SUDDEN EMERGENCE OF A GIANT k-CORE IN A RANDOM GRAPH.

Boris Pittel, Joel Spencer, Nicholas Wormald*

Department of Mathematics
The Ohio State University
Columbus, Ohio 43210

Courant Institute
New York University
New York, NY 10012

and

Department of Mathematics
University of Melbourne
Parkville, VIC 3052
Australia

Abstract. The k-core of a graph is the largest subgraph with minimum degree at least k. For the Erdős-Rényi random graph $G(n,m)$ on n vertices, with m edges, it is known that a giant 2-core grows simultaneously with a giant component, that is when m is close to $n/2$. We show that for $k \geq 3$, with high probability, a giant k-core appears suddenly when m reaches $c_k n/2$; here $c_k = \min_{\lambda>0} \lambda/\pi_k(\lambda)$ and $\pi_k(\lambda) = \mathbf{P}\{\text{Poisson}(\lambda) \geq k-1\}$. In particular, $c_3 \approx 3.35$. We also demonstrate that, unlike the 2-core, when a k-core appears for the first time it is very likely to be giant, of size $\approx p_k(\lambda_k)n$. Here λ_k is the minimum point of $\lambda/\pi_k(\lambda)$ and $p_k(\lambda_k) = \mathbf{P}\{\text{Poisson}(\lambda_k) \geq k\}$. For $k=3$, for instance, the newborn 3-core contains about $0.27n$ vertices. Our proofs are based on the probabilistic analysis of an edge deletion algorithm that always finds a k-core if the graph has one.

1991 *Mathematics Subject Classification.* Primary 05C80, 05C85, 60C05; Secondary 60F10, 60G42, 60J10.

Key words and phrases. Random, graph, multigraph, core, branching process, deletion algorithm, martingales, stopping times, estimates, phase transition.

* Research supported by the Australian Research Council.

Typeset by $\mathcal{AMS}$-TEX

1. Introduction. The random graph process on n vertices is the probability space of all the nested sequences of graphs

$$G(n,0) \subset G(n,1) \subset \cdots \subset G(n,N),$$

$N = \binom{n}{2}$, with vertex set $V = \{1,\dots,n\}$, such that $G(n,m)$ has m edges and each sample sequence has the same probability, $1/N!$. In particular, the random "snapshot" $G(n,m)$ is uniformly distributed on the set of all $\binom{N}{m}$ graphs with m edges. An event H_n occurs in $\{G(n,m)\}$ *with high probability* (*whp*) if $\mathbf{P}H_n \to 1$ as $n \to \infty$. (It is understood that $\{H_n,\ n = 1,2,\dots\}$ is a sequence of events.) According to a classic result by Erdös and Rényi [10] (see also Bollobás [4], Łuczak [19], Janson et al. [14], and Łuczak et al. [22]), for large n the likely structure of $G(n,m)$ undergoes an abrupt change (phase transition) when m passes through $n/2$. Namely, whp this is a birth time of a giant component, that is a component of size of the order of n. More precisely, if $m \approx cn/2$ and $c > 1$ then the giant component whp contains about $\alpha(c)n$ vertices, where $\alpha(c) = 1 - t(c)/c$ and $t(c) \in (0,1)$ is the smaller root of $te^{-t} = ce^{-c}$. Notice that $\alpha(1) = 0$, so the percentage of vertices in the giant component is low if c is close to 1 (from above).

Why is 1 the threshold value of c? A semiformal reasoning goes like this (cf. Karp [17]). Fix a vertex v and consider a subgraph of $G(n, m = cn/2)$ formed by the vertices whose distance from v is at most $\epsilon \log n$. It can be proved that, for ϵ sufficiently small and fixed, whp this subgraph can be looked at as a genealogical tree of the first $\lfloor \epsilon \log n \rfloor$ generations in the Poisson(c) branching process. Such a process either almost surely suffers extinction, or with positive probability ($= 1 - t(c)/c$) survives indefinitely, dependent upon whether $c \le 1$ or $c > 1$. So one should expect that for $c > 1$—with probability $\alpha(c)$—a generic vertex v belongs to a giant component, and the average size of the component is about $\alpha(c)n$.

For $k \ge 2$, does a newborn giant component already contain a k-connected subgraph? If not, how many additional edges later can one expect appearance of such a subgraph? These questions are intimately related to the appearance of the k-core, which was defined in [3] as the unique maximal subgraph with minimum degree at least k. Even the question of the size of the first 2-core was not trivial. This is just the length of the first cycle. Janson [14] derived the limiting distribution of the first cycle size, thus showing that this size is bounded in probability as n tends to infinity. Later Bollobás [6] (see also Bollobás and Rasmussen [7]) rederived this distribution using the martingale techniques. Still later, Flajolet et al. [12] showed the expected length of the first cycle to be asymptotic to $n^{1/6}$. (There is no contradiction here: the tail of the limiting distribution F is such that the corresponding expected value— $\int_0^\infty x\,dF(x)$—is infinite.)

For general k, the results have not been nearly this precise. Bollobás [3] established whp the existence of a k-connected subgraph for $m = cn/2$ with $8 \le k + 3 \le c/2$, $c \ge 67$, indicating that no attempt was made to get the best bounds his approach might deliver. The proof consisted of showing that (a) whp the k-core exists, and (b) whp the k-core is k-connected. Pittel [27] proved that—for every $c > 1$—the giant component contains a 2-connected subgraph of the likely size $\approx \beta(c)n$, with $\beta(c) = (1 - t(c))\alpha(c)$. This 2-connected subgraph cannot, however, be expected to contain a 3-connected subgraph for each $c > 1$! Indeed, Łuczak [18] proved that for $c < 1.24$ whp $G(n, m = cn/2)$ does not contain a subgraph of average degree at least 3. Define c_k as the infimum of $c's$ such that

$G(n,m)$ ($m \geq cn/2$) whp has a k-core. Then Łuczak's result means that $c_3 \geq 1.24$. Chvátal [8]—who introduced the notion of c_k— was able to show that $c_3 \geq 2.88$, and claimed that the same method yielded, for instance, $c_4 \geq 4.52$ and $c_5 \geq 6.06$.

More recently, Łuczak [20], [21] proved for *every* fixed $k \geq 3$ that in the graph process $\{G(n,m)\}$, whp the k-core, if present, is k-connected and contains at least $0.0002n$ vertices. (Being content apparently with establishing any linear bound, Łuczak did not try to get the best bound his method could deliver.)

Using Łuczak bound, Molloy and Reed [25] were able to improve significantly the existing bounds for c_3, c_4 and c_5, showing that $c_3 \in (3.35, 3.59)$, for instance. The proof involved a computer-aided analysis of the recurrence equations which described the mean–values behavior of a few thousand steps of an algorithm that strips away vertices of degree $< k$.

Despite the progress, the question of exact values of c_k and the likely sizes of k-cores ($k \geq 3$) for various values of c has so far remained open. In the present paper we answer these questions, and in addition show that, unlike the 2-core, for $k \geq 3$ the first k-core to appear is whp very large—with approximately $p_k n$ vertices, for some constant p_k which we determine. Our approach is based on probabilistic analysis of a simple algorithm which, for a given graph, either finds a k-core, or correctly diagnoses the absence of a k-core.

Here are our principal results. In the rest of the paper, $k \geq 3$ is a fixed integer.

Given $\lambda > 0$, let $Z(\lambda)$ denote a Poisson distributed random variable with mean λ. Introduce $p_k(\lambda) = \mathbf{P}\{Z(\lambda) \geq k\}$ and $\pi_k(\lambda) = \mathbf{P}\{Z(\lambda) \geq k-1\}$. Define

$$\gamma_k = \inf\{\frac{\lambda}{\pi_k(\lambda)} : \lambda > 0\}. \tag{1.1}$$

Since for $k \geq 3$ the function $\lambda/\pi_k(\lambda)$ approaches ∞ as $\lambda \to 0$ or ∞, the infimum in (1.1) is attained at a point $\lambda_k > 0$. Clearly, the equation

$$c = \frac{\lambda}{\pi_k(\lambda)} \tag{1.2}$$

has no root for λ if $c < \gamma_k$. If $c > \gamma_k$ there are two roots. Let $\lambda_k(c)$ denote the larger root; $\lambda_k(c)$ is a continuous, strictly increasing function of $c > \gamma_k$ and $\lambda_k := \lim_{c \downarrow \gamma_k} \lambda_k(c)$ satisfies

$$\gamma_k = \frac{\lambda_k}{\pi(\lambda_k)}. \tag{1.3}$$

Theorem 1. *Suppose $c \leq \gamma_k - n^{-\delta}$, $\delta \in (0, 1/2)$ being fixed. Let $\epsilon \in (0,1)$ be chosen arbitrarily small. Then the probability that $G(n, m = cn/2)$ has a k-core with at least ϵn vertices is $O(\exp(-n^{\rho}))$, $\forall \rho < (0.5 - \delta) \wedge 1/6$. The probability that there is a k-core of any size ($\geq k+1$, of course) is $O(n^{-(k-2)(k+1)/2})$.*

Theorem 2. *Suppose $c \geq \gamma_k + n^{-\delta}$, $\delta \in (0, 1/2)$ being fixed. Fix $\sigma \in (3/4, 1 - \delta/2)$ and define $\overline{\zeta} = \min\{2\sigma - 3/2, 1/6\}$. Then with probability $\geq 1 - O(\exp(-n^{\zeta}))$ ($\forall \zeta < \overline{\zeta}$), the random graph $G(n, m = cn/2)$ contains a giant k-core of size $np_k(\lambda_k(c)) + O(n^{\sigma})$.*

Theorem 3. *Denote* $p_k = p_k(\lambda_k)$. *Let* $\sigma \in (3/4, 1)$ *be fixed. For every* $\epsilon \in (0, p_k)$, *the probability that a* k*-core of the random graph* $G(n,m)$ *at any "time"* $m \in (0, N)$ *has size from* $[\epsilon n,\, p_k(n - n^\sigma)]$ *is* $O(\exp(-n^\zeta))$, $\forall \zeta < \overline{\zeta}$.

Theorems 1 and 2 taken together imply that $c_k = \gamma_k$. We see also that the random birth time, m_k, of a giant k-core is sharply concentrated (for large n) around $c_k n/2$: with subexponentially high probability m_k is in $[c_k n/2 - n^{1-\delta},\, c_k n/2 + n^{1-\delta}]$, $\forall \delta < 1/2$. Combining Theorems 1, 2 and 3, we obtain that with subexponentially high probability the size of a newborn k-core is close to $p_k n$. So, at a random moment $m \sim c_k n/2$, we observe a sudden appearance ("explosion") of a giant k-core that already contains a positive fraction of all vertices, asymptotic to p_k. For c safely above c_k, the fraction of vertices in the k-core is a continuous function of c.

Numerically, $c_3 \approx 3.35, p_3 \approx 0.27$; $c_4 \approx 5.14, p_4 \approx 0.43$; $c_5 \approx 6.81, p_5 \approx 0.55$. It can be easily shown that for large k

$$c_k = k + \sqrt{k \log k} + O(\log k). \tag{1.4}$$

It has been known, Łuczak [20], that

$$c_k = k + O(k^{1/2+\epsilon}), \quad \forall \epsilon > 0.$$

Note. A recently discovered algorithm for generating minimal perfect hash functions uses random r-uniform hypergraphs, in which the threshold of the appearance of the r-analogue of a cycle is crucial (see Havas et al. [13], where rough estimates on the threshold are derived). The problem of determining this threshold is very similar to that for the k-core of a random graph. A more thorough analysis is made by Majewski et al. [24], where it is shown that a constant analogous to γ_k fits the experimental data very well. No attempt was made there to give a rigorous argument, but the methods of the present paper could possibly be extended to do the job.

The rest of the paper is organized as follows. In Section 2 we discuss a heuristic connection between the deletion processes for the random graph and the genealogical tree of the Poisson branching process. In Section 3 we describe the deletion process for the random graph in full detail and prove that the resulting sequence of states $\{\mathbf{w}(\tau)\}_{\tau \geq 0}$ is a Markov chain. (A state $\mathbf{w}$ is a $(k+1)$-tuple whose components are the numbers of vertices of various degrees and the number of edges in the current graph.) Next (Section 4) we obtain the asymptotic approximations for the one-step transition probabilities of the Markov chain, including the states with the arbitrarily low number of light vertices. We then use these approximations (Section 5) to derive the asymptotic equations for the conditional expectations $\mathrm{E}[\mathbf{w}(\tau+1)|\mathbf{w}(\tau)]$. These equations make it plausible that a corresponding system of $(k+1)$ differential equations has a solution which is whp followed approximately by the sequence $\{\mathbf{w}(\tau)\}$, at least as long as the components of $\mathbf{w}(\tau)$ remain large, of order n that is. This is of course a rather general principle, but a formal justification is not easy in many important cases. Wormald [29] rigorously proved this approximation property for the graph-related random processes, such as ours, under quite general conditions, and his results can be used in our case to get a clear idea as to when the birth of a giant k-core should be expected. (The solution of the differential equations is a particularly sharp approximation when the parameter c is bounded away from the critical value c_k.) However, the deletion process we study is intrinsically difficult in that we need to analyze its almost sure behavior also at the nearterminal moments τ when some of the components of $\mathbf{w}(\tau)$ become small, just of order $O(1)$. (There is a certain analogy here with epidemic processes in a large population triggered by just a few infected individuals, cf. Pittel [26].) Fortunately, however, the differential equations for the deletion process have a pair of remarkable integrals that involve the principal characteristics of $\mathbf{w}(\tau)$. Using this property, and the asymptotics for the transition probabilities, we construct in Section 6 some auxiliary supermartingales of the exponential type, and the desired probabilistic bounds follow then eventually from the Optional Sampling Theorem for supermartingales.

2. Branching Poisson Process Connection. Let a graph $G(V, E)$ be given. Here is a simple algorithm that either finds the k-core in G or establishes its absence. At the first round we delete all the *light* vertices, that is the vertices with degree at most $k-1$; none of them may belong to a k-core. At the next round, we delete all the light vertices in the remaining graph. And so on. The process stops when either the remaining vertices are all *heavy*, that is each with degree at least k, or no vertices are left after the last round. In the first case the remaining graph is the k-core of the graph G; in the second case a k-core does not exist.

Consider the fate of a vertex $v \in V$. If v is heavy it is not deleted in the first round; it will stay after the second round as well provided that it has remained heavy in the graph left after the first round. It is clear intuitively that even if v is very heavy initially, it may be eliminated after several rounds if—in the original graph G—there are too many light vertices dangerously close to v.

Let G be a sample point for the random graph $G(n, m)$. Let $p(n, m)$ denote the probability that a fixed vertex v will survive the deletion process. Clearly, $np(n, m)$ is the expected size of the k-core. It is difficult to estimate $p(n, m)$ rigorously via analyzing the above algorithm. (Later we will achieve this goal by studying a less radical algorithm that deletes at each step only the edges incident to one of the light vertices.) Here is an admittedly loose attempt of such an analysis that suggests—some serious gaps and leaps of faith notwithstanding— an intuitive explanation of why the k-core appears when the number of edges passes through $\gamma_k n/2$. In the light of the algorithm, and by analogy with the giant component phenomenon, we should expect $p(n, m)$ to be close to $\phi(c) := \lim_{j\to\infty} \phi_j(c)$, $c := 2m/n$. Here $\phi_j(c)$ is the probability that the progenitor of the genealogical tree for the Poisson(c) branching process survives a deletion process applied to the first j generations of that tree. In the first round of the process we delete all the members of the j-th generation (if there are any) who have less than $k-1$ children, i.e. the descendants in the $(j+1)$-th generation. (The degree of every such member is at most $k-1$.) Next we delete the members of the $(j-1)$-th generation who have less than $k-1$ children that survived the first round. And so on, until we get to the progenitor himself, who survives if at least k of his children survived the previous rounds. To compute $\phi_j(c)$, introduce also $\varphi_j(c)$ which is the probability that the progenitor has at least $k-1$ surviving children. Since each of the children survives with probability $\varphi_{j-1}(c)$, independently of the other siblings, the total number of surviving children is Poisson$(c\varphi_{j-1}(c))$ distributed. Therefore, for $j \ge 1$

$$\begin{aligned} \varphi_j(c) &= \mathbf{P}\{Z(c\varphi_{j-1}(c)) \ge k-1\}, \\ \phi_j(c) &= \mathbf{P}\{Z(c\varphi_{j-1}(c)) \ge k\}, \end{aligned} \tag{2.1}$$

where $\varphi_0(c) := 1$. There exists $\varphi(c) = \lim_{j\to\infty} \varphi_j(c)$ since $(\varphi_j(c))_{j\ge1}$ is clearly decreasing, and

$$\begin{aligned} \varphi(c) &= \mathbf{P}\{Z(c\varphi(c)) \ge k-1\}, \\ \phi(c) &= \mathbf{P}\{Z(c\varphi(c)) \ge k\}. \end{aligned} \tag{2.2}$$

If $\varphi(c) > 0$, with a notation $\lambda = c\varphi(c)$ the second equation in (2.2) becomes (1.2), which is solvable iff $c \ge \gamma_k$. (!) Thus, we are led to believe that $\lim_{n\to\infty} p(n,m) = 0$ if $c < \gamma_k$, and for $c > \gamma_k$

$$\begin{aligned} \lim_{n\to\infty} p(n,m) &= \phi(c) \\ &= \mathbf{P}\{Z(\lambda) \ge k\} = p_k(\lambda) > 0. \end{aligned}$$

So any given vertex from V belongs to a k-core with probability close to $p_k(\lambda)$, whence whp the core size has to be close to $np_k(\lambda)$. Leaving aside the probabilistic bounds, that is what is claimed in Theorems 1 and 2! (This heuristic derivation is inspired by Karp–Sipser's probabilistic analysis of a greedy matching algorithm, [16].)

We emphasize though that while the reasoning for the subcritical case $c < c_k$ can be made rigorous, we do not see any way to do the same for the supercritical case $c > c_k$. Especially daunting would be to use the connection with the branching process for a proof of the explosion phenomenon (Theorem 3).

3. The edge deletion process and its Markovian structure. Here is a slowed down version of the deletion process that lies in the heart of our proofs. At each step we form a list of all nonisolated light vertices of the current graph, select a vertex i from this list at random uniformly, and delete all the edges incident to i, thus making it isolated. The step is repeated so long as there are edges to be deleted and the current set, H, of heavy vertices is nonempty. At the end, either $H \ne \emptyset$ and so H is the vertex set of the k-core in the initial graph, or $H = \emptyset$ and so there is no k-core.

The idea behind our choice of this particular deletion process is that hopefully its work on the sample point of $G(n,m)$ can be described by a Markov chain amenable to asymptotic study. For this to happen, the state space of such a chain must be sufficiently simple, far simpler than, say, the set of all graphs on V. Here is a natural candidate for such a space: it consists of all $(k+1)$-tuples of nonnegative integers $\mathbf{w} = (\mathbf{v}, \mu)$, $\mathbf{v} = (v_0, v_1, \dots, v_{k-1})$, where v_j is the number of light vertices with degree j, $(0 \le j \le k-1)$, and μ is the total number of edges. Define

$$\overline{v} := n - v_0 - v \text{ where } v := \sum_{j=1}^{k-1} v_j.$$

The number of heavy vertices is then $\overline{v}$. So we want to describe the deletion process by the sequence $\{\mathbf{w}(\tau)\}$ where $\mathbf{w}(\tau)$ is the state after τ steps. Let us denote the corresponding sequence of graphs by $\{G(\tau)\}$. Now, $G(0) = G(n, m)$, that is $G(0)$ is distributed uniformly on the set of all graphs with m edges. Consequently, given the value of the whole tuple $\mathbf{w}(0)$, the conditional distribution of $G(0)$ remains uniform. For us to be able to get back to the sequence $\{G(\tau)\}$, it would be decisively important to have the same property for all τ, namely that $G(\tau)$ is uniform if conditioned on $\mathbf{w}(\tau)$. As we shall see shortly, the process $\{G(\tau)\}$ does have the desired properties.

Note. As an alternative algorithm, at each step one can pick a nonisolated light vertex at random uniformly and delete an edge chosen at random uniformly among all the edges incident to the vertex. In yet another algorithm, each step consists of deletion of an edge chosen at random uniformly among all the edges incident to the light vertices of the current graph. Interestingly, neither of these appealing schemes produces a sequence $\{\mathbf{w}(\tau)\}$ that is Markov!

For a given graph G, introduce $\mathbf{w}(G) = (\mathbf{v}(G), \mu(G))$ where $\mathbf{v}(G) = (v_0(G), \dots, v_{k-1}(G))$, $v_j(G)$ is the total number of vertices with degree j, and $\mu(G)$ is the total number of edges. Similarly define $v(G) = \sum_{j=1}^{k-1} v_j(G)$ (the number of non-isolated light vertices) and $\overline{v}(G) = n - v_0(G) - v(G)$. Given a $(k+1)$-tuple $\mathbf{w}$, define $\mathcal{G}(\mathbf{w}) = \{G : \mathbf{w}(G) = \mathbf{w}\}$, and set $h(\mathbf{w}) = |\mathcal{G}(\mathbf{w})|$. Let us choose the initial graph G from $\mathcal{G}(\mathbf{w})$ at random uniformly and start the deletion process. We obtain a random graph sequence $\{G(\tau)\}$ defined for $0 \le \tau \le T$, where T is the total number of the deletion steps (*stopping time*): either there are no heavy vertices in $G(T)$, or, besides the isolated vertices, there are left only some heavy vertices. Let $\{\mathbf{w}(\tau)\} = \{\mathbf{w}(G(\tau))\}$; clearly, denoting $\mathbf{w}(T) = \mathbf{w}$, we have either

$$\overline{v} = 0,$$

or

$$\overline{v} > 0 \quad \text{but } v = 0.$$

We shall call such $\mathbf{w}$ *terminal*. For convenience we can extend both sequences, setting $G(\tau) \equiv G(T)$, $\mathbf{w}(\tau) \equiv \mathbf{w}(T)$ for all $\tau > T$.

The sequence $\{G(\tau)\}$ is obviously Markov. Given two graphs, G and G', and τ such that $\mathbf{P}\{G(\tau) = G\} > 0$ and $\mathbf{w}(G)$ is not terminal,

$$\mathbf{P}\{G(\tau+1) = G' | G(\tau) = G\} = \frac{1}{v(G)}, \tag{3.1}$$

$v(G)$ being the total number of nonisolated light vertices of G, if G' can be obtained from G by deletion of the edges incident to one of the light vertices of G; otherwise the conditional probability is zero.

Proposition 1. *(a) The sequence $\{\mathbf{w}(\tau)\}$ is also Markov: for every nonterminal $\mathbf{w}$ such that $\mathbf{P}\{\mathbf{w}(\tau) = \mathbf{w}\} > 0$*

$$\begin{aligned} p(\mathbf{w}'|\mathbf{w}) &:= \mathbf{P}\{\mathbf{w}(\tau+1) = \mathbf{w}' | \mathbf{w}(\tau) = \mathbf{w}\} \\ &= \frac{1}{v} \frac{h(\mathbf{w}')}{h(\mathbf{w})} \cdot v_0' \prod_{j=0}^{k} \binom{v_j' - \delta_{j0}}{u_{j+1}}, \quad (v_k' := \overline{v}'), \end{aligned} \tag{3.2}$$

where $\mathbf{u} = \{u_j\}_{1\le j\le k+1}$ *is the solution of the system*

$$\begin{aligned} v_j =& v_{j'} - u_{j+1} + u_j + \delta_{ij}, \quad 0 \le j \le k-1, \, (u_0 := -1), \\ \overline{v} =& \overline{v}' + u_k, \\ \sum_{j=1}^{k+1} u_j =& i := \mu - \mu', \end{aligned} \tag{3.3}$$

provided that $\mathbf{u} \ge \mathbf{0}$*. If* $\mathbf{u} \not\ge \mathbf{0}$ *then* $p(\mathbf{w}'|\mathbf{w}) = 0$*. (In a transition* $G \to G'$*, the parameters* u_j $(1 \le j \le k)$ *and* u_{k+1} *stand for the number of edges in* G *connecting the chosen light vertex with the vertices of degree* j*, and of degree* $> k$*, respectively.)*

(b) For every τ*, conditioned on* $\{\mathbf{w}(\nu)\}_{0\le\nu\le\tau}$*, the random graph* $G(\tau)$ *is distributed uniformly, that is for every* $\{\mathbf{w}^0(\nu)\}_{0\le\nu\le\tau}$ *such that* $\mathbf{P}\{\mathbf{w}(\nu) = \mathbf{w}^0(\nu);$ $0 \le \nu \le \tau\} > 0$*,*

$$\mathbf{P}\{G(\tau) = G|\mathbf{w}(\nu) = \mathbf{w}^0(\nu);\, 0 \le \nu \le \tau\} = \frac{1}{h(\mathbf{w}^0(\tau))}, \quad \forall G \in \mathcal{G}(\mathbf{w}). \tag{3.4}$$

Consequently, if a stopping time $\mathcal{T}$ *adapted to* $\{\mathbf{w}(\nu)\}_{\nu\ge0}$ *and* $\mathbf{w}$ *are such that* $\mathbf{P}\{\mathbf{w}(\mathcal{T}) = \mathbf{w}\} > 0$*, then*

$$\mathbf{P}\{G(\mathcal{T}) = G|\, \mathbf{w}(\mathcal{T}) = \mathbf{w}\} = \frac{1}{h(\mathbf{w})}, \quad \forall G \in \mathcal{G}(\mathbf{w}). \tag{3.5}$$

Proof of Proposition. Suppose that for some $\tau \ge 0$ the sequence $\{\mathbf{w}(\nu)\}_{0\le\nu\le\tau}$ is Markov, with one-step transition probabilities defined by (3.2), and the relation (3.4) holds. (This is definitely so for $\tau = 0$: basis of induction.)

Then for every sequence of nonterminal $\mathbf{w}^0(\nu) = (\mathbf{v}^0(\nu), \mu^0(\nu))$ $(0 \le \nu \le \tau+1)$ such that $\mathbf{P}\{\mathbf{w}(\nu) = \mathbf{w}^0(\nu),\, 0 \le \nu \le \tau\} > 0$ and every $G' \in \mathcal{G}(\mathbf{w}^0(\tau+1))$ we have

$$\begin{aligned} &\mathbf{P}\{G(\tau+1) = G'|\mathbf{w}(\nu) = \mathbf{w}^0(\nu),\, 0 \le \nu \le \tau) \\ =& \sum_{G\in\mathcal{G}(\mathbf{w}^0(\tau))} \mathbf{P}\{G(\tau+1) = G', G(\tau) = G|\mathbf{w}(\nu) = \mathbf{w}^0(\nu), 0 \le \nu \le \tau\} \\ =& \sum_{G\in\mathcal{G}(\mathbf{w}^0(\tau))} \Big(\mathbf{P}\{G(\tau+1) = G'|G(\tau) = G) \\ &\qquad\qquad \times \mathbf{P}\{G(\tau) = G|\mathbf{w}(\nu) = \mathbf{w}^0(\nu),\, 0 \le \nu \le \tau\}\Big) \\ =& \frac{1}{h(\mathbf{w}^0(\tau))} \sum_{G\in\mathcal{G}(\mathbf{w}^0(\tau))} \mathrm{P}\{G(\tau+1) = G'|G(\tau) = G\} \\ =& \frac{1}{h(\mathbf{w}^0(\tau))v^0(\tau)} N(G', \mathbf{w}^0(\tau)), \end{aligned} \tag{3.6}$$

where $N(G', \mathbf{w}^0(\tau))$ is the total number of graphs G from $\mathcal{G}(\mathbf{w}^0(\tau))$, with one nonisolated vertex being marked, such that G' is obtained from G by deleting all the edges incident to this light vertex in G. (In the derivation we have used consecutively the Markov property of $\{G(\nu)\}$, the induction hypothesis, and (3.1).)

It turns out that $N(G', \mathbf{w}^0(\tau))$ is the same for all $G' \in \mathcal{G}(\mathbf{w}^0(\tau+1))$. Here is why. Set for simplicity of notations $\mathbf{w} = \mathbf{w}^0(\tau)$, $\mathbf{w}' = \mathbf{w}^0(\tau+1)$. To get from $G' \in \mathcal{G}(\mathbf{w}')$ back to $G \in \mathcal{G}(\mathbf{w})$ we (1) pick one of the isolated vertices of G', and (2) insert some u_{j+1} edges between the chosen vertex and the set of the remaining $v'_j - \delta_{j0}$ light vertices of degree $0 \le j \le k-1$, and u_{k+1} edges joining the vertex to the set of $\overline{v}'$ heavy vertices of G'. So u_{j+1} vertices now have their degrees increased from j to $j+1$, $(1 \le j \le k)$. $\sum_{1 \le j \le k+1} u_j$, the degree of the chosen vertex in G, equals $\mu - \mu'$, the increase of the total number of edges in the backward transition $G' \to G$. For the given $(k+1)$-tuple $\{u_j\}$, the number of possibilities for the second step is $v'_0 \prod_{j=0}^k \binom{v'_j - \delta_{j0}}{u_{j+1}}$, with $v'_k := \overline{v}'$. As for the tuple $\{u_j\}$, it must satisfy (3.3) since the resulting graph G must be such that $\mathbf{w}(G) = \mathbf{w}$, and the selected vertex has to be one of its nonisolated light vertices. (Consider $1 \le j \le k-1$ for instance. Insertion of the new edges results in appearance of u_j vertices with new degree j, and some u_{j+1} vertices with old degree j have now degree $j+1$. Also, if $\sum_{s=1}^{k+1} u_s = i$ then the chosen vertex now has degree i.) Therefore

$$N(G', \mathbf{w}^0(\tau)) = f(\mathbf{w}^0(\tau), \mathbf{w}^0(\tau+1)),$$
$$f(\mathbf{w}, \mathbf{w}') := v'_0 \prod_{j=0}^{k} \binom{v'_j - \delta_{j0}}{u_{j+1}}, \quad (v'_k := \overline{v}').$$

Thus the conditional probability on the left hand side of (3.6) depends only on $\mathbf{w}^0(\tau)$ and $\mathbf{w}^0(\tau+1)$, and consequently

$$\begin{aligned} \mathbf{P}\{\mathbf{w}(\tau+1) = \mathbf{w}' | \mathbf{w}(\nu) = \mathbf{w}^0(\nu), 0 \le \nu \le \tau\} &= \mathbf{P}\{\mathbf{w}(\tau+1) = \mathbf{w}' | \mathbf{w}(\tau) = \mathbf{w}\} \\ &= \frac{1}{v} \frac{h(\mathbf{w}')}{h(\mathbf{w})} \cdot f(\mathbf{w}, \mathbf{w}'). \end{aligned} \tag{3.7}$$

So, using the induction hypothesis, $\{\mathbf{w}(\nu)\}_{0 \le \nu \le \tau+1}$ is Markov, with one-step transition probabilities $p(\mathbf{w}'|\mathbf{w})$, ($\mathbf{w}$ nonterminal), given by (3.2). Furthermore, if $\{\mathbf{w}^0(\nu)\}_{0 \le \nu \le \tau+1}$ is such that $\mathbf{P}\{\mathbf{w}(\nu) = \mathbf{w}^0(\nu);\, 0 \le \nu \le \tau+1\} > 0$ and $\mathbf{w}' := \mathbf{w}^0(\tau+1)$ is nonterminal, then (denoting $\mathbf{w} = \mathbf{w}^0(\tau)$) for every $G' \in \mathcal{G}(\mathbf{w}')$ according to (3.6) we have:

$$\begin{aligned} &\mathbf{P}\{G(\tau+1) = G' | \mathbf{w}(\nu) = \mathbf{w}^0(\nu);\, 0 \le \nu \le \tau+1\} \\ &\quad = \frac{\mathbf{P}\{G(\tau+1) = G' | \mathbf{w}(\nu) = \mathbf{w}^0(\nu);\, 0 \le \nu \le \tau\}}{\mathbf{P}\{\mathbf{w}(\tau+1) = \mathbf{w}' | \mathbf{w}(\nu) = \mathbf{w}^0(\nu);\, 0 \le \nu \le \tau\}} \\ &\quad = \frac{h(\mathbf{w}')^{-1} p(\mathbf{w}'|\mathbf{w})}{p(\mathbf{w}'|\mathbf{w})} \\ &\quad = \frac{1}{h(\mathbf{w}')}. \end{aligned}$$

The induction proof is complete.

The proof of (3.5) is straightforward, and we omit it. $\square$

Notes. 1. The relations (3.2), (3.4) involve $h(\mathbf{w}) = |\mathcal{G}(\mathbf{w})|$. In the next section, will be able to derive an asymptotic formula of $h(\mathbf{w})$ for the relevant $\mathbf{w}$s that is sufficiently sharp for our purposes.

2. Using (3.3), we can transform the formula for $f(\mathbf{w}, \mathbf{w}')$ into

$$f(\mathbf{w}, \mathbf{w}') = v_i \prod_{j=0}^{k} \frac{v_j'!}{v_j!} \cdot \frac{\left[\prod_{j=1}^{k-1} \binom{v_j - \delta_{ij}}{u_j}\right]}{\dfrac{\overline{v}!}{u_k! u_{k+1}! (\overline{v} - u_k - u_{k+1})!}}, \tag{3.8}$$

where $i = \mu - \mu' = \sum_{s=1}^{k+1} u_s$ and $v_k = \overline{v}$. This formula may look more complicated, but it will work just fine in our estimates. Without the factor $\prod_j v_j'!$, the last product arises naturally if one wants to compute $h(\mathbf{w}')f(\mathbf{w}, \mathbf{w}')$ (the number of transitions $G \to G'$ $(G \in \mathcal{G}(\mathbf{w}), G' \in \mathcal{G}(\mathbf{w}'))$, that is) in a forward fashion, via counting ways to *delete* edges in G.

4. Asymptotics for $h(\mathbf{w})$ and $p(\mathbf{w}'|\mathbf{w})$. Let $\mathbf{d} = (d_1, \dots, d_n)$ be a sequence of nonnegative integers with even sum, and $\hat{h}(\mathbf{d})$ be the total number of labelled graphs with degree sequence $\mathbf{d}$. Then for $\mathbf{w} = (\mathbf{v}, \mu)$ we have obviously

$$h(\mathbf{w}) = \frac{n!}{\prod_{j=0}^{k} v_j!} \sum_{\mathbf{d} \in \mathbf{D}} \hat{h}(\mathbf{d}), \tag{4.1}$$

where $v_k = n - \sum_{j=0}^{k-1} v_j = n - v_0 - v$, and $\mathbf{D} = \mathbf{D}(\mathbf{w})$ is the set of all nonnegative n-tuples $\mathbf{d}$ such that

$$\begin{aligned}
&d_1 = \cdots = d_{v_0} = 0,\\
&d_{v_0+1} = \cdots = d_{v_0+v_1} = 1,\\
&\dots\dots\dots\dots\dots\dots,\\
&d_{\sum_{j=0}^{k-2} v_j + 1} = \cdots = d_{\sum_{j=0}^{k-1} v_j} = k-1,\\
&d_{\sum_{j=0}^{k-1} v_j + 1}, \dots, d_n \ge k,\\
&\sum_{j=1}^{n} d_j = 2\mu.
\end{aligned} \tag{4.2}$$

For $\mathbf{D}$ to be nonempty, it is necessary that

$$\begin{aligned}
t &:= 2\mu - s \ge k\overline{v},\\
s &:= \sum_{j=1}^{k-1} j v_j.
\end{aligned} \tag{4.3}$$

(If $\mathbf{w} = \mathbf{w}(G)$ then s and t are the total degree of light vertices and the total degree of heavy vertices of G, respectively.) No tractable precise formula for $\hat{h}(\mathbf{d})$ is known, but it turns out to be possible to estimate the sum in (4.1) sharply for “likely” $\mathbf{w}$s using the following asymptotic formula due to McKay and Wormald [23].

For $r>0$, define $M_r=\sum_{1\le j\le n}[d_j]_r$ where $[x]_r=x(x-1)\cdots(x-r+1)$, (in particular, $M_1=2\mu=\sum_j d_j$), and $d_{\max}=\max_{1\le j\le n} d_j$. If $M_1\to\infty$ and $d_{\max}=o(M_1^{1/3})$ as $n\to\infty$ then

$$\begin{aligned}\hat{h}(\mathbf{d}) =& \frac{(M_1-1)!!}{\prod_{j=1}^n d_j!}\\ &\cdot \exp[-\frac{M_2}{2M_1}-\frac{M_2^2}{4M_1^2}-\frac{M_2^2M_3}{2M_1^4}+\frac{M_2^4}{4M_1^5}+\frac{M_3^2}{6M_1^3}+O(\frac{d_{\max}^3}{M_1})],\end{aligned}\tag{4.4}$$

where

$$(M_1-1)!!=1\cdot 3\cdots(M_1-1).$$

In the case of bounded degrees, $d_{\max}=O(1)$, the relation yields a formula

$$\hat{h}(\mathbf{d})=(1+o(1))\frac{(M_1-1)!!}{\prod_j d_j!}\exp(-\frac{M_2}{2M_1}-\frac{M_2^2}{4M_1^2})\tag{4.5}$$

obtained earlier by Bender and Canfield [1]. Notice that

$$\hat{h}(\mathbf{d})\le\frac{(M_1-1)!!}{\prod_{j=1}^n d_j!}$$

always. So the exponential factor in (4.4) is at most one. (In fact, Bollobás [2] rederived (4.5) by interpreting that factor as the probability that a certain random pairing on the set $\{1,\dots,M_1\}$ is graph-induced. McKay and Wormald also used the probabilistic approach, which they considerably strengthened by using switching operations on those pairings.)

Let us show —using (4.4)—that for $\mathbf{w}$s likely to appear in the deletion process the $\mathbf{d}$s that dominate in the sum (4.1) are such that (4.4) reduces to (4.5), with $o(1)=O(n^{-1+\epsilon}),\forall\epsilon\in(0,1)$.

Given $\mathbf{w}$ such that $h(\mathbf{w})>0$, introduce $G(\mathbf{w})$ the random graph distributed uniformly on $\mathcal{G}(\mathbf{w})$. For a fixed $b\in(0,1/3)$ define

$$\mathcal{H}_n=\mathcal{H}_n(b):=\{G: d_{\max}(G)\ge n^b \text{ or } \sum_{\{\text{heavy } j\}} d_j^4(G)\ge 2nE(Z^4(c))\},\tag{4.6}$$

(c comes from $m=cn/2$), and consider

$$g(\mathbf{w}):=\mathbf{P}\{G(\mathbf{w})\in\mathcal{H}_n\}.$$

Suppose $\mathbf{w}'$ is such that $p(\mathbf{w}'|\mathbf{w})>0$. Then $h(\mathbf{w}')>0$, too. Deletion of the edges incident to a randomly chosen light vertex of the random graph $G(\mathbf{w})$ produces a random subgraph G'. We know that $\mathbf{P}\{\mathbf{w}(G')=\mathbf{w}'\}=p(\mathbf{w}'|\mathbf{w})>0$, and so, conditioned on this event, $G'\overset{\mathcal{D}}{\equiv}G(\mathbf{w}')$. So, there is a probability space that accomodates both $G(\mathbf{w})$ and $G(\mathbf{w}')$ in such a way that $G(\mathbf{w}')\subset G(\mathbf{w})$. Since the property $\mathcal{H}_n$ is monotone increasing, we therefore obtain:

$$g(\mathbf{w})\ge g(\mathbf{w}'),\ \text{if } p(\mathbf{w}'|\mathbf{w})>0.\tag{4.7}$$

This means that the random sequence $\{g(\mathbf{w}(\tau))\}$ is (almost surely) nondecreasing.

Clearly,

$$\begin{aligned}\frac{1}{n}\mathbf{P}\{g(\mathbf{w}(0)) \ge \frac{1}{n}\} \le& \sum_{\mathbf{w}} g(\mathbf{w})\mathbf{P}\{\mathbf{w}(0) = \mathbf{w}\}\\ =& \sum_{\mathbf{w}} \mathbf{P}\{G(\mathbf{w}) \in \mathcal{H}_n\}\mathbf{P}\{w(G(n,m)) = \mathbf{w}\}\\ =& \mathbf{P}\{G(n,m) \in \mathcal{H}_n\}.\end{aligned}$$

Therefore

$$\begin{aligned}\mathbf{P}\{g(\mathbf{w}(0)) \ge \frac{1}{n}\} &\le n\mathbf{P}\{G(n,m) \in \mathcal{H}_n\}\\ &\le n(P_1 + P_2),\end{aligned} \tag{4.8}$$

where

$$\begin{aligned}P_1 &= \mathbf{P}\{d_{\max}(G(n,m)) \ge n^b\},\\ P_2 &= \mathbf{P}\left\{\sum_{\{\text{heavy } j\}} d_j^4(G(n,m)) \ge 2nE(Z^4(c))\right\}.\end{aligned}$$

To estimate the last probabilities, we use—in sequence—two conditioning techniques.

First of all, the graph process $\{G(n,\mu)\}_{0\le\mu\le m}$ can be viewed as the multigraph process $\{MG(n,\mu)\}_{0\le\mu\le m}$ conditioned on the event $A_n = \{MG(n,m)\text{has no loops and no multiple edges}\}$. (At each stage of the multigraph process, an edge is inserted between two vertices, i and j, drawn uniformly and independently of each other, and of the previous edges; if $i = j$ the multigraph gets a loop at i.) Therefore, using $\mathbf{P}\{U|V\} \le \mathbf{P}\{U\}/\mathbf{P}\{V\}$, and denoting by P_i' the analogous probabilities for $MG(n,m)$, we can write

$$P_i \le \frac{P_i'}{\mathbf{P}\{A_n\}} = O(P_i'), \tag{4.9}$$

since

$$\lim_{n\to\infty} \mathbf{P}\{A_n\} = \lim_{n\to\infty} \frac{\binom{\binom{n}{2}}{m} m! 2^m}{n^{2m}} = \exp(-c/2 - c^2/4) > 0.$$

(This reasoning repeats, in essence, an argument used originally by Chvátal [8] in a similar context. As a proof tool, the random multigraph had been used implicitly by Bollobás [2], and explicitly by Bollobás and Frieze [5].)

Secondly, $\mathbf{d}(MG(n,m)) := (d_1(MG(n,m)),\dots,d_n(G(n,M)))$ is clearly the random sequence of occupancy numbers in the classic allocation scheme "$2m$ distinguishable balls into n boxes". So, using the Poissonization device, we have

$$\mathbf{d}(MG(n,m)) \overset{\mathcal{D}}{\equiv} (Z_1,\dots,Z_n),$$

conditioned on $S_n := \sum_{j=1}^n Z_j = 2m$, where $Z_1,\dots,Z_n$ are independent copies of $Z(c)$, Poisson(c) distributed random variable. Since S_n is Poisson(nc), that is $Z(2m)$,

$$\begin{aligned}\mathbf{P}\{S_n = 2m\} &= \mathbf{P}\{Z(2m) = 2m\}\\ &= e^{-2m}(2m)^{2m}/(2m)! \ge \text{const } m^{1/2}.\end{aligned}$$

Therefore

$$P_i' = O(n^{1/2}P_i''), \tag{4.10}$$

where

$$P_1'' = \mathbf{P}\{\max_{1\le j\le n} Z_j \ge n^b\},$$
$$P_2'' = \mathbf{P}\{\sum_{j=1}^n Z_j^4 \ge 2nE(Z^4(c))\}.$$

By Chernoff's inequality,

$$P_2'' = O(e^{-\alpha(c)n}), \quad \alpha(c) > 0, \tag{4.11}$$

and

$$\begin{aligned} P_1'' &\le n\mathbf{P}\{Z(c) \ge n^b\} \\ &= n\sum_{r\ge n^b} e^{-c}c^r/r! \\ &= O(\exp\{-bn^b \log n/2\}). \end{aligned} \tag{4.12}$$

Combining the estimates (4.8)-(4.12), we obtain then

$$\mathbf{P}\{g(\mathbf{w}(0)) \ge \frac{1}{n}\} = O(e^{-n^b}). \tag{4.13}$$

Thus, see (4.6), (4.7), with probability $\ge 1 - O(e^{-n^b})$, the $\mathbf{w}$s encountered in the deletion process are such that

$$g(\mathbf{w}) \le \frac{1}{n}. \tag{4.14}$$

In other words,

$$h(\mathbf{w}) = (1 + O(\frac{1}{n}))h_1(\mathbf{w}),$$
$$h_1(\mathbf{w}) := |\{G \in \mathcal{G}(\mathbf{w}) : \mathbf{d}(G) \in \mathbf{D}_1\}|.$$

Here the set $\mathbf{D}_1 \subset \mathbf{D}$ is specified by the additional restrictions

$$\begin{aligned} &d_{\max} \le n^b, \\ &\sum_{\{\text{heavy } j\}} d_j^4 \le dn, \quad d := 2\mathrm{E}(Z^4(c)). \end{aligned} \tag{4.15}$$

Hence we may concentrate on the asymptotic behavior of $h_1(\mathbf{w})$. Now, from the McKay-Wormald formula (4.4) and (4.15) it follows that

$$\begin{aligned} h_1(\mathbf{w}) &= (1 + O(n^{-1+3b}))\frac{n!}{\prod_{j=0}^k v_j!}\sum_{\mathbf{d}\in\mathbf{D}_1} \hat{h}_1(\mathbf{d}), \\ \hat{h}_1(\mathbf{d}) &:= \frac{(M_1-1)!!}{\prod_{j=1}^n d_j!}\exp(-\lambda/2 - \lambda^2/4), \\ \lambda &:= \frac{M_2}{M_1}, \end{aligned} \tag{4.16}$$

if—in addition to requiring $g(\mathbf{w}) \le n^{-1}$—we restrict ourselves to $\mathbf{w}$s such that

$$M_1(= 2\mu(\mathbf{w})) \ge an.$$

Here $a > 0$ is fixed, and arbitrarily small. We can afford this restriction since in our subsequent proofs we will not be concerned about the moments t when $\mu(t) = o(n)$. In fact, we go even further and concentrate on $\mathbf{w}$s such that

$$\begin{aligned} &\overline{v} \ge an, \\ &t \ge (k+a)\overline{v}. \end{aligned} \tag{4.17}$$

(compare with (4.3)). The double-conditioning technique quickly reveals that the starting graph $G(n,m)$ meets the conditions (4.17) with exponentially high probability, if

$$\begin{aligned} &a < \mathbf{P}\{Z(c) \ge k\}, \\ &\sum_{j \ge k} j\mathbf{P}\{Z(c) = j\} \ge (k+a)\mathbf{P}\{Z(c) \ge k\}. \end{aligned} \tag{4.18}$$

To obtain a sharp estimate of the sum in (4.16), we use a version of a close derivation done by Pittel and Woyczynski [28]. A key idea is that—just as $G(n,m)$—the degrees of heavy vertices of the random graph $G(\mathbf{w})$ must jointly behave like independent Poissons subject to the total sum condition, each bounded below by k, (c.f. Karp and Sipser [16]).

Introduce a family of $\overline{v}(:= n - v_0 - \sum_{1\le j\le k-1} v_j)$ independent random variables $Y_1, \dots Y_{\overline{v}}$, each being distributed as Poisson$(Z(z))$, conditioned on $\{Z(z) \ge k\}$. Explicitly,

$$\begin{aligned} \mathbf{P}\{Y_j = r\} &= \frac{\mathbf{P}\{Z(z) = r\}}{p_k(z)}, \quad r \ge k, \\ p_k(z) &:= \mathbf{P}\{Z(z) \ge k\}. \end{aligned} \tag{4.19}$$

The parameter $z > 0$ is chosen such that

$$\overline{v}\mathrm{E}(Y) = t. \tag{4.20}$$

Such $z = z(\mathbf{w})$ exists and bounded away from both 0 and ∞, uniformly for all $\mathbf{w}$s satisfying (4.17). ($z(\mathbf{w})$ is unique since $E(Y_1) = ze_k'(z)/e_k(z)$ is strictly increasing; see (5.10).) Using $Y_1, \dots, Y_{\overline{v}}$, we can write (see (4.16)):

$$\begin{aligned} \sum_{\mathbf{d}\in\mathbf{D}_1} \hat{h}_1(\mathbf{d}) = {}& \frac{(M_1-1)!!}{\prod_{j=1}^{k-1}(j!)^{v_j}} \frac{(e_k(z))^{\overline{v}}}{z^t} \\ &\cdot \mathrm{E}[\exp(-\lambda/2 - \lambda^2/4); \sum_{\ell} Y_\ell = t, \mathbf{Y} \in \Upsilon], \\ e_k(z) := {}& \sum_{r\ge k} z^r/r!, \end{aligned} \tag{4.21}$$

where

$$\lambda = \frac{\sum_{j=1}^{k-1} j(j-1)v_j + \sum_{\ell=1}^{\overline{v}} Y_\ell(Y_\ell - 1)}{\sum_{j=1}^{k-1} jv_j + \sum_{\ell=1}^{\overline{v}} Y_\ell},$$

$$\Upsilon := \{\mathbf{y} = (y_1, \dots, y_{\overline{v}}) : \max_\ell y_\ell \le n^b, \sum_\ell y_\ell^4 \le dn\},$$

and we use the notation $\mathrm{E}[U; A] = \mathrm{E}[U\mathbf{1}_A]$. It suffices to estimate sharply $\mathrm{E}[\exp(-\lambda/2-\lambda^2/4); \sum_\ell Y_\ell = t]$, since (cf. (4.11), (4.12))

$$\mathbf{P}\{\mathbf{Y} \notin \Upsilon\} = O(e^{-n^b}). \tag{4.22}$$

Furthermore, the distribution of λ is sharply concentrated around

$$\overline{\lambda} := \frac{\sum_{j=1}^{k-1} j(j-1)v_j + \overline{v}\mathrm{E}(Y_1(Y_1-1))}{\sum_{j=1}^{k-1} jv_j + \overline{v}\mathrm{E}(Y_1)}. \tag{4.23}$$

Indeed, using a large deviation theorem for the sums of i.i.d. random variables, due to Cramér (see Feller [11], Ch. XVI, for instance), we have: uniformly for $\mathbf{w}$ satisfying (4.17),

$$\mathbf{P}\{|\sum_{\ell=1}^{\overline{v}} Y_\ell - \overline{v}\mathrm{E}(Y_1)| \ge \log n\sqrt{n}\} = O[\exp(-\gamma\log^2 n)],$$

$$\mathbf{P}\{|\sum_{\ell=1}^{\overline{v}} Y_\ell(Y_\ell-1) - \overline{v}\mathrm{E}(Y_1(Y_1-1))| \ge \log n\sqrt{n}\} = O[\exp(-\gamma\log^2 n)],$$

for some $\gamma = \gamma(a) > 0$. Thus

$$\lambda - \overline{\lambda} = O(\frac{\log n}{\sqrt{n}})$$

with probability $\ge 1 - \exp(-\gamma\log^2 n)$. Consequently (see also (4.23)) the expectation in (4.21) equals

$$[1 + O(\frac{\log n}{\sqrt{n}})]\exp(-\overline{\lambda}/2 - \overline{\lambda}^2/4)$$

$$\cdot\,\mathbf{P}\{\sum_{\ell=1}^{\overline{v}} Y_\ell = t\} + O[\exp(-\gamma\log^2 n)]. \tag{4.24}$$

Here, by (4.20) and a local limit theorem for the sum of lattice-type i.i.d. random variables ([11], Ch. XVI)

$$\mathbf{P}\{\sum_{\ell=1}^{\overline{v}} Y_\ell = t\} = \frac{1}{\sqrt{\overline{v}\cdot 2\pi\mathrm{Var}(Y_1)}}[1 + O(\frac{1}{\overline{v}})], \tag{4.25}$$

uniformly for $\mathbf{w}$s subject to (4.17). (To be sure, the quoted limit theorem is proved under the only condition that $\mathrm{Var}(Y_1) < \infty$, with the remainder term being simply

$o(1)$. However, an easy refinement of the argument establishes (4.25) under a stronger condition $\mathrm{E}(Y_1^4) < \infty$. This condition obviously holds in our case, and moreover—under (4.17)— $\mathrm{E}(Y_1^4) \le \gamma_2(a) < \infty$, $0 < \gamma_3(a) \le \mathrm{Var}(Y_1) \le \gamma_4(a) < \infty$, which leads to the uniformity of (4.25) for those $\mathbf{w}$s.)

Putting together the relations (4.16), (4.21), (4.24) and (4.25) we obtain the following result.

Proposition 2. *Uniformly for* $\mathbf{w}$ *such that* $h(\mathbf{w}) > 0$ *and the conditions* (4.14),(4.17) *are met,*

$$\begin{aligned} h(\mathbf{w}) = [1 + O(n^{-1+3b} + n^{-1/2}\log n)] \cdot \frac{n!(M_1-1)!!}{\overline{v}!\prod_{j=1}^{k-1}(j!)^{v_j} v_j!} \\ \cdot \frac{(e_k(z))^{\overline{v}}}{z^t} \exp(-\overline{\lambda}/2 - \overline{\lambda}^2/4) \\ \cdot \frac{1}{\sqrt{\overline{v} 2\pi \mathrm{Var}(Y_1)}}; \end{aligned} \tag{4.26}$$

here Y_1, z, $p_k(z)$, *and* $\overline{\lambda}$ *are defined by* (4.19),(4.20) *and* (4.23).

Corollary 1. *Suppose that* $\mathbf{w}$ *is nonterminal, and* $h(\mathbf{w}) > 0$. *If* $\mathbf{w}$ *satisfies the conditions* (4.14), (4.17) *and* $\mathbf{w}'$ *is such that* $\mathbf{u} = \mathbf{u}(\mathbf{w}, \mathbf{w}') \ge \mathbf{0}$ *then*

$$\begin{aligned} p(\mathbf{w}'|\mathbf{w}) =& [1 + O(n^{-1+3b} + n^{-1/2}\log n)] \cdot \left[1 + O\left(\sum_{j=1}^{k-1} \frac{(u_j-1)^+}{v_j+1}\right)\right] \\ & \cdot q(\mathbf{w}'|\mathbf{w}), \end{aligned} \tag{4.27}$$

where

$$q(\mathbf{w}'|\mathbf{w}) := \begin{cases} \dfrac{v_i}{v}\mathbf{P}\{\text{Multin } (i; \overline{\mathbf{p}}) = \mathbf{u}\}, & \text{if } v_i > 0, \\ 0, & \text{if } v_i = 0. \end{cases} \tag{4.28}$$

Here $\mathbf{u} = \{u_j\}_{1 \le j \le k+1}$ *is the solution of* (3.3). Multin $(i; \overline{\mathbf{p}})$ *stands for the multinomially distributed random vector* $\mathbf{X} = \{X_1, \dots, X_{k+1}\}$, *with parameter (number of trials) equal* $i = \mu - \mu'$, *and the probability vector* $\mathbf{p} = \mathbf{p}(\mathbf{w}) = \{p_1, \dots, p_{k+1}\}$ *of* $k+1$ *possible outcomes in each trial given by*

$$\begin{aligned} p_j &= \frac{j(v_j - \delta_{ij})}{2\mu - i}, \quad 1 \le j \le k-1, \\ p_k &= \frac{z^k \overline{v}}{(2\mu - i)(k-1)! e_k(z)}, \\ p_{k+1} &= \frac{z\overline{v}}{2\mu - i}; \end{aligned} \tag{4.29}$$

z *is the root of the equation*(4.20).

Note. According to (4.19), (4.20),

$$\frac{\overline{v}z^k}{(k-1)!e_k(z)}+\overline{v}z=\overline{v}z\left[\frac{\frac{z^{k-1}}{(k-1)!}+e_k(z)}{e_k(z)}\right]$$
$$=\overline{v}\frac{ze_k'(z)}{e_k(z)}=\overline{v}E(Y_1)=t,$$

and

$$\sum_{i=1}^{k-1} iv_i+t=s+t=2\mu. \qquad (!)$$

Therefore

$$\sum_{j=1}^{k+1} p_j=\frac{1}{2\mu-i}(s-i+t)=1.$$

Also, for $v_i>0$, we have $p_j\ge 0\,(1\le j\le k+1)$; so $\{p_j\}$ is indeed a probability distribution.

Proof of Corollary 1. Let $Y_1',z',\overline{\lambda}'$ be for $\mathbf{w}'$ what $Y,z,\overline{\lambda}$ are for $\mathbf{w}$. Since $\|\mathbf{w}-\mathbf{w}'\|=O(1)$ as $n\to\infty$, uniformly for all $\mathbf{w},\mathbf{w}'$ related via (3.3), it follows from (4.17) that

$$|z'-z|=O(\frac{1}{n}),$$

hence

$$\mathrm{Var}(Y_1')=[1+O(\frac{1}{n})]\mathrm{Var}(Y_1),$$
$$\overline{\lambda}'=[1+O(\frac{1}{n})]\overline{\lambda}.$$

Next, introduce

$$f_{\overline{v}t}(y)=\overline{v}\log e_k(y)-t\log y,\quad y>0,$$

so that

$$\frac{(e_k(y))^{\overline{v}}}{y^t}=\exp[f_{\overline{v}t}(y)].$$

By (4.19) and (4.20),

$$\left.\frac{d}{dy}f_{\overline{v}t}(y)\right|_{y=z}=\overline{v}\frac{e_k'(z)}{e_k(z)}-\frac{t}{z}$$
$$=\frac{1}{z}\left[\overline{v}\frac{ze_k'(z)}{e_k(z)}-t\right]$$
$$=\frac{1}{z}[\overline{v}\mathrm{E}(Y_1)-t]=0,$$

so that z is a stationary point of $f_{\overline{v}t}(y)$. (It can be easily proved that $f_{\overline{v}t}(z)=\min\{f_{\overline{v}t}(y):\,y>0\}$.) Consequently,

$$f_{\overline{v}t}(z')-f_{\overline{v}t}(z)=\frac{1}{2}f_{\overline{v}t}(\tilde{z})(z'-z)^2$$
$$(\tilde{z}\text{ is between }z'\text{ and }z)$$
$$=O(n\frac{1}{n^2})=O(\frac{1}{n}).$$

Therefore

$$\begin{aligned} f_{\overline{v}'t'}(z') - f_{\overline{v}t}(z) = & f_{\overline{v}t}(z') - f_{\overline{v}t}(z) \\ & + (\overline{v}' - \overline{v}) \log e_k(z) - (t - t') \log z \\ & + O\left[\left| \log \frac{e_k(z')}{e_k(z)} \right| + \left| \log \frac{z'}{z} \right| \right] \\ & = (\overline{v}' - \overline{v}) \log e_k(z) - (t - t') \log z + O(\frac{1}{n}). \end{aligned}$$

Here (see (3.3))

$$\overline{v} - \overline{v}' = u_k, \quad t - t' = ku_k + u_{k+1}.$$

(About $t-t'$: in the transition $\mathbf{w} \to \mathbf{w}'$, the total degree of heavy vertices decreases by ku_k due to u_k vertices of degree k becoming light, of degree $k-1$; an additional decrease by u_{k+1} is due to some u_{k+1} vertices of degree $> k$ that remain heavy, but each now of degree smaller by one than before.) Thus

$$\begin{aligned} \frac{(e_k(z'))^{\overline{v}'}}{(z')^{t'}} & \div \frac{(e_k(z))^{\overline{v}}}{z^t} \\ & = [1 + O(\frac{1}{n})] \cdot \left(\frac{z^k}{e_k(z)} \right)^{u_k} z^{u_{k+1}}. \end{aligned} \tag{4.30}$$

Further, again using (3.2)

$$\begin{aligned} \frac{\overline{v}! \prod_{j=0}^{k-1} (j!)^{v_j} v_j!}{\overline{v}'! \prod_{j=0}^{k-1} (j!)^{v_j'} v_j'!} = & \left(\prod_{j=0}^{k} \frac{v_j!}{v_j'!} \right) \cdot \left(\prod_{j=0}^{k-1} (j!)^{-u_{j+1}+u_j+\delta_{ij}} \right) \\ & (v_k := \overline{v},\ v_k' := \overline{v}') \\ & = i! \left(\prod_{j=0}^{k} \frac{v_j!}{v_j'!} \right) \cdot \left(\prod_{j=1}^{k-1} j^{u_j} \right) \cdot \left[\frac{1}{(k-1)!} \right]^{u_k}; \end{aligned} \tag{4.31}$$

($u_0 = -1$, as we recall). Finally (see (4.17)),

$$\frac{(2\mu' - 1)!!}{(2\mu - 1)!!} = [1 + O(n^{-1})](2\mu)^{-i} = [1 + O(n^{-1})](2\mu - i)^{-i}. \tag{4.32}$$

Combining Propositions 1,2, the relations (3.8) and (4.30)-(4.32), and using (4.17), we arrive at

$$\begin{aligned} p(\mathbf{w}'|\mathbf{w}) = & \frac{h(\mathbf{w}') f(\mathbf{w}, \mathbf{w}')}{v h(\mathbf{w})} \\ = & [1 + O(n^{-1+3b} + n^{-1/2} \log n)] \cdot \tilde{p}(\mathbf{w}'|\mathbf{w}), \\ \tilde{p}(\mathbf{w}'|\mathbf{w}) := & \frac{v_i i!}{v(2\mu)^i} \prod_{j=1}^{k-1} \binom{v_j - \delta_{ij}}{u_j} \\ & \cdot \frac{1}{u_k!} \left[\frac{z^k \overline{v}}{(k-1)! e_k(z)} \right]^{u_k} \cdot \frac{1}{u_{k+1}!} (z\overline{v})^{u_{k+1}}. \end{aligned} \tag{4.33}$$

Here, since $u_j \le k-1$,

$$\binom{v_j - \delta_{ij}}{u_j} = \frac{(v_j-\delta_{ij})^{u_j}}{u_j!}\left[1 + O\left(\frac{(u_j-1)^+}{v_j+1}\right)\right].$$

Therefore (see the notations (4.29)) the relations (4.27), (4.28) follow. $\square$

Thus $q(\mathbf{w}'|\mathbf{w})$ can be viewed as an one-step transition probability of a Markov chain that evolves on the set of $\mathbf{w}$s defined in (4.14),(4.17), till the moment the process exits this set.

5. Approximate dynamics of $\{E[\mathbf{w}(\tau)]\}$. Let us look carefully at this limiting Markov chain. According to (3.3) and (4.26), (4.27), for the transition probabilities $q(\mathbf{w}'|\mathbf{w})$ and $0 \le j \le k-1$ we have:

$$\begin{aligned} \mathrm{E}_q[v_j(\tau+1)|\mathbf{w}(\tau)=\mathbf{w}] &= \sum_{\mathbf{w}'} v_j' q(\mathbf{w}'|\mathbf{w}) \\ &= \sum_{1\le i\le k-1} \frac{v_i}{v}\mathrm{E}(v_j + X_{j+1} - X_j - \delta_{ij}), \end{aligned} \tag{5.1}$$

($X_0 := -1$). Since $\mathrm{E}(X_j) = ip_j$, $1 \le j \le k+1$, we obtain then

$$\mathrm{E}_q[v_j(\tau+1)|\mathbf{w}(\tau)=\mathbf{w}] = v_j + f_j(\mathbf{w}(\tau)), \quad 0 \le j \le k-1, \tag{5.2}$$

where

$$f_j(\mathbf{w}) = \begin{cases} 1 + \dfrac{v_1 s}{2\mu v}, & \text{if } j=0, \\ \dfrac{(j+1)v_{j+1}s}{2\mu v} - \dfrac{jv_j s}{2\mu v} - \dfrac{v_j}{v}, & \text{if } 1 \le j \le k-2, \\ \dfrac{z^k \overline{v} s}{2\mu v (k-1)! e_k(z)} - \dfrac{(k-1)v_{k-1}s}{2\mu v} - \dfrac{v_{k-1}}{v}, & \text{if } j = k-1. \end{cases} \tag{5.3}$$

(Recall that $s := \sum_{1\le i\le k-1} iv_i$.)Analogously,

$$\begin{aligned} \mathrm{E}_q[\mu(\tau+1)|\mathbf{w}(\tau)=\mathbf{w}] &= \sum_{\mathbf{w}'} \mu' q(\mathbf{w}'|\mathbf{w}) \\ &= \mu - \sum_{1\le i\le k-1} i\frac{v_i}{v} = -\frac{s}{v}. \end{aligned} \tag{5.4}$$

As long as $\mathbf{w}(\tau)$ meets the condition (4.17), the random variables $f_j(\mathbf{w}(\tau))$, $-s(\tau)/v(\tau)$ are all only of order $O(1)$. This makes us expect—though does not actually prove– that with high probability the sample sequence $\{\mathbf{w}(\tau)\}$ must be close to the solution $\tilde{\mathbf{w}}(\tau) = (\tilde{\mathbf{v}}(\tau), \tilde{\mu}(\tau))$ of

$$\begin{aligned} \frac{dv_j(\tau)}{d\tau} &= f_j(\mathbf{w}(\tau)), \\ \frac{d\mu(\tau)}{d\tau} &= -\frac{s(\tau)}{v(\tau)}, \end{aligned} \tag{5.5}$$

subject to the (random) initial conditions

$$\tilde{v}_j(0) = v_j(0),\ (0 \le j \le k-1), \quad \tilde{\mu}(0) = \mu(0). \tag{5.6}$$

At any rate, it is clear that the more we know about this system of ordinary differential equations, the better are our chances (no pun intended) for probabilistic analysis of the random sequence $\{\mathbf{w}(\tau)\}$ itself.

As it turns out, that system has two remarkably simple integrals; namely

$$\frac{z^2}{\mu} \equiv \text{const}, \tag{5.7}$$

$$\frac{\overline{v}}{p_k(z)} \equiv \text{const}. \tag{5.8}$$

(We recall that $ze'_k z/e_k(z) = t/\overline{v}$, $t = 2\mu - s$, $p_k(z) = e^{-z}e_k(z) = \mathbf{P}\{Z(z) \ge k\}$.) Especially surprising is (5.7) since it connects $\mu(\tau)$ the current number of edges and $z(\tau)$ the "hidden"parameter chosen so the Poisson($Z(z)$) conditioned on $\{Z(z) \ge k\}$ has the expected value equal $t(\tau)/\overline{v}(\tau)$ the average degree of a heavy vertex in the current graph. Notice that (5.7) has the same form for all k. We should emphasize though that these are merely the integrals of the *approximate* equations for means $\mathrm{E}[\mathbf{w}(\tau)]$.

Let us prove (5.7). We observe first that (see (4.18)) for every x

$$\mathrm{E}(x^{Y_1}) = \frac{e_k(xz)}{e_k(z)},$$

so that differentiating both sides of this identity twice at $x = 1$ we get

$$\mathrm{E}[Y_1(Y_1 - 1)] = \frac{z^2 e''_k(z)}{e_k(z)}, \tag{5.9}$$

in addition to $\mathrm{E}(Y_1) = ze'_k(z)/e_k(z)$. Therefore

$$\frac{d}{dz}\frac{ze'_k(z)}{e_k(z)} = \frac{1}{z}[\mathrm{E}(Y_1(Y_1-1)) + \mathrm{E}(Y_1) - \mathrm{E}^2(Y_1)] = \frac{1}{z}\mathrm{Var}(Y_1) > 0. \tag{5.10}$$

On the other hand, using the equations (5.5) and $\overline{v} = n - v_0 - v$, we compute

$$\begin{aligned}
\frac{d}{d\tau}\frac{ze_k'(z)}{e_k(z)} &= (\frac{t}{\overline{v}})'_\tau = (\frac{2\mu - s}{\overline{v}})'_\tau \\
&= \frac{1}{\overline{v}}(2\mu' - s') - \frac{\overline{v}'}{\overline{v}^2}(2\mu - s) \\
&= \frac{1}{\overline{v}}\big[-\frac{2s}{\overline{v}} - \sum_{j=1}^{k-2} j\left(\frac{(j+1)v_{j+1}s}{2\mu v} - \frac{jv_j s}{2\mu v} - \frac{v_j}{v}\right) \\
&- (k-1)\left(\frac{z^k\overline{v}s}{2\mu v(k-1)!e_k(z)} - \frac{(k-1)v_{k-1}s}{2\mu v} - \frac{v_{k-1}}{v}\right)\big] \\
&- \frac{2\mu - s}{\overline{v}^2}\left[-\frac{z^k\overline{v}s}{2\mu v(k-1)!e_k(z)}\right] \\
&= \frac{1}{\overline{v}}\big[-\frac{2s}{\overline{v}} + \sum_{j=1}^{k-1}\left(\frac{jv_j s}{2\mu v} + \frac{jv_j}{v}\right) - \frac{(k-1)z^k\overline{v}s}{2\mu v(k-1)!e_k(z)}\big] \\
&+ \frac{s}{2\mu v\overline{v}}(2\mu - s)\frac{z^k}{(k-1)!e_k(z)} \\
&= \frac{1}{\overline{v}}\left[-\frac{s}{v} + \frac{s}{v}\cdot\frac{s}{2\mu} - \frac{(k-1)z^k\overline{v}s}{2\mu v(k-1)!e_k(z)}\right] \\
&+ \frac{s}{2\mu v\overline{v}}(2\mu - s)\frac{z^k}{(k-1)!e_k(z)} \\
&\qquad (\text{using } \frac{2\mu - s}{\overline{v}} = \frac{ze_k'(z)}{e_k(z)}) \\
&= \frac{s}{2\mu v}\left[-\frac{ze_k'(z)}{e_k(z)} - \frac{z^2(z^{k-2}/(k-2)!)}{e_k(z)} + \frac{z^2 e_k'(z)(z^{k-1}/(k-1)!)}{e_k^2(z)}\right] \\
&= \frac{s}{2\mu v}\left[-\frac{ze_k'(z)}{e_k(z)} - \frac{z^2 e_k''}{e_k(z)} + \left(\frac{ze_k'(z)}{e_k(z)}\right)^2\right].
\end{aligned}$$

So, invoking (5.9) and (5.10),

$$\frac{d}{d\tau}\frac{ze_k'(z)}{e_k(z)} = -\frac{sz}{2\mu v}\frac{d}{dz}\frac{ze_k'(z)}{e_k(z)},$$

that is

$$\begin{aligned}
\frac{dz}{d\tau} &= -\frac{sz}{2\mu v} \\
&= \frac{z}{2\mu}\frac{d\mu}{d\tau},
\end{aligned} \tag{5.11}$$

(see (5.5)). Therefore

$$\frac{dz}{d\mu} = \frac{z}{2\mu},$$

and (5.7) follows.

Next,

$$\begin{aligned}\frac{d\overline{v}}{d\tau} &= -\sum_{j=0}^{k-1}\frac{dv_j(\tau)}{d\tau}\\ &= -\frac{z^k\overline{v}s}{2\mu v(k-1)!e_k(z)},\end{aligned}$$

and using (5.11), we have

$$\begin{aligned}\frac{d\overline{v}}{dz} &= \overline{v}\frac{z^{k-1}/(k-1)!}{e_k(z)}\\ &= \overline{v}\frac{dp_k(z)/dz}{p_k(z)}, \qquad (p_k(z) = e^{-z}e_k(z)).\end{aligned}$$

This yields (5.8).

Let us also compute $E_q[s(\tau+1)|\mathbf{w}(\tau)=\mathbf{w}]$. First observe that

$$[s(\tau+1)-s(\tau)]+[t(\tau+1)-t(\tau)] = 2\mu(\tau+1)-2\mu(\tau) = -2i,$$

the last equality holding with (conditional) probability v_i/v, and

$$t(\tau+1)-t(\tau) = -kX_k - X_{k+1}.$$

So (see (4.28)),

$$\begin{aligned}E_q[s(\tau+1)|\mathbf{w}(\tau)=\mathbf{w}] =& s - 2\sum_{i=1}^{k-1}\frac{iv_i}{v}\\ &+ k\left(\sum_{i=1}^{k-1}\frac{iv_i}{v}\,\frac{z^k\overline{v}}{(2\mu-i)(k-1)!e_k(z)}\right) + \left(\sum_{i=1}^{k-1}\frac{iv_i}{v}\,\frac{z\overline{v}}{2\mu-i}\right)\\ =& s - \frac{2s}{v} + [1+O(n^{-1})]\frac{s\overline{v}z}{2\mu v}\left[\frac{k(z^{k-1}/(k-1)!)}{e_k(z)}\right]\\ &+ [1+O(n^{-1})]\frac{s\overline{v}z}{2\mu v}.\end{aligned}$$

Here

$$\begin{aligned}\frac{k(z^{k-1}/(k-1)!)}{e_k(z)}+1 =& \frac{k(z^{k-1}/(k-1)!)}{e_k(z)} + \frac{e_k'(z)}{e_k(z)} - \frac{z^{k-1}/(k-1)!}{e_k(z)}\\ =& \frac{z^{k-1}/(k-2)!}{e_k(z)} + \frac{e_k'(z)}{e_k(z)}\\ & \qquad\qquad (\frac{ze_k'(z)}{e_k(z)} = \frac{2\mu-s}{\overline{v}})\\ =& \frac{1}{e^{-z}e_k(z)}\left[\frac{e^{-z}z^{k-1}}{(k-2)!} - e^{-z}e_k'(z)\right] + 2\frac{2\mu-s}{z\overline{v}}\\ =& -\frac{\pi_k^2(z)}{p_k(z)}\left(\frac{z}{\pi_k(z)}\right)' + 2\frac{2\mu-s}{z\overline{v}}.\end{aligned}$$

(Recall that $p_k(z) := P\{Z(z) \ge k\} = e^{-z}e_k(z)$, $\pi_k(z) := P\{Z(z) \ge k-1\} = e^{-z}e_k'(z)$.) Combining the two relations, and using $z\pi_k(z)/p_k(z) = ze_k'(z)/e_k(z) = t/\overline{v}$, we write

$$E_q[s(\tau+1)|\mathbf{w}(\tau)=\mathbf{w}] = s - \frac{s^2}{\mu v} - \frac{st\pi_k(z)}{2\mu v}\left(\frac{z}{\pi_k(z)}\right)' + O(n^{-1}), \tag{5.12}$$

uniformly for $\mathbf{w}$ in question. Like (5.2)-(5.4), the relation (5.12) motivates us to consider a differential equation

$$\frac{ds}{d\tau} = -\frac{s^2}{\mu v} - \frac{st\pi_k(z)}{2\mu v}\left(\frac{z}{\pi_k(z)}\right)'. \tag{5.13}$$

The equations (5.5), (5.13) will play a critical role in the next (last) section. It is easy to determine s (equivalently, t) as a function of z, without having to integrate the equation (5.13). Indeed,

$$\begin{aligned}\frac{t}{2\mu}\frac{z}{\pi_k(z)} &= \frac{z^2}{2\mu}\frac{t}{p_k(z)}\frac{p_k(z)}{z\pi_k(z)}\\ &= \frac{z^2}{2\mu}\frac{t}{p_k(z)}\frac{\overline{v}}{t}\\ &= \frac{z^2}{2\mu}\frac{\overline{v}}{p_k(z)} \equiv \text{ const},\end{aligned}$$

see (5.7), (5.8). Below we will be using the notations

$$J_1(\mathbf{w}) = \frac{nz^2}{\mu}, \quad J_2(\mathbf{w}) = \frac{\overline{v}}{np_k(z)}, \quad J_3(\mathbf{w}) = \frac{t}{2\mu}\frac{z}{\pi_k(z)}. \tag{5.14}$$

6. Proofs of the main results.

Given $a > 0$, define the set $\mathbf{W} = \mathbf{W}(a)$ by

$$\mathbf{W}(a) := \left\{\mathbf{w} : h(\mathbf{w}) > 0,\, g(\mathbf{w}) \le \frac{1}{n},\, \overline{v} \ge an,\, t \ge (k+a)\overline{v}\right\}, \tag{6.1}$$

and the stopping (exit) time $\mathcal{T} = \mathcal{T}(a)$ by

$$\mathcal{T}(a) = \begin{cases}\min\{\tau < T : \mathbf{w}(\tau) \notin \mathbf{W}(a)\}, & \text{if such } \tau \text{ exist},\\ T, & \text{otherwise.}\end{cases} \tag{6.2}$$

(T is the total number of deletion steps.)

Lemma 1. *Conditioned on* $\{\mathbf{w}(0) \in \mathbf{W}(a)\} \cap \{\mu(\mathbf{w}(0)) = cn/2\}$, *for* $0 < \alpha < \min\{1/2, 1-3b\}$,

$$\mathbf{P}\left\{\max_{\tau \le \mathcal{T}}\left|\frac{J_i(\mathbf{w}(\tau))}{J_i(\mathbf{w}(0))} - 1\right| > x\right\} = O[e^{-xn^{\alpha}}], \quad i = 1, 2, 3, \tag{6.3}$$

uniformly for $x > 0$.

Proof of Lemma 1. Consider $i = 1$, for instance. Introduce the function

$$Q(\mathbf{w}) := \exp\{n^{\alpha}[J_1(\mathbf{w}) - J_1(\mathbf{w}(0))]\}.$$

Let us evaluate

$$\Sigma := \sum_{\mathbf{w}'} Q(\mathbf{w}')p(\mathbf{w}'|\mathbf{w}), \quad \mathbf{w} \in \mathbf{W} = \mathbf{W}(a).$$

By the definition of $\mathbf{W}(a)$, and using (4.7), we see that—for n large enough—$\mathbf{w}' \in \mathbf{W}(a/2)$ whenever $\mathbf{w} \in \mathbf{W}(a)$ and $p(\mathbf{w}'|\mathbf{w}) > 0$. For every point from the line segment connecting $\mathbf{w}$ and $\mathbf{w}'$, the components of grad J_1 are of order n^{-1}, while the second order derivatives are of order n^{-2}. Therefore

$$J_1(\mathbf{w}') = J_1(\mathbf{w}) + (\mathbf{w}' - \mathbf{w})^{*}\text{grad } J_1(\mathbf{w}) + O(n^{-2}).$$

(* stands for transposition operation.) So, expanding the exponential function,

$$Q(\mathbf{w}') = Q(\mathbf{w}) \left[1 + n^{\alpha}(\mathbf{w}' - \mathbf{w})^{*} \text{ grad } J_1(\mathbf{w}) + O(n^{2(\alpha-1)})\right],$$

and consequently

$$\Sigma = Q(\mathbf{w}) \left\{1 + n^{\alpha}E[\mathbf{w}' - \mathbf{w}|\mathbf{w}]^{*}\text{grad } J_1(\mathbf{w}) + O(n^{2(\alpha-1)})\right\}. \tag{6.4}$$

Recall now that $J_1(\tilde{\mathbf{w}}(\tau))$ remains constant along the trajectory $\tilde{\mathbf{w}}(\tau)$ of the differential equations system (5.5). Geometrically, this means that

$$E_q[\mathbf{w}' - \mathbf{w} \,|\, \mathbf{w}] \perp \text{grad } J_1(\mathbf{w}),$$

so that

$$\begin{aligned} E[\mathbf{w}' - \mathbf{w} \,|\, \mathbf{w}]^{*}\text{grad } J_1(\mathbf{w}) = &\sum_{\mathbf{w}'} (\mathbf{w}' - \mathbf{w})^{*}\text{grad } J_1(\mathbf{w}) \\ &\cdot [p(\mathbf{w}'|\mathbf{w}) - q(\mathbf{w}'|\mathbf{w})]. \end{aligned} \tag{6.5}$$

By Corollary 1,

$$\begin{aligned} |p(\mathbf{w}'|\mathbf{w}) - q(\mathbf{w}'|\mathbf{w})| = &O(n^{-1+3b} + n^{-1/2}\log n) \\ &+ O\left[q(\mathbf{w}'|\mathbf{w}) \sum_{j=1}^{k-1} \frac{(u_j - 1)^{+}}{v_j + 1}\right]. \end{aligned} \tag{6.6}$$

Now (see (4.28), (4.29)), for $1 \le j \le k-1$,

$$\begin{aligned} \sum_{\mathbf{w}'} q(\mathbf{w}'|\mathbf{w}) \frac{(u_j - 1)^{+}}{v_j + 1} = &\sum_{i=1}^{k-1} \frac{v_i}{v} E\left[\frac{(X_j - 1)^{+}}{v_j + 1}\right] \\ = &O\left(\frac{v_j}{\mu^2}\right) = O(n^{-1}). \end{aligned}$$

So the estimate (6.6) becomes

$$|p(\mathbf{w}'|\mathbf{w}) - q(\mathbf{w}'|\mathbf{w})| = O(n^{-1+3b} + n^{-1/2}\log n).$$

Since $\|\text{grad } J_1(\mathbf{w})\| = O(n^{-1})$, we obtain then from (6.5)

$$E[\mathbf{w}' - \mathbf{w}|\mathbf{w}]^* \text{grad } J_1(\mathbf{w}) = O(n^{-2+3b} + n^{-3/2}\log n). \tag{6.7}$$

Thus (see (6.4))

$$\begin{aligned} \Sigma =& Q(\mathbf{w})[1 + O(n^{-\omega}\log n)], \\ \omega :=& \min\{2 - 3b - \alpha,\, 3/2 - \alpha,\, 2(1-\alpha)\} > 1, \end{aligned} \tag{6.8}$$

since $\alpha < \min(1/2, 1-3b)$.

Probabilistically, the relation (6.8) means the following. Introduce the random sequence

$$\{R(\tau)\} := \{Q(\mathbf{w}(\tau))\}.$$

Then, for $\mathbf{w}(\tau) \in \mathbf{W}$,

$$E[R(\tau+1)|\, \mathbf{w}(\tau)] = [1 + O(n^{-\omega})]R(\tau), \tag{6.9}$$

that is $\{R(\tau)\}$ is *almost* a martingale sequence, as long as $\mathbf{w}(\tau-1) \in \mathbf{W}$.

Since the total number of steps is at most n, it follows then from (6.9), that the sequence

$$\{\tilde{R}(\tau)\} := \{(1 + n^{-\omega}\log^2 n)^{-\tau} R(\tau)\} \tag{6.10}$$

is a *supermartingale*, as long as $\mathbf{w}(\tau - 1) \in W$. Fix $x > 0$ and introduce a stopping time

$$\mathcal{T}' = \begin{cases} \min\{\tau \le \mathcal{T} :\ J_1(\mathbf{w}(\tau)) - J_1(\mathbf{w}(0)) > x\}, & \text{if such } \tau \text{ exist,} \\ \mathcal{T} + 1, & \text{otherwise.} \end{cases}$$

Now, applying the Optional Sampling Theorem (Durrett [9]) to the supermartingale $\{\tilde{R}(\tau)\}$ and the stopping time $\mathcal{T} \wedge \mathcal{T}'$, and going back to $\{R(\tau)\}$, we get

$$\begin{aligned} E[Q(\mathbf{w}(\mathcal{T} \wedge \mathcal{T}'))] \le& (1 + n^{-\omega}\log^2 n)^n \cdot E[Q(\mathbf{w}(0))] \\ =& (1 + n^{-\omega}\log^2 n)^n = O(1), \text{ as } n \to \infty. \end{aligned}$$

Since obviously

$$E[Q(\mathbf{w}(\mathcal{T} \wedge \mathcal{T}'))] \ge e^{xn^\alpha} \cdot \mathrm{P}\ \{\mathcal{T}' \le \mathcal{T}\},$$

we have then

$$\begin{aligned} \mathrm{P}\ \{\max_{\tau \le \mathcal{T}} [J_1(\mathbf{w}(\tau)) - J_1(\mathbf{w}(0))] > x\} =& \mathbf{P}\{\mathcal{T}' \le \mathcal{T}\} \\ =& O(e^{-xn^\alpha}), \end{aligned} \tag{6.11}$$

uniformly for $x > 0$. Analogously,

$$\mathrm{P}\ \{\min_{\tau \le \mathcal{T}} [J_1(\mathbf{w}(\tau)) - J_1(\mathbf{w}(0))] < -x\} = O(e^{-xn^\alpha}). \tag{6.12}$$

The estimate (6.3) follows immediately from (6.11) and (6.12), since $0 < c_1(a) \le J_1(\mathbf{w}) \le c_2(a) < \infty$, $\forall w \in \mathbf{W} = W(a)$. $\square$

Corollary 2. *For* $0 < \beta < \alpha < \min\{1/2, 1-3b\}$,

$$\mathbf{P}\{A|\,\mathbf{w}(0) \in \mathbf{W}(a),\, \mu(0) = cn/2\} = O(e^{-n^{\alpha-\beta}}), \tag{6.13}$$

where

$$A := \left\{ \max_{\substack{1\le i\le 3\\ \tau\le T}} \left| \frac{J_i(\mathbf{w}(\tau))}{J_i(\mathbf{w}(0))} - 1 \right| > n^{-\beta} \right\}. \tag{6.14}$$

Proof of Theorem 1. Given $\epsilon > 0$, let P_n denote the probability that the deletion algorithm applied to the random graph $G(n, m = cn/2)$ delivers a k-core of size $\ge \epsilon n$. We have to show that P_n is subexponentially small if

$$c \le \gamma_k - n^{-\delta}, \quad \delta \in (0, 1/2). \tag{6.15}$$

Using the double-conditioning device, we proved (see (4.13)) that

$$\mathbf{P}\{B\} \ge 1 - O(e^{-n^b}), \quad B := \{g(\mathbf{w}(G(n,m))) \le 1/n\}. \tag{6.16}$$

(Recall that $g(\mathbf{w}) := \mathbf{P}\{G(\mathbf{w}) \in \mathcal{H}_n\}$, and $\mathcal{H}_n = \mathcal{H}_n(b)$ is a subset of graphs G such that, in particular, $d_{\max}(G) \ge n^b$; here $b \in (0, 1/3)$ is fixed.) The same method plus Cramér's large deviation theorem for the sums of i.i.d. random variables can be used to show that, for every $b_1 < 1/2$,

$$\mathbf{P}\{C\} \ge 1 - O(e^{-n^{b_1}}), \tag{6.17}$$

where

$$\begin{aligned} \mathrm{C} &:= \mathrm{C}_1 \cap \mathrm{C}_2, \\ \mathrm{C}_1 &:= \left\{ |\overline{v}(G(n,m)) - np_k(c)| \le n^{\frac{1+b_1}{2}} \right\}, \\ \mathrm{C}_2 &:= \left\{ |t(G(n,m)) - nc\pi_k(c)| \le n^{\frac{1+b_1}{2}} \right\}. \end{aligned} \tag{6.18}$$

Notice right now that on the event C

$$z(0) = z(G(n,m)) = c + O(n^{-(1-b_1)/2}) \tag{6.19}$$

(see (4.20)). Clearly, $\mathbf{w}(G(n,m)) \in \mathbf{W}(a)$ on the event $\mathrm{B} \cap \mathrm{C}$, if

$$a < \min\left[p_k(c), \frac{c\pi_k(c)}{p_k(c)} - k \right].$$

Choose a even smaller, so that $a < \epsilon$.

Suppose $G(n,m)$ has a k-core of size $\ge \epsilon n$. If the event $\overline{A}\cap B\cap C$ takes place as well, then $\mathcal{T}=T$, provided that a is chosen sufficiently small. To demonstrate this, suppose $\mathbf{w}(\tau)\notin \mathbf{W}$ for some $\tau<T$. Since $h(\mathbf{w}(\tau))>0$, $g(\mathbf{w}(\tau))\le 1/n$, and $\overline{v}(\tau)\ge\overline{v}(T)\ge\epsilon n>an$, this can happen only if $t(\tau)<(k+a)\overline{v}(\tau)$. But then

$$\frac{z(\tau)e_k'(z(\tau))}{e_k(z(\tau))}=\frac{t(\tau)}{\overline{v}(\tau)}<k+a,$$

and consequently

$$z(\tau)\le\chi a;$$

here (and below) $\chi>0$ (with or without various attributes) stands for an absolute positive constant. Then, using the definition of A in (6.14), and $J_2(\cdot)$ in (5.14), we conclude

$$\begin{aligned}\overline{v}(\tau)\le&2\overline{v}(0)\frac{p_k(z(\tau))}{p_k(z(0))}\\=&\left[2n+O(n^{(1+b')/2})\right]\frac{e^{-z(\tau)}z^k(\tau)}{k!}[1+O(z(\tau))]\\\le&\chi_1 na^k,\end{aligned}$$

(see (6.19)). If $a<\epsilon$ is chosen sufficiently small (which we may assume) then the last inequality is incompatible with $\overline{v}(\tau)\ge\epsilon n$. So indeed, $\mathcal{T}=T$.

Consequently, $\mathbf{w}(T-1)\in\mathbf{W}(a)$, so that $\overline{v}(T-1)$, $t(T-1)$ are of order n, while $2\mu(T)=t(T)$. Since $s(T)=0$, we have

$$2\mu(T-1)=2\mu(T)+O(1)=t(T)+O(1)=t(T-1)+O(1).$$

So, by the definition of $J_3(\cdot)$ in (5.14), we obtain

$$\begin{aligned}J_3(\mathbf{w}(T-1))=&[1+O(n^{-1})]\frac{z(T-1)}{\pi_k(z(T-1))}\\\ge&\gamma_k\cdot[1+O(n^{-1})].\end{aligned}\tag{6.20}$$

(Recall that $\gamma_k:=\min z/\pi_k(z)$.) By the definition of the events C_1 and C_2 in (6.15), and (6.16), we also have

$$J_3(\mathbf{w}(0))=c\cdot\left[1+O(n^{-(1-b_1)/2})\right].\tag{6.21}$$

Putting (6.20) and (6.21) together, and using the definition of the event A, (or rather its complement $\overline{A}$), we arrive at

$$\gamma_k\le c+O(n^{-\beta}+n^{-(1-b_1)/2}).$$

However, in view of (6.15), this is impossible if we choose β and b_1 such that

$$\delta<\beta \text{ and } \delta<\frac{1-b_1}{2}.\tag{6.22}$$

Therefore, for this choice of the parameters β and b_1 the events $\overline{A} \cap B \cap C$ and $\{G(n,m)$ has a core of size $\geq \epsilon n\}$ are disjoint! Hence (see (6.13), (6.16) and (6.17))

$$\begin{aligned} P_n \leq \mathbf{P}\{\overline{\overline{A} \cap B \cap C}\} &\leq \mathbf{P}\{\overline{B \cap C}\} + \mathbf{P}\{A \cap (B \cap C)\} \\ &\leq \mathbf{P}\{\overline{B}\} + \mathbf{P}\{\overline{C}\} + \mathbf{P}\{A | B \cap C\} \\ &= O(e^{-n^{\rho}}), \\ \rho &:= \min\{\alpha - \beta,\, b_1,\, b\}. \end{aligned} \tag{6.23}$$

Besides the restrictions (6.22), the parameters here are also subject to the constraints

$$\begin{aligned} &\beta < \alpha < \min\{1/2, 1 - 3b\}, \\ &b < 1/3, \quad b_1 < 1/2. \end{aligned} \tag{6.24}$$

By taking b, b_1, α sufficiently close from *below* to $1/6$, $\min\{1/2, 1-2\delta\}$, and $\min\{1/2, 1-3b\}$ respectively, and β from *above* to δ, we can make the parameter ρ in (6.23) arbitrarily close (from below) to $\min\{1/2-\delta, 1/6\}$.

To obtain the bound $O(n^{-(k-2)(k+1)/2})$ for the probability that $G(n,m)$ has a k-core of any size, it is enough now to handle the sizes $\leq \epsilon n$, where ϵ can be selected arbitrarily small. The corresponding probability is bounded above by the expected number of k-cores of those small sizes, which turns out to be of the above order, if ϵ is appropriately small. (The dominant contribution to the expectation comes from possible k-cores of the minimum size, $k+1$ that is.) We omit the details. $\square$

Proof of Theorem 2. Now we have to consider the case

$$c \geq \gamma_k + n^{-\delta}, \quad \delta \in (0, 1/2). \tag{6.25}$$

Let the parameters α, β, b and b_1 satisfy the conditions (6.22) and (6.24). Then, by (6.25), on the event C (see (6.18).(6.19))

$$\begin{aligned} z(0) &\geq \gamma_k + n^{-\delta} + O(n^{-(1-b_1)/2}) \\ &\geq \gamma_k + \frac{1}{2} n^{-\delta} \\ &\geq \lambda_k + \frac{1}{2} n^{-\delta}, \end{aligned} \tag{6.26}$$

because $\gamma_k = \lambda_k / \pi_k(\lambda_k) > \lambda_k$. Since $c > \gamma_k$, the minimum value of $\lambda/\pi_k(\lambda)$, the equation

$$\frac{\lambda}{\pi_k(\lambda)} = c$$

has two roots. Let $\lambda_k(c)$ denote the larger root. How far is $\lambda_k(c)$ from λ_k? Since

$$\left(\frac{z}{\pi_k(z)}\right)'\bigg|_{z=\lambda_k} = 0, \qquad \left(\frac{z}{\pi_k(z)}\right)''\bigg|_{z=\lambda_k} > 0, \tag{6.27}$$

using (6.25) we obtain

$$\lambda_k(c) - \lambda_k \geq \chi n^{-\delta/2}. \tag{6.28}$$

We want to show that, with high probability, $t(T) = 2\mu(T)$ and $z(T)$ is close to $\lambda_k(c)$, so that $\overline{v}(T)$ (the size of the k- core) is about $np_k(\lambda_k(c))$. To this end, fix $\nu > 0$ and introduce $\hat{z}$ by

$$\begin{aligned} \hat{z} =& \lambda_k(c) + \nu n^{-\theta}, \\ \theta :=& \min\{\beta/2,\ (1-b_1)/4\}. \end{aligned} \tag{6.29}$$

Notice right now that

$$\begin{aligned} \frac{\hat{z}}{\pi_k(\hat{z})} =& c + \left(\frac{\lambda}{\pi_k(\lambda)}\right)'\Bigg|_{\lambda=\lambda_k(c)} \nu n^{-\theta} + O(n^{-2\theta}) \\ =& c + O(n^{-\eta}), \qquad (\eta := \delta/2 + \theta), \\ \frac{\hat{z}}{\pi_k(\hat{z})} \geq& c + \left[\left(\frac{\lambda}{\pi_k(\lambda)}\right)'\Bigg|_{\lambda=\lambda_k} + \chi_1 n^{-\delta/2}\right] \nu n^{-\theta} + O(n^{-2\theta}) \\ \geq& c + \chi_2 \nu n^{-\eta}, \end{aligned} \tag{6.30}$$

since $\delta/2 < \theta$ (see (6.22)).

Next, set in (6.1), (6.2)

$$a = \min\left\{\frac{1}{1+\nu n^{-\eta}} p_k(\hat{z}),\ \frac{1}{1+n^{-1/2}} \frac{\hat{z}\pi_k(\hat{z})}{p_k(\hat{z})} - k\right\}. \tag{6.31}$$

We claim that if n is large enough, then—with high probability —there exists $\hat{\tau} < \mathcal{T}(a)$ such that

$$z(\hat{\tau} - 1) > \hat{z}, \quad z(\hat{\tau}) \leq \hat{z}. \tag{6.32}$$

To prove this, let us suppose the event $B \cap C$ happens. Then , since $c > \hat{z}$, see (6.29)), $\mathbf{w}(G(n,m)) \in \mathbf{W}(a)$, and (Corollary 2) the event $\overline{A}$ takes place with conditional probability $\geq 1 - O(e^{-n^{\alpha-\beta}})$. Assuming simultaneous occurrence of all three events, $\overline{A}$, B and C, consider two possible alternatives.

1. $\mathcal{T}(a) = T$. In this case, $\overline{v}(T) \geq an$, so that the algorithm delivers a giant k-core. Then $2\mu(T) = t(T)$ and

$$\begin{aligned} \frac{z(T)}{\pi_k(z(T))} =& \frac{t(T)}{2\mu(T)} \frac{z(T)}{\pi_k(z(T))} \\ =& (1 + O(n^{-\beta})) J_3(\mathbf{w}(G(n,m))) \\ =& c\left[1 + O(n^{-\beta} + n^{-(1-b_1)/2})\right] \\ =& c(1 + O(n^{-2\theta})). \end{aligned}$$

Now, $z(T)/z(T-1) = 1 + O(n^{-1})$, because $\|\mathbf{w}(T) - \mathbf{w}(T-1)\| = O(1)$, and $\mathbf{w}(T-1)$, $\mathbf{w}(T) \in \mathbf{W}(a)$; therefore

$$\frac{z(T-1)}{\pi_k(z(T-1))} = c(1 + O(n^{-2\theta})),$$

as well. Since $2\theta > \eta$, we see that, for n large enough

$$\frac{\hat{z}}{\pi_k(\hat{z})} > \frac{z(T)}{\pi_k(z(T))}.$$

The last inequality certainly implies existence of τ that satisfies (6.32). (At this moment of the proof, we do not know yet that is very unlikely that $z(T)$ is close not to $\hat{z}$, but to $\tilde{z} < \hat{z}$ defined by $\tilde{z}/\pi_k(\tilde{z}) = \hat{z}/\pi_k(\hat{z})$.)

2. $\mathcal{T}(a) < T$. Then, for some $\tau < T$, we have $\mathbf{w}(\tau) \in \mathbf{W}(a)$, but either $\overline{v}(\tau+1) < an$ or $t(\tau+1) < (k+a)\overline{v}(\tau+1)$. In the first case,

$$\begin{aligned}\frac{\overline{v}(\tau)}{p_k(z(\tau))} &= (1 + O(n^{-1}))\frac{\overline{v}(\tau+1)}{p_k(z(\tau+1))}\\ &\ge (1 + O(n^{-\beta}))(1 + O(n^{-(1-b_1)/2}))n\\ &\ge n(1 + O(n^{-2\theta})).\end{aligned}$$

So, by the definition of a,

$$\frac{1}{1+\nu n^{-\eta}}\,\frac{p_k(\hat{z})}{p_k(z(\tau))} \ge 1 + O(n^{-2\theta});$$

It follows then $p_k(\hat{z}) > p_k(z(\tau))$, whence $\hat{z} > z(\tau)$, if n is large enough. In the second case,

$$\begin{aligned}\frac{z(\tau)\pi_k(z(\tau))}{p_k(z(\tau))} &= (1 + O(n^{-1}))\frac{z(\tau+1)\pi_k(z(\tau+1))}{p_k(z(\tau+1))}\\ &= (1 + O(n^{-1}))\frac{t(\tau+1)}{\overline{v}(\tau+1)}\\ &\le (1 + O(n^{-1}))(k+a) \le \frac{1+O(n^{-1})}{1+n^{-1/2}}\,\frac{\hat{z}\pi_k(\hat{z})}{p_k(\hat{z})}\\ &< \frac{\hat{z}\pi_k(\hat{z})}{p_k(\hat{z})},\end{aligned}$$

which shows that $z(\tau) < \hat{z}$ since $z\pi_k(z)/p_k(z)$ is strictly increasing (see (5.10)).

Thus,

$$\begin{aligned}&\mathbf{P}\{\exists\hat{\tau} < \mathcal{T}(a) : z(\hat{\tau}-1) > \hat{z},\, z(\hat{\tau}) \le \hat{z}\}\\ &\qquad \ge 1 - O(e^{-n^{\rho}}),\\ &\qquad \rho = \min\{\alpha - \beta,\, b,\, b_1\},\end{aligned} \tag{6.33}$$

(see (6.23)).

We extend the definition of $\hat{\tau}$ setting $\hat{\tau} = \mathcal{T}(a)$ if $z(\tau)$ never falls below $\hat{z}$. Observe that $\hat{\tau}$ is a stopping time adopted to $\{\mathbf{w}(\tau)\}$.

Let us have a close look at the sequence $\{\mathbf{w}(\tau)\}_{\tau \ge \hat{\tau}}$, conditioned on the event $\overline{A} \cap B \cap C$. First of all,

$$z(\hat{\tau}) = (1 + O(n^{-1}))\hat{z}.$$

So,

$$\begin{aligned} \frac{t(\hat{\tau})}{2\mu(\hat{\tau})} =&(1 + O(n^{-\beta}))J_3(\mathbf{w}(G(n,m))) \div \frac{z(\hat{\tau})}{\pi_k(\hat{\tau})} \\ =&(1 + O(n^{-2\theta}))c \,\div\, c(1 + O(n^{-\eta})) \\ =&1 + O(n^{-\eta}), \end{aligned} \tag{6.34}$$

that is

$$s(\hat{\tau}) := 2\mu(\hat{\tau}) - t(\hat{\tau}) = O(n^{1-\eta}).$$

(Recall that $s(\tau)$ is the total degree of light vertices of $G(\tau)$.) However, $\overline{v}(\hat{\tau})$, the total degree of heavy vertices of $G(\hat{\tau})$ is still of order n. More precisely,

$$\begin{aligned} \overline{v}(\hat{\tau}) =&p_k(z(\hat{\tau}))\frac{\overline{v}(0)}{p_k(z(0))}(1 + O(n^{-\beta})) \\ =&(1 + O(n^{-2\theta}))np_k(\hat{z}) \\ =&(1 + O(n^{-\theta}))np_k(\lambda_k(c)). \end{aligned} \tag{6.35}$$

What remains to show is that, with high probability, the deletion process will end within at most $n^{\sigma}, (\sigma \in (0,1))$, steps, delivering a giant k-core having about $np_k(\lambda_k(c))$ vertices.

We will specify σ shortly. Whatever σ is, it is clear that for $\hat{\tau} \le \tau \le \hat{\tau} + n^{\sigma}$,

$$|\mu(\tau) - \mu(\hat{\tau})|, \quad |\overline{v}(\tau) - \overline{v}(\hat{\tau})|, \quad |t(\tau) - t(\hat{\tau})| = O(n^{\sigma}),$$

so that $\mu(\tau)$, $\overline{v}(\tau)$, and $t(\tau)$ are all of order n, while

$$s(\tau) = O(n^{\sigma_1}), \quad \sigma_1 = \max\{1 - \eta, \sigma\}.$$

A little reflection based on the equation $z\pi_k(z)/p_k(z) = t/\overline{v}$ (see (4.20)) and (5.10) shows then that

$$z(\tau) = \hat{z} + O(n^{-(1-\sigma)}).$$

So, comparing with (6.28),(6.29), and remembering that $\delta/2 < \theta$,

$$\begin{aligned} z(\tau) - \lambda_k \ge&\chi n^{-\delta/2} + O(n^{-(1-\sigma)}) \\ \ge&\chi_3 n^{-\delta/2}, \end{aligned} \tag{6.36}$$

if we require that

$$\frac{\delta}{2} < 1 - \sigma_1. \tag{6.37}$$

Denoting

$$\Delta s(\tau) = s(\tau+1) - s(\tau),$$

and using (5.12), we have then: for $s(\tau) > 0$,

$$\begin{aligned} E_q[\Delta s(\tau)|\mathbf{w}(\tau)] \le & - \frac{s(\tau)t(\tau)\pi_k(z(\tau))}{2\mu(\tau)v(\tau)} \left(\frac{z}{\pi_k(z)} \right)' \Bigg|_{z=z(\tau)} + O(n^{-1}) \\ & (\frac{s(\tau)}{v(\tau)} \ge 1) \\ \le & - \chi_4 n^{-\delta/2}. \end{aligned} \tag{6.38}$$

We notice that $\Delta s(\tau) \le 2(k-1)$ always. So, invoking (6.6) and the estimate that follows it, we get from (6.38)

$$E[\Delta s(\tau)|\mathbf{w}(\tau)] \le \chi_5 n^{-\delta/2}. \tag{6.39}$$

($\delta < \min\{1/2, 1-3b\}$, see (6.22),(6.24).)

The rest is short. For $y > 0$, it follows from (6.39) that

$$\begin{aligned} E[e^{ys(\tau+1)}|\mathbf{w}(\tau)] =& e^{ys(\tau)} \exp[yE(\Delta s(\tau)|\mathbf{w}(\tau))] \\ & \cdot E\left\{\exp[y(\Delta s(\tau) - E(\Delta s(\tau)|\mathbf{w}(\tau)))]|\mathbf{w}(\tau)\right\} \\ & \le e^{ys(\tau)} \cdot \exp\left(-y\chi_5 n^{-\delta/2} + 2y^2k^2\right). \end{aligned} \tag{6.40}$$

(We have used a well-known estimate

$$E(e^{yY}) \le e^{y^2d^2/2},$$

provided that $|Y| \le d$ and $E(Y) = 0$.) Set

$$y = y_n = \frac{\chi_5 n^{-\delta/2}}{4k^2};$$

y_n minimizes the second exponent on the right in (6.40). Then (6.40) becomes

$$E(e^{y_n s(\tau+1)}|\mathbf{w}(\tau)) \le e^{y_n s(\tau)} \cdot e^{-\chi n^{-\delta}}, \qquad \chi := \frac{(\chi_5)^2}{8k^2}.$$

Therefore, the sequence

$$\{S(\tau)\}_{\tau\ge\hat{\tau}} := \left\{\exp\left[y_n s(\tau) + (\tau-\hat{\tau})\chi n^{-\delta}\right]\right\}_{\tau\ge\hat{\tau}}$$

is a supermartingale, as long as $s(\tau) > 0$. Hence (the Optional Sampling Theorem again !)

$$\begin{aligned} E\left[S(\hat{\tau} + n^{\sigma} \wedge (T-\hat{\tau}))|\,\mathbf{w}(\hat{\tau})\right] \le & S(\hat{\tau}) \\ =& e^{y_n s(\hat{\tau})} \\ =& \exp\left[\chi^* n^{1-\eta-\delta/2}\right]. \end{aligned}$$

Thus

$$\mathbf{P}\{T-\hat{\tau}\ge n^{\sigma}|\mathbf{w}(\hat{\tau})\}\le \exp[\chi^{*}n^{1-\eta-\delta/2}-\chi n^{\sigma-\delta}] \le \exp[-\chi^{**}n^{\sigma-\delta}], \tag{6.41}$$

if we require

$$\sigma-\delta>1-\eta-\delta/2,\quad (\eta=\delta/2+\theta),$$

or equivalently

$$\sigma>1-\theta. \tag{6.42}$$

For the estimate (6.41) to be useful, we also have to satisfy

$$\delta<\sigma. \tag{6.43}$$

Using (6.33), (6.41), and collecting the constraints (6.22),(6.24),(6.35), (6.37), (6.42), and (6.43), we can state now the following result.

With probability $\ge e^{-n^{\zeta}}$,

$$\zeta:=\min\{\alpha-\beta,b,b_1,\sigma-\delta\}, \tag{6.44}$$

the edge deletion process finds a giant k-core of size $np_k(\lambda_k(c))+O(n^{\phi})$,

$$\phi:=\max\{1-\theta,\sigma\},\quad (\theta=\min\{\beta/2,(1-b_1)/4\}). \tag{6.45}$$

Here

$$\begin{gathered} b<1/3,\quad b_1<1/2,\quad \beta<\alpha<\min\{1/2,1-3b\},\\ \delta<\sigma<1,\\ \delta<2(1-\sigma)<\min\{\beta,(1-b_1)/2\}. \end{gathered} \tag{6.46}$$

It is easy to see that, for every $\delta<1/2$ and $\sigma\in(3/4,1-\delta/2)$, we can satisfy the restrictions (6.46) by choosing b, b_1, α sufficiently close to (but less than) $1/6$, $\min\{1/2,4\sigma-3\}$, $1/2$ respectively, and β sufficiently close to (but more than) $2(1-\sigma)$. This way, we can make ζ in (6.44) arbitrarily close from below to $\min\{2\sigma-3/2,1/6\}$, and ϕ in (6.45) arbitrarily close from below to σ.

This observation completes the proof of Theorem 2. □

Finally,

Proof of Theorem 3. Let $\sigma\in(3/4,1)$ and $\epsilon\in(0,p_k(\lambda_k))$ be given. Denote by $P(n,m)$ the probability that the random graph $G(n,m)$ has a k-core of size in the interval $[n\epsilon,np_k(\lambda_k)-n^{\sigma}]$. We need to show that $P(n,m)$ is subexponentially small, uniformly for $m\le\binom{n}{2}$.

Fix $\delta \in (2(1-\sigma), 1/2)$. According to Theorems 1 and 2, it suffices to consider $m = cn/2$ with

$$|c - \gamma_k| = O(n^{-\delta}).$$

Suppose $G(n,m)$ has a k-core of size in question, $\geq n\epsilon$ in particular. Following line by line the proof of Theorem 1, we obtain: on the event $\overline{A} \cap B \cap C$,

$$\begin{aligned} \frac{z(T)}{\pi_k(z(T))} =& c\left[1 + O(n^{-\min\{\beta,(1-b_1)/2\}})\right] \\ =& \gamma_k\left[1 + O(n^{-\min\{\beta,(1-b_1)/2,\delta\}})\right] \\ =& \gamma_k[1 + O(n^{-\delta})], \end{aligned}$$

provided that

$$\delta < \beta, \text{ and } \delta < \frac{1-b_1}{2}. \tag{6.47}$$

Consequently,

$$|z(T) - \lambda_k| = O(n^{-\delta/2}),$$

and

$$\frac{p_k(z(T))}{p_k(\lambda_k)} = 1 + O(n^{-\delta/2}). \tag{6.48}$$

Furthermore,

$$\begin{aligned} \frac{\overline{v}(T)}{p_k(z(T))} = J_2(\mathbf{w}(T)) =& [1 + O(n^{-\beta})] \cdot J_2(\mathbf{w}(0)) \\ =& n\left[1 + O(n^{-\min\{\beta,(1-b_1)/2\}})\right]. \end{aligned}$$

Since $\overline{v}(T) \leq np_k(\lambda_k) - n^{\sigma}$, the previous estimate yields

$$\begin{aligned} \frac{p_k(\lambda_k)}{p_k(z(T))} \geq& [1 + \chi n^{-(1-\sigma)}] \cdot [1 + O(n^{-\min\{\beta,(1-b_1)/2\}})] \\ =& 1 + \chi' n^{-(1-\sigma)}. \end{aligned} \tag{6.49}$$

(We know that $1 - \sigma < \delta/2 < \min\{\beta, (1-b_1)/2\}$.)

The relations (6.48) and (6.49) are incompatible since $1 - \sigma < \delta/2$. Therefore

$$\begin{aligned} P(n,m) \leq& \mathbf{P}\{\overline{\overline{A} \cap B \cap C}\} \\ =& O(e^{-n^{\rho}}), \\ \rho =& \min\{\alpha - \beta, b, b_1\}. \end{aligned} \tag{6.50}$$

Recall also that

$$\begin{aligned} \beta <& \alpha < \min\{1/2, 1 - 3b\}, \\ & b < 1/3, \quad b_1 < 1/2. \end{aligned} \tag{6.51}$$

Like two times before, it is easy to choose—subject to constraints $\delta > 2(1 - \sigma)$, (6.47), (6.51)—the values of δ, α, β, b and b_1 such that ρ gets arbitrarily close (from below) to $\min\{2\sigma - 3/2, 1/6\}$. □

Acknowledgement. The authors thank Béla Bollobás, Alan Frieze and Tomasz Łuczak for their interest in this study. The work of the first author benefited from a partial relief of teaching in 1994–1995. A referee's comments were very helpful.

References

1. E. A. Bender and E. A. Canfield, *The asymptotic number of labelled graphs with given degree sequences*, J. Comb. Theory (A) **24** (1978), 296–307.
2. B. Bollobás, *A probabilistic proof of an asymptotic formula for the number of labelled regular graphs*, Europ. J. Combinatorics **1** (1980), 311–316.
3. ———, *The evolution of sparse graphs*, Graph Theory and Combinatorics. Proc. Cambridge Combinatorial Conf. in honour of Paul Erdös (B. Bollobás, ed.), Academic Press, 1984, pp. 35–57.
4. ———, *Random Graphs*, Academic Press, 1985.
5. B. Bollobás and A. M. Frieze, *On matchings and Hamiltonian cycles in random graphs*, Random Graphs' 83, Ann. Discrete Math. 28 (M. Karoński and A. Ruciński, eds.), 1985, pp. 23–46.
6. B. Bollobás, *Martingales, isoperimetric inequalities and random graphs*, Combinatorics, Eger (Hungary) 1987, Coll. Math. Soc. J. Bolyai, 52, Akad. Kiadó, Budapest, 1988, pp. 113–139.
7. B. Bollobás and S. Rasmussen, *First cycles in random directed graph processes*, Discrete Math. **75** (1989), 55–68.
8. V. Chvátal, *Almost all graphs with 1.44 edges are 3-colourable*, Random Structures and Algorithms **2** (1991), 11–28.
9. R. Durrett, *Probability: Theory and Examples*, Wadsworth & Brooks. Pacific Grove, 1991.
10. P. Erdös and A. Rényi, *On the evolution of random graphs*, Magy. Tud. Akad. Mat. Kut. Intéz. Közl. **5** (1960), 17–61.
11. W. Feller, *An Introduction to Probability Theory and Its Applications, II, 2nd Edition*, John Wiley & Sons, 1971.
12. P. Flajolet, D. E. Knuth, and B. Pittel, *The first cycles in an evolving graph*, Discrete Math. **75** (1989), 167–215.
13. G. Havas, B.S. Majewski, N.C. Wormald, and Z.J. Czech, *Graphs, Hypergraphs and Hashing*, 19th International Workshop on Graph-Theoretic Concepts in Computer Science (WG' 93), Utrecht, 1993, vol. 790, Springer–Verlag, Lecture Notes in Computer Science, 1994, pp. 153–165.
14. S. Janson, *Poisson convergence and Poisson processes with applications to random graphs*, Stochastic Processes and their Applications **26** (1987), 1-30.
15. S. Janson, D. E. Knuth, T. Łuczak, and B. Pittel, *The birth of the giant component*, Random Structures and Algorithms (Issue in Honour of Dr. Paul Erdös) **4** (1993), 233–358.
16. R. M. Karp and M. Sipser, *Maximum matchings in sparse random graphs*, IEEE Conference on Computing, 1981, pp. 364–375.
17. R. M. Karp, *The transitive closure of a random digraph*, Random Structures and Algorithms **1** (1990), 73–93.
18. T. Łuczak. Private communication. (1989).
19. ———, *Component behavior near the critical point of the random graph process*, Random Structures and Algorithms **1** (1990), 287–310.
20. ———, *Size and connectivity of the k-core of a random graph*, Discrete Math. **91** (1991), 61–68.
21. ———, *Sparse random graphs with a given degree sequence*, Random Graphs, 2. (A. Frieze and T. Łuczak, eds.), John Wiley & Sons, 1992, pp. 165–182.
22. T. Łuczak, B. Pittel, and J. C. Wierman, *The structure of a random graph at the point of phase transition*, Transactions of the American Mathematical Society **341** (1994), 721–748.
23. B. D. McKay and N. C. Wormald, *Asymptotic enumeration by degree sequence of graphs with degrees $o(n^{1/2})$*, Combinatorica **11** (1991), 369–382.
24. B.S. Majewski, N.C. Wormald, G. Havas, and Z.J. Czech, *A Family of Perfect Hashing Methods*, (submitted).
25. M. Molloy, *M. Math. Thesis, University of Waterloo* (1992).
26. B. Pittel, *On a Daley–Kendall model of random rumours*, J. Appl. Prob. **27** (1990), 14–27.

27. B. Pittel, *On trees census and the giant component in sparse random graphs*, Random Structures and Algorithms **1** (1990), 311–342.

28. B. Pittel and W. A. Woyczynski, *A graph–valued Markov process as rings–allowed polymerization model: subcritical behavior*, SIAM J. Appl. Math. **50** (1990), 1200–1220.

29. N. C. Wormald, *Differential equations for random processes and random graphs*, Annals of Applied Probability (submitted).