n)$ of

$$q_d(z) = \frac{dn_d z^d}{\sum_{d'} d' n_{d'} z^{d'}}.$$

The bound (3.7) follows by noting that

$$\sum_d (d-1)q_d(z) = \frac{\sum d(d-1)n_d z^d}{\sum_d dn_d z^d} = \frac{z^2 f''(z)}{zf'(z)} = \frac{zf''(z)}{f'(z)}.$$

The variance bound is immediate from the fact that by time t we have 'used up' $O(t)$ vertices of each degree, changing the conditional variance by at most $O(t/n)$. □

Let us now define the idealized trajectory that we have in mind. Recall that $f(z)=f_n(z)$ is defined by (3.3), and $z=z(\tau)$ by

$$f(z)+\tau = 1 \tag{3.9}$$

for $0\leqslant \tau \leqslant 1$, so

$$\frac{\mathrm{d}z}{\mathrm{d}\tau} = -\frac{1}{f'(z)}. \tag{3.10}$$

We shall think of τ as a rescaled time parameter, taking $\tau = t/n$, but will write our trajectory as a function of z. In the light of our assumption (2.6), we need only consider $\tau \leqslant c_0/2$. It is easy to check that this implies $z\geqslant 1/2$. In fact, we can always assume that $\tau = o(1)$ and so $z = 1-o(1)$.

https://doi.org/10.1017/S0963548311000666 Published online by Cambridge University Press

Let $\mu_1 = f'(1) = n^{-1}\sum dn_d$ be the average degree in our graph $G = G_{\mathbf{d}}$, and define functions $x(\tau)$ and $u(\tau)$ by

$$x = zf'(z) - \frac{f'(z)^2}{\mu_1} \tag{3.11}$$

and

$$u = zf'(z). \tag{3.12}$$

Using elementary calculus, it is straightforward to check that these functions satisfy $x = 0$ when $\tau = 0$, *i.e.*, when $z = 1$, and

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} = \frac{\mathrm{d}x}{\mathrm{d}z}\frac{\mathrm{d}z}{\mathrm{d}\tau} = -1 + \frac{zf''(z)}{f'(z)}\left(1 - \frac{2x}{u}\right). \tag{3.13}$$

Recall that f may depend on n. Since at least $c_0 n$ vertices have degree between 1 and $d_{\max}$, for $z \geqslant 1/2$ (which is the only range we consider), we have $f'(z) \geqslant c_0 2^{-d_{\max}}$, so $f'(z)$ is bounded below away from zero. Also, any given derivative of $f(z)$ is bounded by a constant depending only on $d_{\max}$. Using (3.10) it follows easily that the derivative of any fixed order of x with respect to τ is bounded uniformly in (large enough) n.

One can check that

$$-(f')^3\mu_1\frac{\mathrm{d}^2}{\mathrm{d}\tau^2}x = 2(f')^2f''' + \mu_1 z(f'')^2 - \mu_1 f'f'' - \mu_1 zf'f'''.$$

Substituting $z = 1$ and noting that $\mu_i = f^{(i)}(z)|_{z=1}$, we have

$$-\mu_1^4\frac{\mathrm{d}^2}{\mathrm{d}\tau^2}x\bigg|_{z=1} = 2\mu_1^2\mu_3 + \mu_1\mu_2^2 - \mu_1^2\mu_2 - \mu_1^2\mu_3 = \mu_1(\mu_1\mu_3 + \mu_2^2 - \mu_1\mu_2).$$

By assumption $\lambda = \mu_2/\mu_1 \sim 1$, and as noted in Section 1, $\mu_3 = \Theta(1)$ (see (1.7)). Hence

$$\frac{\mathrm{d}^2}{\mathrm{d}\tau^2}x\bigg|_{\tau=0} \sim -\frac{\mu_3}{\mu_1^2} = -\Theta(1). \tag{3.14}$$

Let us collect together some basic properties of the trajectory x. We write $\dot{x}$ for the derivative of x with respect to τ.

Lemma 3.4. *Suppose that* $\mathbf{d}$ *satisfies the assumptions* (1.1) *and* (1.2), *and that* $\lambda = \lambda(\mathbf{d}) \to 1$. *Then*

(i) $x(0) = 0$,
(ii) $\dot{x}(0) = \lambda - 1$ *and*
(iii) $\ddot{x}(\tau) = -\frac{\mu_3}{\mu_1^2} + o(1)$, *uniformly in* $\tau = o(1)$.

Suppose in addition that $\varepsilon = \lambda - 1 > 0$ *and that* $\varepsilon^3 n \to \infty$, *and let* $\rho = \rho_n$ *be defined by* (1.9). *Then also*

(iv) $x(\rho) = 0$,
(v) $\rho \sim \frac{2\mu_1^2}{\mu_3}\varepsilon$,
(vi) $\dot{x}(\rho) \sim -\varepsilon$ *and*
(vii) $x(\tau) > \varepsilon\tau/2$ *whenever* $0 \leqslant \tau = o(\varepsilon)$.

https://doi.org/10.1017/S0963548311000666 Published online by Cambridge University Press

Proof. We have noted (i) already. Substituting $z = 1$ (corresponding to $\tau = 0$) into (3.13) gives (ii). Further, (iii) follows from (3.14) and the fact (noted above) that the third derivative of x is uniformly bounded over $\tau = o(1)$.

Turning specifically to the supercritical case, (iv) follows easily from (3.11) and (1.9). Indeed, recalling that f is the generating function of our degree distribution, (1.8) says exactly that z is the smallest positive solution to $z = f'(z)/\mu$. From (3.11) we have $x = 0$ at this value of z. Now (1.9) says that $\rho = 1 - f(z)$ which, by our change of variable formula (3.9), is exactly the corresponding value of τ.

Finally, (v)–(vii) follow from (i)–(iv) and Taylor's theorem. □

For $1 \leqslant t \leqslant c_0 n/2$ set $x_t = nx(t/n)$ and $u_t = nu(t/n)$; our aim is to show that X_t will be close to x_t and U_t close to u_t; the functions x and u are the corresponding 'scaling limits'. Since, as a function of τ, x has uniformly bounded second derivative, we have

$$x_{t+1} - x_t = \left.\frac{\mathrm{d}x}{\mathrm{d}\tau}\right|_{\tau=t/n} + O(1/n).$$

Hence, from (3.13), writing z for $z(t/n)$, and defining $w_t = \frac{zf''(z)}{f'(z)}$, we have

$$x_{t+1} - x_t = -1 + w_t\left(1 - \frac{2x_t}{u_t}\right) + O(1/n), \tag{3.15}$$

which is strongly reminiscent of (2.10). Our aim is to show that (X_t) will w.h.p. remain close to (x_t), and use this, and the asymptotic normality of the deviations, to prove Theorem 1.1.

For the rest of the section we fix some $t_{\max} = t_{\max}(n) = o(n)$ (see the next two sections for specific values). In what follows, we shall only consider values of t up to $t_{\max}$.

Set

$$D_{t+1} = \mathbb{E}(X_{t+1} - X_t \mid \mathcal{F}_t) \quad \text{and} \quad \Delta_{t+1} = X_{t+1} - X_t - D_{t+1},$$

noting that D_{t+1} is random. Corollary 3.3 shows that for any deterministic $t = t(n)$ with $t \leqslant t_{\max}$, there is some random K_t satisfying

$$K_t = O_{\mathrm{p}}(\sqrt{t}/n) \tag{3.16}$$

such that $|U_i - u_i|/n \leqslant K_t$ and $|\mathbb{E}(\eta_{i+1} - 1 \mid \mathcal{F}_i) - w_i| \leqslant K_t$ for all $i \leqslant t$. Note that (taking the smallest value of K_t for which these bounds hold), we may assume that K_t is increasing in t. Recalling that $U_i \geqslant c_0 n/2$ in the range we consider, and noting that η_{i+1} is bounded by $d_{\max}$, it follows using (2.10) and (3.15) that for $i \leqslant t \leqslant t_{\max}$ we have

$$|D_{i+1} - (x_{i+1} - x_i)| \leqslant c\left(\frac{|X_i - x_i|}{n} + K_t + \frac{C_i}{n}\right), \tag{3.17}$$

for some constant c that depends only on $d_{\max}$ and c_0. Since $X_0 = 0$, we may write X_t as

$$X_t = \sum_{i \leqslant t}(D_i + \Delta_i) = \sum_{i \leqslant t} D_i + S_t,$$

https://doi.org/10.1017/S0963548311000666 Published online by Cambridge University Press

where

$$S_t = \sum_{i \leqslant t} \Delta_i.$$

Let

$$\widetilde{X}_t = x_t + S_t, \tag{3.18}$$

which we shall think of as a (rather precise) random approximation to X_t, and define the 'error term' E_t by

$$E_t = X_t - \widetilde{X}_t = \sum_{i \leqslant t} D_i - x_t. \tag{3.19}$$

Recall that x_t is deterministic. The key point is that the distribution of S_t is easy to control, since (S_t) is a martingale with bounded differences. Let

$$M_t = \max_{0 \leqslant i \leqslant t} |S_i|.$$

Lemma 3.5. *For any (deterministic) $t = t(n)$ and any $m \geqslant 0$ we have*

$$\mathbb{P}(M_t \geqslant m) \leqslant d_{\max}^2 t/m^2. \tag{3.20}$$

In particular, $M_t = O_p(\sqrt{t})$. Furthermore, for any $\alpha = \alpha(n) > 0$ and any $\omega = \omega(n) \to \infty$, the event $\{M_t \leqslant \alpha t/4$ for all $\omega/\alpha^2 \leqslant t \leqslant t_{\max}\}$ holds w.h.p.

Proof. Since the differences Δ_i are bounded by $d_{\max}$, their (conditional) variances are at most $d_{\max}^2$, so $\mathrm{Var}(S_t) \leqslant d_{\max}^2 t$. Applying Doob's maximal inequality gives (3.20). That $M_t = O_p(\sqrt{t})$ follows immediately.

For the last part, let $t_i = 2^i \omega/\alpha^2$, and let $\mathcal{E}_i$ be the event that $M_{t_{i+1}} \geqslant \alpha t_i/4$. Then (3.20) gives

$$\mathbb{P}(\mathcal{E}_i) \leqslant \frac{d_{\max}^2 t_{i+1}}{\alpha^2 t_i^2/16} \leqslant \frac{32 d_{\max}^2}{2^i \omega}.$$

It follows that $\sum \mathbb{P}(\mathcal{E}_i) = o(1)$, so w.h.p. no $\mathcal{E}_i$ holds, giving the result. □

Lemma 3.6. *Let $t = t(n)$ satisfy $t \to \infty$ and $t = o(n)$. Then S_t is asymptotically normal with mean 0 and variance $v_0 t$, where v_0 is given by (1.6).*

Proof. Recall that (S_t) is a martingale with $S_0 = 0$, and the differences Δ_i are bounded. Moreover, by Corollary 3.3, $\mathrm{Var}(\Delta_{i+1} \mid \mathcal{F}_i) \sim v_0$ when $i = o(n)$. The result thus follows from a standard martingale central limit theorem such as Brown [7, Theorem 2]. □

We now turn to the error terms E_t. Using the second expression for E_t in (3.19), summing (3.17) over $1 \leqslant i < t$, and noting that $|X_i - x_i| = |S_i + E_i|$, for $t \leqslant t_{\max}$ we see

https://doi.org/10.1017/S0963548311000666 Published online by Cambridge University Press

that

$$|E_t| \leqslant c\left(\frac{t}{n}\max_{i<t}|X_i - x_i| + tK_t + \frac{tC_{t-1}}{n}\right)$$
$$\leqslant \frac{ct}{n}\max_{i<t}|E_i| + c\left(\frac{tM_t}{n} + tK_t + \frac{tC_{t-1}}{n}\right).$$

Since $ct/n \leqslant ct_{\max}/n = o(1)$ is less than $1/2$ if n is large, then for n large enough (which we assume from now on), it follows by induction on i that

$$|E_i| \leqslant 2c\left(\frac{tM_t}{n} + tK_t + \frac{tC_{t-1}}{n}\right) \tag{3.21}$$

for $0 \leqslant i \leqslant t$.

Set

$$M_t^* = 2c\left(\frac{tM_t}{n} + tK_t\right). \tag{3.22}$$

Note that M_t^* is increasing in t. Also, for $0 \leqslant t \leqslant t_{\max}$, (3.21) (applied with $i = t$) gives

$$|E_t| \leqslant M_t^* + 2ctC_{t-1}/n. \tag{3.23}$$

For any (deterministic) $t = t(n)$, it follows from (3.16) and Lemma 3.5 that

$$M_t^* = O_{\mathrm{p}}(t^{3/2}/n). \tag{3.24}$$

4. The critical case

Using the bounds from the previous section, it is very easy to prove Theorem 1.3. The hardest part is establishing that the description of the limit actually makes sense; this follows from the results of Aldous [1] by simple rescaling. Since all probabilistic technicalities are same as in [1], we shall not mention them. Indeed, we take a combinatorial point of view in the estimates that follow.

Proof of Theorem 1.3. Recall that by assumption $n^{1/3}(\lambda - 1) \to \alpha_1 \in \mathbb{R}$, while $\mu_3/\mu_1 \to \alpha_0$ and $\mu_3/\mu_1^2 \to \alpha_2$ with $\alpha_0, \alpha_2 > 0$. For the moment, let ω be a large constant. A little later we shall allow ω to tend to infinity slowly.

Define a random function $S(s)$ on $[0,\omega]$ by setting $S(t/n^{2/3}) = n^{-1/3}S_t$ for $0 \leqslant t \leqslant \omega n^{2/3}$, and interpolating linearly between these values. Recall that S_t is a martingale with bounded differences and that for $t \leqslant \omega n^{2/3} = o(n)$, the conditional variances of the differences are $v_0 + o(1) \sim \mu_3/\mu_1 \sim \alpha_0$ (see (1.7)). It follows easily that S converges to $\alpha_0^{1/2}W$, where W is a standard Brownian motion on $[0,\omega]$, in the sense that these random functions can be coupled so that $\sup_{s\in[0,\omega]}|S(s) - \alpha_0^{1/2}W(s)|$ converges to 0 in probability. (To see this, one can apply a functional martingale central limit theorem, or simply subdivide into suitable short intervals and use a standard martingale CLT.)

Recalling that $\widetilde{X}_t = x_t + S_t = nx(t/n) + S_t$, define a random function $\widetilde{X}(s)$ on $[0,\omega]$ by setting

$$\widetilde{X}(s) = n^{-1/3}\widetilde{X}_t = n^{2/3}x(sn^{-1/3}) + S(s)$$

https://doi.org/10.1017/S0963548311000666 Published online by Cambridge University Press

whenever $s = t/n^{2/3}$ for integer $t \leqslant \omega n^{2/3}$, and again interpolating linearly. Note that $\widetilde{X}$ is given by adding a deterministic function to S. From Lemma 3.4 and Taylor's theorem, we have

$$n^{2/3}x(sn^{-1/3}) = 0 + n^{2/3}(\lambda - 1)sn^{-1/3} - n^{2/3}\left(\frac{\mu_3}{\mu_1^2} + o(1)\right)\frac{s^2n^{-2/3}}{2}$$
$$= \alpha_1 s - \frac{\alpha_2}{2}s^2 + o(1),$$

uniformly in $s \in [0, \omega]$. It follows that $\widetilde{X}$ converges to the inhomogeneous Brownian motion $W_{\alpha_0,\alpha_1,\alpha_2}$ defined in (1.18).

Setting $t_{\max} = \omega n^{2/3}$, by (3.24) the quantity $M^*_{t_{\max}}$ defined in (3.22) satisfies

$$M^*_{t_{\max}} = O_{\mathrm{p}}(t_{\max}^{3/2}/n) = O_{\mathrm{p}}(1) = o_{\mathrm{p}}(n^{1/3}).$$

Pick some (deterministic) $k = k(n)$ with $k/n^{1/3} \to \infty$. Suppose that we explore more than k components in the first $t_{\max}$ steps. When we finish exploring the kth component we have $X_t = -2k$ and $C_t = k$. Since $t_{\max} = o(n)$, the bound (3.23) thus gives $|E_t| \leqslant O_{\mathrm{p}}(1) + o(k)$. In particular, w.h.p. $|E_t| \leqslant k$. It then follows that $|\widetilde{X}_t| \geqslant |X_t| - |E_t| \geqslant k$. From the convergence of $(\widetilde{X}_t)$ to $W_{\alpha_0,\alpha_1,\alpha_2}$ we have $\sup_{t \leqslant t_{\max}} |\widetilde{X}_t| = O_{\mathrm{p}}(n^{1/3})$. It follows that $C_{t_{\max}} = O_{\mathrm{p}}(n^{1/3})$. Using (3.23) again, this gives $\sup_{t \leqslant t_{\max}} |E_t| = O_{\mathrm{p}}(1)$. In other words, the 'idealized random trajectory' $\widetilde{X}_t$ is an extremely close approximation to X_t up to time $t_{\max}$. Using (X_t) to define a function X on $[0, \omega]$ as for S and $\widetilde{X}$ above, it follows that X also converges to $W_{\alpha_0,\alpha_1,\alpha_2}$.

Finally, recalling that $X_t = A_t - 2C_t$, and that X_t first hits $-2k$ when we finish exploring the kth component, *i.e.*, just before C_t increases to $k+1$, it is easy to check that for $t > 0$ we have $A_t = X_t - \min_{i<t} X_i + O(1)$. Defining a final function A on $[0, \omega]$ using the A_t, we see that A converges to the function B defined in (1.19). So far ω was fixed, but convergence for all fixed ω implies convergence for $\omega \to \infty$ sufficiently slowly. Now the sizes of the components explored during the first $\omega n^{2/3}$ steps are simply $n^{2/3}$ times the excursion lengths of A, which converge to the excursion lengths of B. (This follows from basic properties of B.)

For the component sizes, it remains only to show that when $\omega \to \infty$, for any $\delta > 0$, w.h.p. there are no components of size at least $\delta n^{2/3}$ in the rest of the graph. This follows from Theorem 1.2 by an argument similar to that at the end of Section 5.

Finally, we also claimed convergence for the nullities, or numbers of back-edges, to appropriate Poisson parameters. For $t = o(n)$, which implies $U_t \sim \mu_1 n$ and $\mathbb{E}(\eta_{t+1} - 1 \mid \mathcal{F}_t) \sim \mathbb{E}(\eta - 1) = \lambda \sim 1$, from (2.9) we have

$$\mathbb{E}(\theta_{t+1} \mid \mathcal{F}_t) = (1 + o(1))\frac{A_t}{\mu_1 n} + O(1/n).$$

In terms of the rescaled function A on $[0, \omega]$, this corresponds to formation of back-edges at rate $A(s)/\mu_1$, and joint convergence of the component sizes and back-edge counts to the excursion lengths and mark counts claimed in the theorem follows easily as in [1]. □

5. The supercritical case

We follow the argument of Bollobás and the author in [6] closely, and attempt to use the same notation where possible. Throughout this section we fix a function $\omega = \omega(n)$ tending to infinity slowly, in particular with $\omega^6 = o(\varepsilon^3 n)$ and $\omega^2 = o(1/\varepsilon)$. As in [6], set

$$\sigma_0 = \sqrt{\varepsilon n}$$

and

$$t_0 = \omega\sigma_0/\varepsilon,$$

ignoring, as usual, the irrelevant rounding to integers. Note for later that $t_0 = o(\varepsilon n)$, and that $t_0 \geqslant \omega/\varepsilon^2$ if n is large. We shall apply the results of Section 3 with, say,

$$t_{\max} = 4\frac{\mu_1^2}{\mu_3}\varepsilon n \sim 2\rho n = \Theta(\varepsilon n); \tag{5.1}$$

see Lemma 3.4(v).

Lemma 5.1. *Let Z denote the number of components completely explored by time t_0, and let $T_0 = \inf\{t : X_t = -2Z\}$ be the time at which we finish exploring the last such component. Then $Z \leqslant \omega/\varepsilon \leqslant \sigma_0/\omega$ and $T_0 \leqslant \omega/\varepsilon^2 \leqslant \sigma_0/(\varepsilon\omega)$ hold w.h.p.*

Proof. Set $k = \omega/\varepsilon$ and $t_0' = \omega/\varepsilon^2$. It is easy to check that $k \leqslant \sigma_0/\omega$ and $t_0' \leqslant \sigma_0/(\varepsilon\omega)$, so it suffices to prove that $Z \leqslant k$ and $T_0 \leqslant t_0'$ hold w.h.p.

Note first that $t_0 = \omega(n/\varepsilon)^{1/2}$, so $t_0^{3/2}/n = \omega^{3/2}(\varepsilon^3 n)^{-1/4} = o(\omega^{3/2}) = o(k)$. Let $\mathcal{A}_1$ denote the event that $M_{t_0}^* < k/8$, so $\mathcal{A}_1$ holds w.h.p. by (3.24). Let $\mathcal{A}_2$ be the event that $M_{t_0'} < k/8$, and $\mathcal{A}_3$ the event that $M_t \leqslant \varepsilon t/4$ for all $t_0' \leqslant t \leqslant t_0$. Then, noting that $k/\sqrt{t_0'} = \sqrt{\omega} \to \infty$, the events $\mathcal{A}_2$ and $\mathcal{A}_3$ hold w.h.p. by Lemma 3.5.

From Lemma 3.4 we have $x_t = nx(t/n) \geqslant \varepsilon t/2 \geqslant 0$ for all $t \leqslant t_0 = o(\varepsilon n)$. Suppose that $\mathcal{A} = \mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{A}_3$ holds. Then we have $x_t + S_t \geqslant x_t - M_t \geqslant \varepsilon t/4 - k/8$ for all $t \leqslant t_0$, using $\mathcal{A}_2$ for $t \leqslant t_0'$, and $\mathcal{A}_3$ for $t > t_0'$.

For $t \leqslant t_0$, from (3.23) and the fact that $t_0 = o(n)$ we have

$$|E_t| \leqslant M_{t_0}^* + 2ctC_t/n \leqslant k/8 + C_t$$

if $\mathcal{A}_1$ holds and n is large.

At time T_0 we have $X_{T_0} = -2Z$ (see (2.11)) and $C_{T_0} = Z$. Suppose that $\mathcal{A}$ holds. Then $\widetilde{X}_{T_0} = X_{T_0} - E_t \leqslant -Z + k/8$. Since $\widetilde{X}_{T_0} = x_{T_0} + S_{T_0} \geqslant \varepsilon T_0/4 - k/8$, it follows that

$$-Z + k/8 \geqslant \widetilde{X}_{T_0} \geqslant \varepsilon T_0/4 - k/8.$$

Rearranging gives $Z \leqslant k/4 - \varepsilon T_0/4 \leqslant k/4$ and hence, since $Z \geqslant 0$, $T_0 \leqslant k/\varepsilon$, completing the proof. □

Let $T_1 = \inf\{t : X_t = -2(Z+1)\}$, so T_1 is the first time after t_0 at which we finish exploring a component. In particular, there is a component with $T_1 - T_0$ vertices.

https://doi.org/10.1017/S0963548311000666 Published online by Cambridge University Press

Lemma 5.2. *$T_1 - T_0$ is asymptotically normally distributed with mean ρn and variance $2\mu_1\varepsilon^{-1}n$, where ρ is defined by* (1.9).

Proof. For $t \leqslant T_1$ we have $C_t \leqslant Z + 1$, which is w.h.p. at most $\omega/\varepsilon + 1$ by Lemma 5.1. For $t \leqslant \min\{T_1, t_{\max}\}$ it follows that $tC_t/n = O(\omega)$. Since $t_{\max}^{3/2}/n = \Theta(\varepsilon^{3/2}n^{1/2})$, the bounds (3.23) and (3.24) imply that w.h.p.

$$\max_{t \leqslant \min\{T_1, t_{\max}\}} |E_t| \leqslant \omega\sqrt{\varepsilon^3 n} \leqslant \sigma_0/\sqrt{\omega}, \tag{5.2}$$

say.

Ignoring the irrelevant rounding to integers, let $t_1 = \rho n$. Let $t_1^- = t_1 - t_0$ and $t_1^+ = t_1 + t_0$. Recalling (5.1), we have $t_1^+ \leqslant t_{\max} = O(\varepsilon n) = O(\sigma_0^2)$. Hence, by Lemma 3.5, $M_{t_1^+} = O_{\mathrm{p}}(\sigma_0)$. Since $X_t = x_t + S_t + E_t$ it follows that

$$\max_{t \leqslant \min\{T_1, t_1^+\}} |X_t - x_t| \leqslant \sqrt{\omega}\sigma_0 \tag{5.3}$$

holds w.h.p.

Let $a = -\dot{x}(\rho)$, so from Lemma 3.4,

$$a = -\dot{x}(\rho) \sim \varepsilon. \tag{5.4}$$

Since $x(\rho) = 0$ and $\ddot{x}$ is uniformly bounded, recalling that $t_0 = o(\varepsilon n)$ it follows easily that $x_{t_1^-}$ and $x_{t_1^+}$ are both of order $\varepsilon t_0 = \omega\sigma_0$. To be concrete, if n is large enough, then we certainly have

$$x_{t_1^-} \geqslant 10\sqrt{\omega}\sigma_0 \quad \text{and} \quad x_{t_1^+} \leqslant -10\sqrt{\omega}\sigma_0,$$

say. By Lemma 3.4 we have $x_{t_0} \geqslant \varepsilon t_0/2 \geqslant 10\sqrt{\omega}\sigma_0$. Also x_t increases near (within $o(\varepsilon n)$ of) $t = 0$, decreases near $t = t_1$, and is of order $\Theta(\varepsilon^2 n)$ in between. It follows that $\inf_{t_0 \leqslant t \leqslant t_1^-} x_t \geqslant 10\sqrt{\omega}\sigma_0$. Let $\mathcal{B}$ denote the event described in (5.3). Then, whenever $\mathcal{B}$ holds, we have $X_t \geqslant 0$ for $t_0 \leqslant t \leqslant \min\{T_1, t_1^-\}$. Since $X_{T_1} \leqslant -2(Z+1) < 0$, and $T_1 > t_0$ by definition, this implies $T_1 > t_1^-$.

Recall from Lemma 5.1 that $Z \leqslant \sigma_0$ w.h.p. Suppose $Z \leqslant \sigma_0$, $\mathcal{B}$ holds, and $T_1 > t_1^+$. Then from $\mathcal{B}$ and the bound on $x_{t_1^+}$ we have $X_{t_1^+} \leqslant -9\sqrt{\omega}\sigma_0 < -2Z - 2$, contradicting $T_1 > t_1^+$. It follows that $T_1 \leqslant t_1^+$ holds w.h.p.

We claim that

$$\sup_{|t-t_1| \leqslant t_0} |\widetilde{X}_t - \widetilde{X}_{t_1} - a(t_1 - t)| = o_{\mathrm{p}}(\sigma_0). \tag{5.5}$$

From (3.18) we may write $\widetilde{X}_t - \widetilde{X}_{t_1}$ as

$$(x_t - x_{t_1}) + (S_t - S_{t_1}).$$

Recalling that $t_1 = \rho n$, $\dot{x}(\rho) = -a$ and that $\ddot{x}$ is uniformly bounded, the difference between the first term and $a(t_1 - t)$ is $O(|t - t_1|^2/n) = O(t_0^2/n) = o(\sigma_0)$. Since $(S_t - S_{t_1^-})_{t=t_1^-}^{t_1^+}$ is a martingale with final variance $O(t_0) = o(\sigma_0^2)$, Doob's maximal inequality gives $\sup_{|t-t_1| \leqslant t_0} |S_t - S_{t_1}| = o_{\mathrm{p}}(\sigma_0)$, and (5.5) follows.

Recall from Lemma 5.1 that Z, the number of components explored by time t_0, satisfies $Z = o_{\mathrm{p}}(\sigma_0)$. We have shown above that w.h.p. $T_1 = \inf\{t : X_t = -2(Z+1)\}$ lies between t_1^- and t_1^+. From (5.2), X_t is within $o_{\mathrm{p}}(\sigma_0)$ of $\widetilde{X}_t$ at least until T_1. It follows that at time T_1, we have $\widetilde{X}_t = o_{\mathrm{p}}(\sigma_0)$. Since $a = \Theta(\varepsilon)$, (5.5) thus gives

$$T_1 = t_1 + \widetilde{X}_{t_1}/a + o_{\mathrm{p}}(\sigma_0/\varepsilon). \tag{5.6}$$

From Lemma 3.6, (3.18) and the fact that $x(\rho) = 0$, we have that $\widetilde{X}_{t_1}$ is asymptotically normal with mean 0 and variance $v_0\rho n$. Hence $\widetilde{X}_{t_1}/a$ is asymptotically normal with mean 0 and variance

$$v_0\rho n/a^2 \sim 2\mu_1\varepsilon^{-1}n,$$

using (1.7), (1.11) and (5.4). Since this variance is of order $\varepsilon^{-1}n = \varepsilon^{-2}\sigma_0^2$, the $o_{\mathrm{p}}(\sigma_0/\varepsilon)$ error term in (5.6) is irrelevant, and T_1 is asymptotically normal with mean $t_1 = \rho n$ and variance $2\mu_1\varepsilon^{-1}n$. Finally, from Lemma 5.1 we have $T_0 = o_{\mathrm{p}}(\sigma_0/\varepsilon)$. It follows that $T_1 - T_0$ is asymptotically normal with the parameters claimed in the theorem. □

We are now ready to complete the proof of Theorem 1.1.

Proof of Theorem 1.1. Let $\mathcal{C}$ denote the component explored from time T_0 to T_1. We have already shown that $\mathcal{C}$ has the size claimed; two tasks remain, namely to study the nullity of $\mathcal{C}$, and to show that there are no other 'large' components.

Recall from (2.9) that the conditional expected number of back-edges added at each step satisfies

$$\mathbb{E}(\theta_{t+1} \mid \mathcal{F}_t) = \mathbb{E}(\eta_{t+1} - 1 \mid \mathcal{F}_t)A_t/U_t + O(1/n). \tag{5.7}$$

For the nullity, we shall consider only $t \leqslant \min\{T_1, t_{\max}\}$, recalling that w.h.p. $T_1 \leqslant t_{\max} = O(\varepsilon n)$. In this range, we have $C_t \leqslant Z + 1 \leqslant 2\omega/\varepsilon$ w.h.p. by Lemma 5.1. Also $|E_t| \leqslant \omega\sqrt{\varepsilon^3 n}$ w.h.p. by (5.2). Since $\varepsilon \to 0$ and $\varepsilon^3 n \to \infty$, if $\omega \to \infty$ sufficiently slowly then both these bounds are $o(\sqrt{\varepsilon n})$. Recalling that $\widetilde{X}_t = X_t - E_t = A_t - 2C_t - E_t$, it follows that w.h.p.

$$|\widetilde{X}_t - A_t| = o(\sqrt{\varepsilon n}) \tag{5.8}$$

throughout our range.

For $t \leqslant t_{\max}$, Lemma 3.4 gives $x_t = O(\varepsilon^2 n)$, and it follows easily from the bounds above that $\widetilde{X}_t$ is w.h.p. $O(\varepsilon^2 n)$. Recalling that $f''(1)/f'(1) = \mu_2/\mu_1 \sim 1$, and defining z by $f(z) + t/n = 1$ as before, by Corollary 3.3 the maximum relative error (for $t \leqslant t_{\max}$) in approximating U_t by $nzf'(z)$ or $\mathbb{E}(\eta_{t+1} - 1 \mid \mathcal{F}_t)$ by $zf''(z)/f'(z)$ is $O_{\mathrm{p}}(\sqrt{\varepsilon n}/n) = o_{\mathrm{p}}(1/\sqrt{\varepsilon^3 n})$. Using (5.7) and (5.8), it follows that w.h.p.

$$D_{t+1}^* = \mathbb{E}(\theta_{t+1} \mid \mathcal{F}_t) = \frac{\widetilde{X}_t}{n}\frac{f''(z)}{f'(z)^2} + o(\sqrt{\varepsilon/n}) \tag{5.9}$$

for all $t \leqslant T_1$.

Recalling that $x_{t_1} = 0$, the bounds in the proof of Lemma 5.2 show that w.h.p. $|\widetilde{X}_t| \leqslant \omega\sigma_0$ for $t_1^- \leqslant t \leqslant T_1$, and that w.h.p. $T_1 \leqslant t_1^- + 2t_0$. It follows from (5.9) that w.h.p. no more than, say, $\omega^2\sigma_0 t_0/n = \omega^3\sigma_0^2/(\varepsilon n) = \omega^3 = o(\sqrt{\varepsilon^3 n})$ back-edges are added between time t_1^-

https://doi.org/10.1017/S0963548311000666 Published online by Cambridge University Press

and time T_1. A similar bound holds for steps up to t_0 and for steps between T_1 and $t_1 = \rho n$. Let $Y = Y_{\rho n}$ be the total number of back-edges found up to time ρn. Using (2.12), it follows that

$$|Y - n(\mathcal{C})| = o_{\mathrm{p}}(\sqrt{\varepsilon^3 n}). \tag{5.10}$$

Let us write θ_{t+1} as $D^*_{t+1} + \Delta^*_{t+1}$, so by definition $\mathbb{E}(\Delta^*_{t+1} \mid \mathcal{F}_t) = 0$. Then $Y = D^* + S^*$, where $D^* = \sum_{t \leqslant \rho n} D^*_t$ and $S^* = \sum_{t \leqslant \rho n} \Delta^*_t$. Note that D^* is random. Recalling that $\widetilde{X}_t = x_t + S_t$ and that $\mathbb{E}S_t = 0$, it follows from (5.9) and the definition of $x_t = x(t/n)$ (see (3.11)) that

$$\mathbb{E}D^* = \sum_{t=0}^{\rho n - 1} \left(z f'(z) - \frac{f'(z)^2}{\mu_1} \right) \frac{f''(z)}{f'(z)^2} + o_{\mathrm{p}}(\sqrt{\varepsilon^3 n}),$$

where, as usual, $z = z(t)$ is defined by $f(z) + t/n = 1$.

It is not hard to see that the sum above is sufficiently well approximated by the corresponding integral. Recalling that with $\tau = t/n$ we have $\frac{\mathrm{d}\tau}{\mathrm{d}z} = -f'(z)$, it follows that

$$\mathbb{E}D^* = n \int_{z=z_1}^{1} (z - f'(z)/\mu_1) f''(z) \mathrm{d}z + o_{\mathrm{p}}(\sqrt{\varepsilon^3 n}),$$

where z_1 corresponds to $\tau = \rho$. In other words, z_1 is the value of z defined in (1.8), which, as noted in the proof of Lemma 3.4, satisfies $z_1 = f'(z_1)/\mu_1$. The integrand above is the derivative of $z f'(z) - f(z) - f'(z)^2/(2\mu_1)$. Hence

$$\mathbb{E}D^* = n \left(f(z_1) - \frac{\mu_1 z_1^2}{2} - 1 + \frac{\mu_1}{2} \right) + o_{\mathrm{p}}(\sqrt{\varepsilon^3 n}) = \rho^* n + o_{\mathrm{p}}(\sqrt{\varepsilon^3 n}), \tag{5.11}$$

recalling (1.10). Since $\mathbb{E}\Delta^* = 0$, this gives $\mathbb{E}Y = \rho^* n + o_{\mathrm{p}}(\sqrt{\varepsilon^3 n})$.

From Lemma 3.4(v) and (3.10) it follows easily that $\delta = 1 - z_1 \sim \rho/\mu_1 \sim 2\varepsilon\mu_1/\mu_3$. We may write ρ^* as $f(1-\delta) - 1 + \mu_1\delta - \mu_1\delta^2/2$. Expanding $f(z)$ about $z = 1$, using $f(1) = 1$, $f'(1) = \mu_1$, $f''(1) = \mu_2 = (1+\varepsilon)\mu_1$, $f'''(1) = \mu_3$ and $f^{(4)} = O(1)$, a little calculation establishes that

$$\rho^* \sim \frac{2\mu_1^3}{3\mu_3^2} \varepsilon^3. \tag{5.12}$$

We now turn to the variance and covariance estimates. Here we can be much less careful, as a $o(1)$ relative error does not affect our conclusions.

Recall that D^* is random. From (5.9), (5.11) and the fact that $\widetilde{X}_t = \mathbb{E}\widetilde{X}_t + S_t$ we can write D^* as $\rho^* n + D' + o_{\mathrm{p}}(\sqrt{\varepsilon^3 n})$, where

$$D' = \sum_{t=0}^{\rho n - 1} \frac{S_t}{n} \frac{f''(z)}{f'(z)^2}.$$

Thus

$$Y = D^* + S^* = \rho^* n + D' + S^* + o_{\mathrm{p}}(\sqrt{\varepsilon^3 n}).$$

https://doi.org/10.1017/S0963548311000666 Published online by Cambridge University Press

Throughout the relevant range, $f''(z)/f'(z)^2 \sim f''(1)/f'(1)^2 = \mu_2/\mu_1^2 \sim 1/\mu_1$. Since $S_t = \sum_{i\leqslant t} \Delta_i$, it follows that

$$D' = \sum_{i=0}^{\rho n-1} a_i \Delta_i,$$

for some constants $a_i = a_i(n)$ satisfying

$$a_i \sim (\rho n - i)/(\mu_1 n).$$

Since the Δ_i are martingale differences with variances $v_0 + o(1)$, it follows that

$$\mathrm{Var}(D') \sim \sum_{i=0}^{\rho n-1} v_0 a_i^2 \sim \frac{v_0}{\mu_1^2} \sum_{i=0}^{\rho n-1} \frac{(\rho n - i)^2}{n^2} \sim \frac{v_0 \rho^3}{3\mu_1^2} n \sim \frac{8\mu_1^3}{3\mu_3^2} \varepsilon^3 n,$$

using (1.7) and (1.11). Hence $D^* = \rho^* n + O_{\mathrm{p}}(\sqrt{\varepsilon^3 n}) = (1 + o_{\mathrm{p}}(1))\rho^* n$.

Recalling that $\theta_{t+1} \geqslant 2$ is much less likely than $\theta_{t+1} = 1$, we have

$$\mathrm{Var}(\Delta_{t+1}^* \mid \mathcal{F}_t) = \mathrm{Var}(\theta_{t+1} \mid \mathcal{F}_t) \sim \mathbb{E}(\theta_{t+1} \mid \mathcal{F}_t).$$

Summing, it follows that

$$s^2 = \sum_{t=0}^{\rho n-1} \mathrm{Var}(\Delta_{t+1}^* \mid \mathcal{F}_t) = (1 + o_{\mathrm{p}}(1))D^* = (1 + o_{\mathrm{p}}(1))\rho^* n.$$

In particular $s^2/(\rho^* n)$ converges in probability to 1.

Recall that given $\mathcal{F}_t$, Δ_{t+1} and Δ_{t+1}^* both have conditional expectation 0. Recalling (2.8), their conditional covariance, which is just that of $\eta_{t+1} - 1$ and θ_{t+1}, is asymptotically $\mathrm{Var}(\eta_{t+1} - 1 \mid \mathcal{F}_t)A_t/U_t \sim v_0 A_t/U_t = \Theta(\varepsilon^2)$. This is much smaller than the square root of the product of their variances. It follows easily that, after appropriate normalization, the joint distribution of $S = S_{\rho n} = \sum_{t\leqslant \rho n} \Delta_t$ and $N' = D' + S^* = \sum_{t\leqslant \rho n}(a_t \Delta_t + \Delta_t^*)$ is asymptotically multivariate normal, with

$$\mathrm{Var}(N') \sim \mathrm{Var}(D') + \mathrm{Var}(S^*) \sim \frac{10\mu_1^3}{3\mu_3^2} \varepsilon^3 n$$

and

$$\mathrm{Cov}(N', S) \sim \mathrm{Cov}(D', S) \sim \sum_{t=0}^{\rho n-1} a_t v_0 \sim \frac{\rho^2 v_0}{2\mu_1} n \sim \frac{2\mu_1^2}{\mu_3} \varepsilon^2 n.$$

Recalling that the nullity of $\mathcal{C}$ is $Y + o_{\mathrm{p}}(\sqrt{\varepsilon^3 n}) = \rho^* n + N' + o_{\mathrm{p}}(\sqrt{\varepsilon^3 n})$ and that $|\mathcal{C}| = \rho n + S/a + o_{\mathrm{p}}(\sqrt{\varepsilon n})$ where $a = -\dot{x}(\rho) \sim \varepsilon$, it follows that $|\mathcal{C}|$ and $n(\mathcal{C})$ are jointly asymptotically normal with the means, variances and covariance claimed in Theorem 1.1.

It remains only to prove that all components other than $\mathcal{C}$ have size bounded by

$$O_{\mathrm{p}}\big(\varepsilon^{-2} \log(\varepsilon^3 n)\big) = o(\varepsilon n)$$

as in (1.12). Lemma 5.1 shows that $T_0 = O_{\mathrm{p}}(\varepsilon^{-2})$, so w.h.p. by time T_1 we have found no second component larger than this.

https://doi.org/10.1017/S0963548311000666 Published online by Cambridge University Press

Let us stop the exploration at time T_1. Then the unexplored part of the graph is simply the configuration multigraph on the degree sequence $\mathbf{d}'$ given by the vertices not yet reached. Note that $\lambda(\mathbf{d}')$ is exactly the expected value, given the history, of the degree η_{T_1+1} of the vertex about to be chosen. Since we have explored $O(\varepsilon n) = o(n)$ vertices, $\mathbf{d}'$ satisfies the assumption (1.2) (with a slightly reduced c_0), and it is still bounded. Since $\sqrt{\varepsilon n}/n = o(\varepsilon)$, by Corollary 3.3 $\mathbb{E}(\eta_{T_1+1} - 1 \mid T_1, \mathcal{F}_{T_1}) = \dot{x}(\rho) + o_{\mathrm{p}}(\varepsilon)$, so we find that $\lambda(\mathbf{d}') = 1 - a + o_{\mathrm{p}}(\varepsilon)$ with a as in (5.4). Since $a = \Theta(\varepsilon)$, Theorem 1.2 thus tells us that the largest component remaining has size $O_{\mathrm{p}}(\varepsilon^{-2}\log(\varepsilon^3 n))$, as required. □

Remark. Considering a random walk with independent increments with distribution $\eta - 2$, one would expect that the probability that our random walk (X_t) 'takes off' without hitting -2 near the start is roughly the expected degree of the initial vertex times $2\varepsilon/v_0$. (To see this, simply solve the recurrence relation for the probability of hitting -2 as a function of the initial value.) The expected degree of the initial vertex (which is chosen with probability proportional to degree) is roughly 2, giving a 'take-off' probability of roughly $4\varepsilon/v_0 \sim 4\varepsilon\mu_1/\mu_3$; it is not hard to check that this is asymptotically correct in the actual process (X_t).

One should expect this probability to be simply related to (or at first sight equal to) ρ; here the difference is that ρ is the probability that a *uniformly* chosen random vertex is in the giant component. It is not hard to check that the giant component is 'tree-like', and in particular has average degree $2 + o(1)$, so the probability that a vertex chosen with probability proportional to degree is in the giant component is around $2\rho/\mu_1$. From (1.11) this is asymptotically $4\varepsilon\mu_1/\mu_3$, as it should be.

This comment illustrates a strange feature of the trajectory tracking arguments here and in the papers of Nachmias and Peres [23] and Bollobás and Riordan [6]: there is a related viewpoint using branching processes which more easily gives the approximate size of the giant component, essentially by considering the probability that a vertex is in a large component, *i.e.*, that the random trajectory 'takes off'. One can prove quite accurate bounds by this method without worrying about when the trajectory will hit zero eventually; see Bollobas and Riordan [5]. However, for the distributional result, it seems easier to follow the whole trajectory.

6. A local limit theorem

One of the ingredients of the proof of Theorem 1.2 is a local limit theorem (Lemma 6.3 below) that may perhaps be known, but that we have not managed to find in the literature. This concerns a sequence (S_n) of sums of independent random variables. As usual in the combinatorial setting, each S_n involves different variables; we cannot make the more usual assumption in probability theory that each S_n is the sum of the first n terms of a single sequence. This makes little difference to the proofs, however.

We start from Esseen's inequality in the following form, also known as the Berry–Esseen theorem; see, for example, Petrov [24, Ch. V, Theorem 3]. We write $\phi(x)$ and $\Phi(x)$ for the density and distribution functions of the standard normal random variable.

https://doi.org/10.1017/S0963548311000666 Published online by Cambridge University Press

Theorem 6.1. *Let $Z_1, \dots, Z_t$ be independent random variables with $\rho = \sum_{i=1}^t \mathbb{E}(|X_i|^3)$ finite, and let $S = \sum_{i=1}^t Z_i$. Then*

$$\sup_x |\mathbb{P}(S \leqslant x) - \Phi((x-\mu)/\sigma)| \leqslant A\rho/\sigma^3,$$

where μ and σ^2 are the mean and variance of S, and A is an absolute constant. □

Given an integer-valued random variable Z, let $b_r(Z)$ denote the rth *Bernoulli part* of Z, defined by

$$b_r(Z) = 2 \sup_i \min\{\mathbb{P}(Z = i), \mathbb{P}(Z = i + r)\}. \tag{6.1}$$

It is easy to check that for any $p \leqslant b_1(Z)$ we can write Z in the form

$$Z = Z' + IB, \tag{6.2}$$

where $I \sim \mathrm{Bern}(p)$, $B \sim \mathrm{Bern}(1/2)$, and B is independent of the pair (Z', I). Here $\mathrm{Bern}(p)$ denotes the Bernoulli distribution assigning probability p to the value 1 and probability $1-p$ to the value 0. Intuitively, I is the indicator random variable of the event that we managed to include the Bernoulli variable B into Z. Similarly, for $p \leqslant b_r(Z)$ we can write Z in the form $Z' + rIB$ with the same assumptions on Z', I and B.

We should like a 'local limit theorem' giving, under mild conditions, an asymptotic formula for $\mathbb{P}(S = \lfloor \mu \rfloor)$, say, where S is a sum of independent random variables and $\mu = \mathbb{E}S$. As is well known (see, *e.g.*, [24, Ch. VII]), when the summands take integer values in a finite range, the only obstruction is their taking values in a non-trivial arithmetic progression, in which case the sum *cannot* take certain values. Results similar to the next lemma are stated in [24], but the conditions are different in important ways.

Lemma 6.2. *Let $k \geqslant 1$ be fixed. Suppose that for each n we have a sequence $(Z_{ni})_{i=1}^{t(n)}$ of independent random variables taking values in $\{-k, -k+1, \dots, k\}$. Let $S_n = \sum_{i=1}^{t(n)} Z_{ni}$, and let μ_n and σ_n^2 denote the mean and variance of S_n. Suppose that $\sigma_n^2 = \Theta(t(n))$, that $t(n) \to \infty$, and that*

$$b_{1,n} = \sum_{i=1}^{t(n)} b_1(Z_{ni}) \to \infty.$$

Then, for any sequence (x_n) of integers satisfying $x_n = \mu_n + O(\sigma_n)$ we have

$$\mathbb{P}(S_n = x_n) \sim p_0(n) = \frac{1}{\sigma_n \sqrt{2\pi}} \exp\left(-\frac{(x_n - \mu_n)^2}{2\sigma_n^2}\right). \tag{6.3}$$

Proof. The condition $|Z_{ni}| \leqslant k$ ensures that $\mathbb{E}(|Z_{ni}|^3) < k^3 = O(1)$, while by assumption $\sigma_n^2 = \Theta(t(n))$, so Theorem 6.1 gives

$$\sup_x |\mathbb{P}(S_n \leqslant x) - \Phi((x - \mu_n)/\sigma_n)| = O(t(n)^{-1/2}).$$

We shall not use exactly this bound, instead applying Theorem 6.1 to a slightly different sum of independent variables.

Let $\omega = \omega(n)$ be an integer chosen so that $\omega \to \infty$, but $\omega^6 \leqslant b_{1,n}$ and $\omega^{24} \leqslant t(n)$, say. Choose $p_i = p_{ni}$ so that $0 \leqslant p_i \leqslant b_1(Z_{ni})$ and $\sum p_i = \omega^6$.

Suppressing the dependence on n in the notation, let us write $Z_i = Z_{ni}$ in the form $Z_i = Z_i' + I_i B_i$ as in (6.2), where $I_i \sim \mathrm{Bern}(p_i)$, $B_i \sim \mathrm{Bern}(1/2)$, B_i is independent of (Z_i', I_i), and variables associated to different i are independent. Let $\mathbf{I} = (I_1, \ldots, I_t)$, and let $N = |\mathbf{I}| = \sum I_i$.

The idea is simply to condition on $\mathbf{I}$, and thus on N. Let $S^0 = S_n^0 = \sum Z_i'$, and $S^1 = S_n^1 = \sum_{i:I_i=1} B_i$, so $S_n = S^0 + S^1$. Then, given $\mathbf{I}$, S^0 and S^1 are independent. Furthermore, the conditional distribution of S^1 is binomial $\mathrm{Bi}(N, 1/2)$.

In the following argument we shall consider values of N in three separate ranges: $N \geqslant t^{1/3}$, $N \leqslant \omega^6/2$, and the 'typical' range $\omega^6/2 \leqslant N \leqslant t^{1/3}$.

Since N is a sum of independent indicators with $\mathbb{E}N = \omega^6 \leqslant t^{1/4}$, standard results (*e.g.*, the Chernoff bounds) imply that $\mathbb{P}(N \geqslant t^{1/3})$ is exponentially small in $t^{1/3}$, and hence, extremely crudely, that $\mathbb{P}(N \geqslant t^{1/3}) = o(t^{-1/2})$. Thus

$$\mathbb{P}\big(S_n = x_n \wedge N \geqslant t^{1/3}\big) = o(t^{-1/2}). \tag{6.4}$$

Note that $\mathbb{E}S^0 = \mu_n - \mathbb{E}N/2 = \mu_n - \omega^6/2 = \mu_n + o(\sqrt{t})$. Similarly, $\mathrm{Var}\, S^0 = \mathrm{Var}\, S_n + o(\sqrt{t}) = \sigma_n^2 + o(\sqrt{t})$. Let $\tilde{\mu}$ and $\tilde{\sigma}^2$ denote the conditional mean and variance of S^0 given $\mathbf{I}$. Given $\mathbf{I}$, the summands Z_i' in S^0 are independent, but their individual distributions depend on the I_i. Changing one I_i only affects the distribution of one summand, and all Z_i' are bounded by $\pm k$, so we see that $\tilde{\mu}$ and $\tilde{\sigma}^2$ change by $O(1)$ if one entry of $\mathbf{I}$ is changed. It follows easily that $\tilde{\mu} = \mathbb{E}S^0 + O(N + \mathbb{E}N)$ and $\tilde{\sigma}^2 = \mathrm{Var}\, S^0 + O(N + \mathbb{E}N)$. Hence, when $N \leqslant t^{1/3}$ we have

$$\tilde{\mu} = \mu_n + o(t^{1/2}) \quad \text{and} \quad \tilde{\sigma} \sim \sigma_n. \tag{6.5}$$

Given $\mathbf{I}$, S^0 is a sum of t independent random variables whose third moments are all bounded by k^3. By Theorem 6.1 it follows that when $N \leqslant t^{1/3}$ then

$$\mathbb{P}(S^0 \leqslant x \mid \mathbf{I}) = \Phi((x - \tilde{\mu})/\tilde{\sigma}) + O(t^{-1/2}) \tag{6.6}$$

uniformly in x. Considering consecutive (integer) values of x, it follows that

$$\sup_x \mathbb{P}(S^0 = x \mid \mathbf{I}) = O(t^{-1/2}) \tag{6.7}$$

whenever $N = |\mathbf{I}| \leqslant t^{1/3}$.

To handle the case $N \leqslant \omega^6/2$, recall that after conditioning on $\mathbf{I}$, the sums S^0 and S^1 are independent. Thus

$$\begin{aligned}\mathbb{P}(S_n = x_n \mid N \leqslant \omega^6/2) &\leqslant \sup_{\mathbf{I}_0 : |\mathbf{I}_0| \leqslant \omega^6/2} \sup_m \mathbb{P}(S^0 = x_n - m \mid \mathbf{I} = \mathbf{I}_0, S^1 = m) \\ &= O(t^{-1/2}),\end{aligned}$$

using (6.7) for the final bound. Since $\mathbb{P}(N \leqslant \omega^6/2) = o(1)$, this gives

$$\mathbb{P}\big(S_n = x_n \wedge N \leqslant \omega^6/2\big) = o(t^{-1/2}). \tag{6.8}$$

Finally, consider the 'typical' case, where $\omega^6/2 \leqslant N \leqslant t^{1/3}$. Condition on $\mathbf{I}$, assuming that N is in this range. Let I be an 'interval' consisting of ω consecutive integers. By (6.6)

https://doi.org/10.1017/S0963548311000666 Published online by Cambridge University Press

we have

$$\mathbb{P}(S^0 \in I \mid \mathbf{I}) = \Phi(y + \omega/\tilde{\sigma}) - \Phi(y) + O(t^{-1/2}),$$

where $y = (\min I - \tilde{\mu})/\tilde{\sigma}$. If the endpoints of I are within $o(\sqrt{t})$ of $x_n = \mu_n + O(\sigma_n)$, then using (6.5) we have $y \sim (x_n - \mu_n)/\sigma_n$, and it follows easily that

$$\mathbb{P}(S^0 \in I \mid \mathbf{I}) = \omega p_0 + O(t^{-1/2}) \sim \omega p_0, \tag{6.9}$$

where p_0 is as in (6.3), and we used $p_0 = \Theta(t^{-1/2})$ and $\omega \to \infty$ in the final approximation.

For $a = 0, 1, 2, \dots$, let J_a be the interval $[a\omega, (a+1)\omega - 1]$, and let $I_a = x_n - J_a$. Recalling that S^0 and S^1 are conditionally independent, we have

$$\mathbb{P}(S_n = x_n \mid \mathbf{I}) \geqslant \sum_{a=0}^{N/\omega+1} \mathbb{P}(S^0 \in I_a \mid \mathbf{I}) \min_{j \in J_a} \mathbb{P}(S^1 = j \mid \mathbf{I}),$$

and a corresponding upper bound with min replaced by max. Since $N \leqslant t^{1/3} = o(\sqrt{t})$, from (6.9) we have $\mathbb{P}(S^0 \in I_a \mid \mathbf{I}) \sim \omega p_0$ for all $a \leqslant N/\omega + 1$, so we obtain

$$\mathbb{P}(S_n = x_n \mid \mathbf{I}) \geqslant (1 + o(1))\omega p_0 \sum_a \min_{j \in J_a} \mathbb{P}(S^1 = j \mid \mathbf{I}),$$

and a corresponding upper bound with min replaced by max. Recall that, conditional on $\mathbf{I}$, the distribution of S^1 is simply binomial $\mathrm{Bi}(N, 1/2)$. Since the standard deviation $\sqrt{N}/2$ of S^1 is much larger than ω, elementary properties of the binomial distribution imply that

$$\sum_a \min_{j \in J_a} \mathbb{P}(S^1 = j \mid \mathbf{I}) \sim \sum_a \max_{j \in J_a} \mathbb{P}(S^1 = j \mid \mathbf{I}) \sim 1/\omega.$$

(The bulk of each sum comes near the middle of the binomial distribution, where the point probabilities hardly change within one interval; all that is actually needed here is $\omega = o(\sqrt{N})$.) This gives us an estimate for $\mathbb{P}(S_n = x_n \mid \mathbf{I})$ valid whenever $\omega^6/2 \leqslant N \leqslant t^{1/3}$, and it follows that

$$\mathbb{P}\big(S_n = x_n \wedge \omega^6/2 < N < t^{1/3}\big) \sim p_0 \mathbb{P}(\omega^6/2 < N < t^{1/3}) \sim p_0. \tag{6.10}$$

Combining (6.4), (6.8) and (6.10) gives the result. □

Results somewhat similar to Lemma 6.2 are certainly known; see, for example, McDonald [19], where Bernoulli parts are used to deduce a local limit theorem from a central limit theorem. However, the assumptions are different, and the Bernoulli part needed is much larger.

Note that uniform boundedness is not really needed in Lemma 6.2; a suitable condition on the third moments should suffice. Also, we may replace the condition $\sum_i b_1(Z_{ni}) \to \infty$ by $\sum_i b_r(Z_{ni}) \to \infty$ for all $r \in R$, where R is any set of integers with highest common factor 1. This latter condition is 'almost' necessary (after passing to a subsequence): without it there is some d (the highest common factor of the integers in R) such that even distribution modulo d will only happen because of a 'coincidence'; see the discussion in [24, Ch. VII].

We shall need a result along the lines of Theorem 6.2 but away from the central part of the distribution. This follows easily using a trick called 'exponential tilting', introduced by

Cramér [8], and suggested to us by Paul Balister. Let Z be a random variable, here with finite support, such that $\mathbb{P}(Z > 0)$ and $\mathbb{P}(Z < 0)$ are both positive. Consider the function

$$f(\alpha) = \mathbb{E}(Ze^{\alpha Z}) = \sum_x p_x x e^{\alpha x},$$

where $p_x = \mathbb{P}(Z = x)$. Note that

$$f'(\alpha) = \mathbb{E}(Z^2 e^{\alpha Z}) > 0.$$

Also, if the support of Z is contained in $[-k, k]$, then $|f''(\alpha)| = |\mathbb{E}(Z^3 e^{\alpha Z})| \leqslant k^3 e^{|\alpha| k}$.

Since $f(\alpha)$ is increasing and tends to $\pm\infty$ as $\alpha \to \pm\infty$, there is a unique $a = a(Z)$ such that $f(a) = 0$. Define

$$c = c(Z) = \mathbb{E}(e^{aZ}) = \sum_x p_x e^{ax}, \tag{6.11}$$

and let Z' be the random variable with

$$\mathbb{P}(Z' = x) = \mathbb{P}(Z = x)e^{ax}/c,$$

noting that these probabilities sum to 1 by the definition of c, and that $\mathbb{E}(Z') = 0$ by the definition of a. It is easy to check that if S_t denotes the sum of t independent copies of Z and S_t' the sum of t independent copies of Z', then

$$\mathbb{P}(S_t' = x) = \mathbb{P}(S_t = x)e^{ax}/c^t. \tag{6.12}$$

Recall that $b_1(Z)$ is the Bernoulli part of Z, defined by (6.1).

Lemma 6.3. *Let F be a finite set of integers, and let Z_n, $n \geqslant 1$, be a sequence of probability distributions supported on F, with $\liminf \mathbb{P}(Z_n < 0) > 0$ and $\liminf \mathbb{P}(Z_n > 0) > 0$. Suppose that $t = t(n) \to \infty$, and that $tb_1(Z_n) \to \infty$. Let S_n denote the sum of t independent copies of Z_n, and define $a_n = a(Z_n)$, $c_n = c(Z_n)$ and Z_n' as above. Then*

$$\mathbb{P}(S_n = x) \sim \frac{1}{\tilde{\sigma}_n \sqrt{2\pi t}} e^{-a_n x} c_n^t$$

uniformly in integer $x = o(\sqrt{t})$, where $\tilde{\sigma}_n^2$ is the variance of Z_n'.

Proof. In the light of (6.12), it suffices to prove that $\mathbb{P}(S_n' = x) \sim 1/(\tilde{\sigma}_n \sqrt{2\pi t})$, where S_n' is the sum of t independent copies of Z_n'.

Passing to a subsequence, we may suppose that $\mathbb{P}(Z_n = i)$ converges for each i, and that there are $i < 0$ and $j > 0$ for which the limit is strictly positive. It follows that the 'tilting amounts' $a = a(Z_n)$ are bounded. Hence $b_1(Z_n') = \Theta(b_1(Z_n))$, so $tb_1(Z_n') \to \infty$. Also, $\mathrm{Var}(Z_n')$ is bounded below by some positive number. Lemma 6.2 thus applies to the sum of t independent copies of Z_n', giving the result. □

We finish this section by noting some basic properties of tilting applied to random variables whose mean is close to zero.

Lemma 6.4. *Let k be fixed. If Z_n is a sequence of distributions on $\{-k, \ldots, k\}$ with $\varepsilon_n = \mathbb{E}Z_n \to 0$ and $\mathrm{Var}(Z_n) = \sigma_n^2 = \Theta(1)$, then the quantities $a_n = a(Z_n)$ and $c_n = c(Z_n)$ defined*

https://doi.org/10.1017/S0963548311000666 Published online by Cambridge University Press

above satisfy

$$a_n \sim -\varepsilon_n/\sigma_n^2 \tag{6.13}$$

and

$$1 - c_n \sim \varepsilon_n^2/(2\sigma_n^2). \tag{6.14}$$

Furthermore, $\mathrm{Var}(Z_n') \sim \sigma_n^2$*, and if* W_n *is supported on* $\{-k, \ldots, k\}$ *and may be coupled to agree with* Z_n *with probability* $1 - p_n$ *where* $p_n = o(\varepsilon_n)$*, then*

$$|a(W_n) - a_n| = O(p_n) \quad \text{and} \quad |c(W_n) - c_n| = O(\varepsilon_n p_n). \tag{6.15}$$

Proof. Suppressing the dependence on n, let $f(\alpha) = \mathbb{E}(Ze^{\alpha Z})$ as above. Then $f(0) = \mathbb{E}Z = \varepsilon$, and $f'(0) = \mathbb{E}Z^2 = \sigma^2 + \varepsilon^2 \sim \sigma^2$. Also, $f''(\alpha)$ is uniformly bounded for $\alpha \in [-1, 1]$, say. It follows easily that a_n, the solution to $f(\alpha) = 0$, satisfies (6.13). Similarly, let $g(\alpha) = \mathbb{E}(e^{\alpha Z})$. Then $g(0) = 1$, $g'(0) = \varepsilon$, $g''(0) = f'(0) \sim \sigma^2$ and $g'''(\alpha)$ is bounded for $\alpha = O(1)$. It follows that $c_n = g(a_n) = 1 + a_n\varepsilon + a_n^2\sigma^2/2 + O(\varepsilon^3)$, giving (6.14).

To see that $\mathrm{Var}(Z_n') \sim \sigma_n^2$ it is enough to note that $a_n \to 0$. The final part may be proved using the fact that for each fixed j and all $\alpha \in [-1, 1]$, we have $|\mathbb{E}(Z_n^j e^{-\alpha Z_n}) - \mathbb{E}(W_n^j e^{-\alpha W_n})| = O(p_n)$. □

7. The subcritical case

In this section we prove Theorem 1.2. Although we do use the exploration process considered in the rest of the paper, we do not track the deviations of this process from its expectation; instead we use stochastic domination arguments to 'sandwich' the process between two processes with independent increments.

We start with a lemma concerning the tail of the distribution of the time that a certain random walk with independent increments takes to first hit a given value. In the application we shall essentially take Z_n to be the distribution of $\eta - 2$, where η is the degree of a vertex of our graph chosen with probability proportional to its degree (see Section 2). In fact, we shall adjust the distribution slightly both to allow us to use stochastic domination, and to meet the condition $tb_1(Z_n) \to \infty$. In what follows we often suppress dependence on n in the notation. Recall that the Bernoulli part $b_1(Z)$ of a distribution Z is defined by (6.1).

Lemma 7.1. *Let* $k \geqslant 1$ *be fixed, and let* Z_n *be a sequence of probability distributions on* $\{-1, 0, 1, \ldots, k\}$ *converging in distribution to some* Z *with* $\mathrm{Var}(Z) > 0$*, such that* $\varepsilon = \varepsilon(n) = -\mathbb{E}Z_n > 0$ *and* $\varepsilon \to 0$*. Let* $\mathbf{W}_n = (W_t)_{t \geqslant 0}$ *be a random walk with* $W_0 = 0$ *and the increments independent with distribution* Z_n*, and let* $\tau_r = \tau_r(n) = \inf\{t : W_t = -r\}$*. Suppose that* $r \geqslant 1$ *is fixed, and that* $t = t(n)$ *is such that* $\varepsilon^2 t \to \infty$ *and* $tb_1(Z_n) \to \infty$*. Then*

$$\mathbb{P}(\tau_r \geqslant t) \sim c_{r,Z}\,\delta^{-1} t^{-3/2} e^{-\delta t}, \tag{7.1}$$

where $\delta = \delta_n = -\log(c(Z_n))$ *with* $c(Z_n)$ *defined as in* (6.11)*, and* $c_{r,Z} > 0$ *is some constant depending on* r *and* Z*.*

https://doi.org/10.1017/S0963548311000666 Published online by Cambridge University Press

Proof. We start with the case $r = 1$. Here Spitzer's lemma [27] gives $\mathbb{P}(\tau_1 = t) = \mathbb{P}(W_t = -1)/t$; indeed, given a sequence of possible values $x_1, \ldots, x_t$ of the first t increments summing to -1, there is exactly one cyclic permutation of $(x_1, \ldots, x_t)$ such that the walk with the permuted increments stays non-negative up to step t.

Lemma 6.4 gives $a_n \to 0$ and $c_n = 1 - \Theta(\varepsilon^2)$. Thus $\delta = \Theta(\varepsilon^2)$ and, writing $\tilde{\sigma}^2$ for the variance of Z'_n, $\tilde{\sigma} \sim \sigma$. Lemma 6.3 thus gives

$$\mathbb{P}(W_t = -1) \sim \frac{1}{\sigma\sqrt{2\pi t}} e^{-\delta t},$$

whenever $t \to \infty$ with $t b_1(Z_n) \to \infty$. Hence

$$\mathbb{P}(\tau_1 = t) \sim t^{-3/2} \frac{1}{\sigma\sqrt{2\pi}} e^{-\delta t}. \tag{7.2}$$

When $\delta t \to \infty$, summing over t easily gives

$$\mathbb{P}(\tau_1 \geqslant t) \sim \delta^{-1} t^{-3/2} \frac{1}{\sigma\sqrt{2\pi}} e^{-\delta t}. \tag{7.3}$$

Indeed, the sum is dominated by the first $O(\delta^{-1}) = O(\varepsilon^{-2})$ terms, and in this range $t^{-3/2}$ hardly changes.

For general r we note that τ_r is distributed as the sum of r independent copies $\tau^{(1)}, \ldots, \tau^{(r)}$ of τ_1. Since $\sum_{t \geqslant 1} t^{-3/2}$ converges, using (7.2) it is easy to see that the dominant contribution to $\mathbb{P}(\tau_r \geqslant t)$ comes from the case that one of the $\tau^{(i)}$ is large and the others are $O(1)$. Convergence in distribution of Z_n implies that for each j, $\mathbb{P}(\tau_1 = j)$ converges to some limit, so (7.1) follows from (7.3). □

We are now ready to prove Theorem 1.2.

Proof of Theorem 1.2. Let $Z = Z_n$ denote the distribution of $\eta - 2$, recalling that η is the degree of a vertex chosen with probability proportional to its degree. Passing to a subsequence, we may assume that Z_n converges in distribution to some distribution Z^*. Nonetheless, in what follows we must work with the actual distribution Z_n rather than the limit, since the bounds are sensitive to small changes in the distribution of Z_n.

Note that $Z = Z_n$ is supported on $\{-1, 0, \ldots, d_{\max} - 2\}$, and that by (1.7) we have $\mathrm{Var}(Z) = \Theta(1)$. Also, $\mathbb{E}Z = -\varepsilon$, where by assumption $\varepsilon \to 0$ and $\varepsilon^3 n \to \infty$. Let $\delta = \delta(Z) = -\log(c(Z))$ be defined as above, noting that

$$\delta \sim \frac{\varepsilon^2}{2v_0} = \Theta(\varepsilon^2) \tag{7.4}$$

by Lemma 6.4. Note also that δ is exactly the quantity δ_n appearing in the statement of Theorem 1.2.

Let $\Lambda = \varepsilon^3 n$, recalling that $\Lambda \to \infty$ by assumption, and set

$$t^+ = \delta^{-1}(\log \Lambda - \tfrac{5}{2} \log\log \Lambda + \omega)$$

for some $\omega \to \infty$ with $\omega = o(\log\log \Lambda)$.

Let Z' be defined in the same way as Z, except that we first remove the $2d_{\max} t^+ = o(n)$ non-isolated vertices of lowest degree from our degree sequence. Then, conditional on the

https://doi.org/10.1017/S0963548311000666 Published online by Cambridge University Press

first $t \leqslant 2t^+$ steps of our process, using (2.2) the distribution of the next increment $X_{t+1} - X_t = \eta_{t+1} - 2 - 2\theta_{t+1} \leqslant \eta_{t+1} - 2$ is stochastically dominated by Z'. Let $\gamma = \varepsilon^{3/2}/\log\Lambda$, say, noting that $\gamma \to 0$, $\gamma t^+ \sim \varepsilon^{-1/2} \to \infty$, and $\varepsilon\gamma t^+ \to 0$. Define Z^+ by modifying the distribution of Z' as follows. Pick some k such that $\mathbb{P}(Z' = k) \geqslant 2\gamma$, and shift mass γ from k to $k+1$. Note that $b_1(Z^+) \geqslant \gamma$, so $t^+ b_1(Z^+) \to \infty$. Also Z^+ stochastically dominates Z', and Z^+ and Z may be coupled to agree with probability $1 - p$, where

$$p = O(t^+/n + \gamma).$$

Note that $\varepsilon^{-1}t^+/n = \Theta((\log\Lambda)/\Lambda) = o(1)$, and clearly $\gamma = o(\varepsilon)$, so $p = o(\varepsilon)$.

Considering the random walk with independent increments distributed as Z^+, writing $\mathcal{C}_1$ for the first component revealed by our exploration, stochastic domination and Lemma 7.1 give

$$\mathbb{P}(|\mathcal{C}_1| \geqslant t^+) = \mathbb{P}\big(\inf\{t : X_t = -2\} \geqslant t^+\big) \leqslant (1+o(1))c(\delta^+)^{-1}(t^+)^{-3/2}e^{-\delta^+ t^+},$$

where δ^+ is defined as δ, but using Z^+ in place of Z, and $c = c_{2,Z^*}$ is a positive constant.

Lemma 6.4 gives $\delta^+ \sim \delta$, and indeed $|\delta^+ - \delta| = O(\varepsilon p)$. Thus $|\delta^+ t^+ - \delta t^+| = O(\varepsilon p t^+)$, which is easily seen to be $o(1)$. Thus the bound above can be written more simply as

$$\begin{aligned}\mathbb{P}(|\mathcal{C}_1| \geqslant t^+) &\leqslant (1+o(1))c\delta^{-1}(t^+)^{-3/2}e^{-\delta t^+}\\ &= c\frac{t^+}{n}n\delta^{-1}(t^+)^{-5/2}\Lambda^{-1}(\log\Lambda)^{5/2}e^{-\omega}\\ &\sim c\frac{t^+}{n}(2v_0)^{-3/2}e^{-\omega}\\ &= \frac{t^+}{n}\Theta(e^{-\omega}) = o(t^+/n).\end{aligned}$$

If our graph $G_{\mathbf{d}}$ contains a component of order at least t^+, then the probability that we explore this component first is at least $(2t^+ - 2)/(d_{\max}n) = \Theta(t^+/n)$. It follows that $\mathbb{P}(L_1 \geqslant t^+) = o(1)$, proving the upper bound in (1.17).

Turning to the lower bound, we use stochastic domination in the other direction. This time we must account for back-edges. At a given step $t \leqslant 2t^+$, the (conditional, given the history) probability of forming a back-edge is $O(t^+/n)$, simply because there can only be $O(t^+)$ active stubs. It follows that we can define a distribution Z^- that may be coupled to agree with Z with probability $1 - O(t^+/n + \gamma)$ so that the conditional distribution of $X_{t+1} - X_t$ stochastically dominates Z^- whenever $t \leqslant 2t^+$. Setting

$$t^- = \delta^{-1}(\log\Lambda - \tfrac{5}{2}\log\log\Lambda - \omega), \tag{7.5}$$

the argument above adapts easily to prove that

$$\mathbb{P}(|\mathcal{C}_1| \geqslant t^-) \sim c\frac{t^-}{n}(2v_0)^{-3/2}e^{\omega}.$$

Let $I = [t^-, t^+]$ and write $t = (t^- + t^+)/2$, say. Noting that $t^- \sim t^+ \sim t$, the bounds above combine to give $\mathbb{P}(|\mathcal{C}_1| \in I) \sim c'e^{\omega}t/n$, where $c' = c(2v_0)^{-3/2}$. Let N denote the number of components $\mathcal{C}$ with $|\mathcal{C}| \in I$. It is easy to check that with high probability no such component will have significantly more than $|\mathcal{C}|$ edges. Since our initial vertex is chosen

https://doi.org/10.1017/S0963548311000666 Published online by Cambridge University Press

with probability proportional to its degree, it follows that $\mathbb{P}(|\mathcal{C}_1| \in I) \sim (\mathbb{E}N)2t/(\mu_1 n)$, where μ_1 is the overall average degree. Hence

$$\mathbb{E}N \sim c'\mu_1 e^{\omega}/2 \to \infty.$$

Finally, with $\mathcal{C}_2$ the second component explored by our process, we have

$$\mathbb{P}(|\mathcal{C}_1|, |\mathcal{C}_2| \in I) \sim \mathbb{E}(N(N-1))(2t/(\mu_1 n))^2. \tag{7.6}$$

The estimates above apply just as well to bound $\mathbb{P}(|\mathcal{C}_2| \in I \mid |\mathcal{C}_1| \in I)$: throughout, we only needed that at most $2t^+$ vertices had been 'used up'. We find that the left-hand side in (7.6) is asymptotically $(c'e^{\omega}t/n)^2$, so it follows that

$$\mathbb{E}(N(N-1)) \sim (c'\mu_1 e^{\omega}/2)^2 \sim (\mathbb{E}N)^2.$$

Since $\mathbb{E}N \to \infty$, this gives $\mathbb{E}(N^2) \sim (\mathbb{E}N)^2$, so Chebyshev's inequality implies that $\mathbb{P}(N > 0) \to 1$, completing the proof of (1.17).

The argument for (1.15) is essentially the same; we simply replace $-\omega$ by $+x$ in the definition (7.5) of t^-. With N the number of components with order between this new t^- and t^+ it follows as above that $\mathbb{E}N \sim \alpha = c'\mu_1 e^{-x}/2$. Moreover, arguing as for $N(N-1)$ above, for each fixed r the rth factorial moment of N converges to α^r. It follows by standard results that N converges in distribution to a Poisson distribution with mean α, so $\mathbb{P}(N = 0) \to e^{-\alpha}$. Note that the constant c in (1.15) is

$$c = c'\mu_1/2 = c_{2,Z^*}(2v_0)^{-3/2}\mu_1/2 \sim c_{2,Z^*}2^{-5/2}\mu_1^{5/2}\mu_3^{-3/2}, \tag{7.7}$$

with c_{2,Z^*} as in Lemma 7.1. Since we were considering a subsequence on which Z_n (the distribution of $\eta - 2$) converges to Z^*, we have $c_{2,Z_n} \to c_{2,Z^*}$, and so can take $c = c_{2,Z_n}2^{-5/2}\mu_1^{5/2}\mu_3^{-3/2}$ in (1.15). □

Acknowledgement

The author would like to thank Béla Bollobás for many helpful discussions, as well as for the invitation to visit the University of Memphis, and Paul Balister for suggesting the use of 'tilting' in the proof of Lemma 6.3.

References

[1] Aldous, D. (1997) Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.* **25** 812–854.

[2] Bender, E. A. and Canfield, R. E. (1978) The asymptotic number of labeled graphs with given degree sequences. *J. Combin. Theory Ser. A* **24** 296–307.

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

[4] Bollobás, B. (1984) The evolution of random graphs. *Trans. Amer. Math. Soc.* **286** 257–274.

[5] Bollobás, B. and Riordan, O. (2009) Random graphs and branching processes. In *Handbook of Large-Scale Random Networks* (B. Bollobás, R. Kozma and D. Miklós eds), Vol. 18 of *Bolyai Soc. Math. Studies*, pp. 15–115.

[6] Bollobás, B. and Riordan, O. (2012) Asymptotic normality of the size of the giant component via a random walk. *J. Combin. Theory Ser. B* **102** 53–61.

[7] Brown, B. M. (1971) Martingale central limit theorems. *Ann. Math. Statist.* **42** 59–66.
[8] Cramér, H. (1938) Sur un nouveau théorème-limite de la théorie des probabilités. *Actualités Scientifiques et Industrielles* **736** 5–23.
[9] Ding, J., Kim, J. H., Lubetzky, E. and Peres, Y. (2010) Diameters in supercritical random graphs via first-passage percolation. *Combin. Probab. Comput.* **19** 729–751.
[10] Doob, J. L. (1953) *Stochastic Processes*, Wiley/Chapman and Hall.
[11] Fountoulakis, N. (2007) Percolation on sparse random graphs with given degree sequence. *Internet Mathematics* **4** 329–356.
[12] Goerdt, A. (2001) The giant component threshold for random regular graphs with edge faults. *Theoret. Comput. Sci.* **259** 307–321.
[13] Janson, S. (2009) On percolation in random graphs with given vertex degrees. *Electron. J. Probab.* **14** 87–118.
[14] Janson, S. and Luczak, M. J. (2009) A new approach to the giant component problem. *Random Struct. Alg.* **34** 197–216.
[15] Kang, M. and Seierstad, T. G. (2008) The critical phase for random graphs with a given degree sequence. *Combin. Probab. Comput.* **17** 67–86.
[16] Karp, R. M. (1990) The transitive closure of a random digraph. *Random Struct. Alg.* **1** 73–93.
[17] Łuczak, T. (1990) Component behavior near the critical point of the random graph process. *Random Struct. Alg.* **1** 287–310.
[18] Martin-Löf, A. (1986) Symmetric sampling procedures, general epidemic processes and their threshold limit theorems. *J. Appl. Probab.* **23** 265–282.
[19] McDonald, D. R. (1979) On local limit theorem for integer valued random variables. *Teor. Veroyatnost. i Primenen.* **24** 607–614. See also *Theory Probab. Appl.* **24** (1980), 613–619.
[20] Molloy, M. and Reed, B. (1995) A critical point for random graphs with a given degree sequence, *Random Struct. Alg.* **6** 161–179.
[21] 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.
[22] Nachmias, A. and Peres, Y. (2007) Component sizes of the random graph outside the scaling window. *ALEA Lat. Amer. J. Probab. Math. Stat.* **3** 133–142.
[23] Nachmias, A. and Peres, Y. (2010) Critical percolation on random regular graphs. *Random Struct. Alg.* **36** 111–148.
[24] Petrov, V. V. (1975) *Sums of Independent Random Variables*, translated from the Russian by A. A. Brown, Vol. 82 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*, Springer.
[25] Pittel, B. (2008) Edge percolation on a random regular graph of low degree. *Ann. Probab.* **36** 1359–1389.
[26] Pittel, B. and Wormald, C. (2005) Counting connected graphs inside-out. *J. Combin. Theory Ser. B* **93** 127–172.
[27] Spitzer, F. (1956) A combinatorial lemma and its application to probability theory. *Trans. Amer. Math. Soc.* **82** 323–339.