# On the threshold for $k$-regular subgraphs of random graphs


Paweł Prałat
Department of Mathematics
West Virginia University
Morgantown, WV, USA

Jacques Verstraëte
Department of Mathematics
University of California
San Diego, CA, USA

Nicholas Wormald *
Department of Combinatorics and Optimization
University of Waterloo
Waterloo ON, Canada



**Abstract**

The $k$-core of a graph is the largest subgraph of minimum degree at least $k$. We show that for $k$ sufficiently large, the threshold for the appearance of a $k$-regular subgraph in the Erdős-Rényi random graph model $\mathcal{G}(n, p)$ is at most the threshold for the appearance of a nonempty $(k+2)$-core. In particular, this pins down the point of appearance of a $k$-regular subgraph to a window for $p$ of width roughly $2/n$ for large $n$ and moderately large $k$. The result is proved by using Tutte's necessary and sufficient condition for a graph to have a $k$-factor.


## 1 Introduction

In this paper, we study the appearance of $k$-regular subgraphs of random graphs. The *$k$-core* of a graph $G$ is the unique largest subgraph of $G$ of minimum degree at least $k$ (note that the $k$-core may be empty). Evidently, the $k$-core of a graph can be found be repeatedly deleting vertices of degree less than $k$ from the graph. In the case $k = 2$, this corresponds to the appearance of cycles in $\mathcal{G}(n, p)$, which is well-researched, and precise results concerning the distribution of cycles may be found in Janson [7] and Flajolet, Knuth and Pittel [6]. When discussing thresholds of $k$-cores and $k$-regular subgraphs, we mean thresholds for nonempty $k$-cores and nonempty $k$-regular subgraphs. By analysing the vertex deletion algorithm for the Erdős-Rényi model $\mathcal{G}(n, p)$ of random graphs, Pittel, Spencer and Wormald [11] proved that for fixed $k \geq 3$, $c_k/n$ is a sharp threshold for a $k$-core in $\mathcal{G}(n, p)$ where the constant $c_k$


*Supported by the Canada Research Chairs Program and NSERC

is given by

$$c_k = \frac{\lambda_k}{\pi_k(\lambda_k)}\,, \tag{1}$$

$\pi_k(\lambda)$ is defined by

$$\pi_k(\lambda) = \sum_{j\geq k-1} \frac{e^{-\lambda}\lambda^j}{j!}\,, \tag{2}$$

and $\lambda_k$ is the positive number minimising the right hand side of (1). Hence, any threshold of appearance of a $k$-regular subgraph, for $k \geq 3$, must be essentially at least $c_k/n$. Recently, a number of simpler proofs establishing the threshold $c_k/n$ for the $k$-core have been published (see Kim [10], Cain and Wormald [5], and Janson and Łuczak [8]).

Our main approach relies on a parity-free weakening of Tutte's necessary and sufficient condition for a graph to have a $k$-factor (see Theorem 2).

### 1.1 Regular subgraphs

In what follows, we write a.a.s. to denote an event which occurs with probability tending to one as $n \to \infty$. In comparison to studying the $k$-core in random graphs, it appears to be substantially more difficult to analyse the appearance of $k$-regular subgraphs when $k \geq 3$. One reason is that it is NP-hard to determine whether a graph contains such a subgraph, and there is no analogue of the simple vertex deletion algorithm which produces the $k$-core. As every $k$-regular subgraph of a graph is contained in the $k$-core of the graph, we deduce that $\mathcal{G}(n,p)$ a.a.s. does not contain a $k$-regular subgraph whenever $p$ is below the threshold for the $k$-core described in (1) and (2). Bollobás, Kim and Verstraëte [3] showed that $\mathcal{G}(n,p)$ a.a.s. contains a $k$-regular subgraph when $p$ is, roughly, larger than $4c_k/n$, and conjectured a sharp threshold for the appearance of $k$-regular subgraphs in $\mathcal{G}(n,p)$. In the same paper it was shown that for some $c > c_3$, the 3-core of $\mathcal{G}(n,c/n)$ has no 3-regular subgraph a.a.s., whereas for $c > c_4$, the 4-core of $\mathcal{G}(n,c/n)$ contains a 3-regular subgraph a.a.s. In support of the conjecture of a sharp threshold, Pretti and Weigt [12] numerically analysed equations arising from the cavity method of statistical physics to conclude empirically that indeed, there is a sharp threshold for the appearance of a $k$-regular subgraph of a random graph. For $k > 3$ they concluded that it is the same as the threshold for the $k$-core, which is at odds with [3, Conjecture 1.3]. For $k = 3$, these thresholds differ, as shown using the first moment method in [3].

In this paper, we improve the window of the threshold for $k$-regular subgraphs in $\mathcal{G}(n,p)$ by proving Theorem 1 and its corollary below. A *$k$-factor* of a graph is a spanning $k$-regular subgraph, and a graph is *$k$-factor critical* if, whenever we delete a vertex from the graph, we obtain a graph which has a $k$-factor.

**Theorem 1.** *There exists an absolute constant $k_0$ such that for $k \geq k_0$ and constant $c > c_{k+2}$ with $c < k + 2\sqrt{k \log k}$, the $(k+2)$-core of a random graph $\mathcal{G}(n,c/n)$ is a.a.s. nonempty and either contains a $k$-factor or is $k$-factor-critical.*

Theorem 1 will be proved in Section 4. We remark that the first nonempty $k$-core of the random graph process a.a.s. contains many vertices of degree $k+1$ adjacent to $k+1$ vertices of degree $k$, so the $k$-core cannot contain a $k$-factor and cannot be $k$-factor critical a.a.s. Bollobás, Cooper, Fenner and Frieze [2] conjectured that the $(k+1)$-core contains $\lfloor k/2 \rfloor$ edge disjoint hamiltonian cycles a.a.s., so Theorem 1 supports this conjecture.

The value of $c_k$ can be determined approximately for large $k$ as follows. This corrects, and sharpens, the error term of the formula given in [11]. All logarithms in this paper are natural, and $\mathbb{N}$ is the set of positive integers.

**Lemma 1.** *For any $k \in \mathbb{N}$, let $q_k = \log k - \log(2\pi)$. Then*

$$c_k = k + (kq_k)^{1/2} + \Big(\frac{k}{q_k}\Big)^{1/2} + \frac{q_k - 1}{3} + O\Big(\frac{1}{\log k}\Big) \quad \text{as } k \to \infty.$$

Lemma 1 is proved in Section 5. Hence, for sufficiently large $k$ we have $k < c_k < 4k/3$, and consequently Theorem 1 applies to a nonempty range of $c$. Since containment of a $k$-regular subgraph is a monotone property, we thus have the following.

**Corollary 1.** *There exists an absolute constant $k_0$ such that for $k \geq k_0$, $\varepsilon > 0$ and any function $p(n) > c_{k+2}(1+\varepsilon)/n$, the random graph $\mathcal{G}(n,p)$ a.a.s. has a $k$-regular subgraph.*

It follows immediately from Lemma 1 that

$$c_{k+2} = c_k + 2 + O\Big(\frac{1}{\log k}\Big).$$

Hence, we have pinned down the threshold for the appearance of $k$-regular subgraphs in $\mathcal{G}(n,p)$ to a window for $p$ of width $2/n + O(1/n \log k)$. The following two problems remain unsettled.

**Conjecture 1.** [3] *There is a sharp threshold for the appearance of a $k$-regular graph in $\mathcal{G}(n,p)$, in other words, there exists a constant $\rho$ such that for any $\epsilon > 0$, if $p > (1+\epsilon)\rho/n$ then $\mathcal{G}(n,p)$ a.a.s. contains a $k$-regular subgraph, and if $p < (1-\epsilon)\rho/n$ then $\mathcal{G}(n,p)$ a.a.s. does not contain a $k$-regular subgraph.*

Some empirical evidence for this conjecture is given in Pretti and Weigt [12]. The second problem is to show that the $(k+1)$-core of a random graph, when it is a.a.s. nonempty, contains a $k$-factor or is $k$-factor critical a.a.s.

**Notation.** Throughout the paper, we denote by $\mathcal{G}(n,p)$ the Erdős-Rényi model of random graphs with independent edges having probability $p = p(n)$ each. If $G$ is a graph with vertex set $V(G)$, then $\lambda_G(S,T)$ denotes the number of edges of $G$ with one endpoint in $S$ and one endpoint in $T$, where $S, T \subseteq V(G)$. If $S = T$, we write $\lambda_G(S)$ instead. Also, $G[S]$ denotes the subgraph induced by $S$. The number of components of a graph $G$ is denoted by $\omega(G)$. Let $d_G(v)$ denote the degree of $v$ in $G$. Apart from Section 2, where graphs may have multiple edges, all graphs considered are simple (they have no multiple edges).

## 2 Factors of Graphs

Let $G$ be a graph (which, in this section, may have multiple edges) and let $k \in \mathbb{N}$. Recall that a *$k$-factor* of $G$ is a spanning subgraph of $G$ all of whose vertices have degree $k$. A graph is *$k$-factor-critical* if the deletion of any vertex of $G$ results in a graph with a $k$-factor. It is convenient for us to define

$$\delta_k(G) = \begin{cases} 0 & \text{if } k|V(G)| \text{ is even} \\ 1 & \text{if } k|V(G)| \text{ is odd.} \end{cases}$$

The following sufficient condition for a graph to have a $k$-factor follows immediately from the $f$-factor theorem of Tutte [13] (see [1, Corollary 3.11, p.78]).

**Theorem 2.** *Let $k \in \mathbb{N}$, and let $G$ be a graph with $\delta_k(G) = 0$ such that for every pair of disjoint sets $S, T \subseteq V(G)$ for which $S \cup T \neq \emptyset$,*

$$\sum_{v \in T} d_G(v) + k|S| \geq \omega(G - (S \cup T)) + k|T| + \lambda_G(S, T). \tag{3}$$

*Then $G$ has a $k$-factor.*

In fact, Tutte's theorem specialised to $k$-factors says that if $\omega$ is replaced by the number of components having the correct parity, then the condition is both necessary and sufficient. It is straightforward to give constructions of graphs which show that the condition (3) given in this theorem is tight in many instances. Perhaps the simplest construction is to take the multigraph consisting of $K_{1,3}$ with each edge multiplied $\mu$ times, and then join two non-adjacent vertices with a single edge. For $k \geq 2$, it is straightforward to check that this graph has no $k$-factor, whereas if $S$ is the central vertex of degree $3\mu$ and $T$ is the unique vertex of degree $\mu$, then the left hand side of (3) is $\mu + k$ whereas the right hand side is exactly $1 + k + \mu$. The following simple corollary will be used, which applies especially when a graph has an odd number of vertices and therefore cannot have a $k$-factor if $k$ is odd.

**Corollary 2.** *Let $k \in \mathbb{N}$ and let $G$ be a graph with $\delta_k(G) = 1$. Suppose that for all $x \in V(G)$ and disjoint sets $S, T \subseteq V(G - x)$ with $S \cup T \neq \emptyset$ and $R = S \cup \{x\}$,*

$$\sum_{v \in T} d(v) + k|S| \geq \omega(G - (R \cup T)) + k|T| + \lambda_G(R, T). \tag{4}$$

*Then $G$ is $k$-factor critical.*

*Proof.* Let $J = G - \{x\}$ and let $S, T \subseteq V(J)$ be disjoint sets for which $S \cup T \neq \emptyset$. Let $R = S \cup \{x\}$ and let $\lambda_G(x, T)$ denote the number of neighbours of $x$ in $T$. By (4),

$$\begin{aligned} \sum_{v \in T} d_J(v) + k|S| &= \sum_{v \in T} d(v) + k|S| - \lambda_G(x, T) \\ &\geq \omega(G - (R \cup T)) + k|T| + \lambda_G(R, T) - \lambda_G(x, T) \\ &= \omega(J - (S \cup T)) + k|T| + \lambda_J(S, T). \end{aligned}$$

Since $\delta_k(J) = 0$, we conclude from Theorem 2 that $J$ has a $k$-factor, regardless of the choice of $x$. So $G$ is $k$-factor critical. ∎

## 3 Structure of the $k$-core

In this section we describe the structure of the $k$-core in $\mathcal{G}(n,p)$; this material will be used throughout the proof of Theorem 1. We will assume throughout that $p = c/n$ where $c > c_k$, so that the $k$-core of $\mathcal{G}(n,p)$ is a.a.s. nonempty. We let $K$ denote this nonempty $k$-core.

In the first lemma, $\partial X$ denotes the set of edges of $K$ with exactly one endpoint in a set $X \subset V(K)$. The lemma seems to be well known, and follows, for example, from Benjamini, Kozma and Wormald [4, Lemma 5.3]. (That lemma concerns graphs with a given degree sequence, all degrees between 3 and $n^{0.02}$. See the proof of [4, Theorem 4.2] to find the connection with the following.)

**Lemma 2.** *There is a positive constant $\gamma$ such that the following holds. Fix $k \geq 3$. Then a.a.s. every set $X \subset V(K)$ of at most $\frac{1}{2}|V(K)|$ vertices satisfies*

$$|\partial X| \geq \gamma k |X|.$$

Throughout the rest of the paper, $\gamma$ denotes the constant appearing in Lemma 2.

**Lemma 3.** *For fixed $k$ and $C$, a.a.s. every set $Y \subset V(K)$ of size at most $t(n) = C \log n / \log\log n$ induces a graph with at most $|Y|$ edges.*

*Proof.* The expected number of subgraphs of $y$ vertices with at least $y+1$ edges, for some $y \leq t(n)$, is at most

$$\begin{aligned}
\sum_{y \leq t(n)} \binom{n}{y} \binom{\binom{y}{2}}{y+1} p^{y+1} &\leq \sum_{y \leq t(n)} \left(\frac{en}{y}\right)^y \left(\frac{ey^2}{2(y+1)}\right)^{y+1} \left(\frac{c}{n}\right)^{y+1} \\
&< \sum_{y \leq t(n)} \frac{(e^2c)^{y+1} y}{n} = n^{-1+o(1)}.
\end{aligned}$$

So the claim follows from Markov's inequality. ∎

**Lemma 4.** *Let $k > 2/\gamma$. Then a.a.s. for every set $Y \subset V(K)$ of size at most $s(n) = \log n / 2ec \log\log n$, $K - Y$ contains a component with more than $|V(K)| - 2s(n)$ vertices.*

*Proof.* By Lemma 3, we may assume that all sets of $2y \leq 2s(n)$ vertices in $K$ induce at most $2y$ edges. We may also assume that a.a.s. the property in Lemma 2 holds.

It suffices to show that if $X$ is the vertex set of a union of components of $K - Y$ and $|X| \leq \frac{1}{2}|V(K)|$, then $|X| < s(n)$. Suppose $|X| \geq s(n)$. From the property in Lemma 2, $|\partial X| \geq \gamma k |X|$, and since $X$ is a union of components, this means

$$\lambda(Y \cup X) \geq \gamma k |X|.$$

Suppose $|Y| = y$. By averaging over sets $Z \subset X$ of size $y$, some such set satisfies

$$\lambda(Y \cup Z) \geq \gamma k y > 2y.$$

However $|Y \cup Z| = 2y$, which is a contradiction. ∎

Our final lemma is a large deviation result for the degrees of the vertices of the $k$-core. Essentially, the degree of a vertex in $K$ has (asymptotically) a truncated Poisson distribution, which gives a precise bound on the number of vertices which deviate from degree $c$ in $K$.

**Lemma 5.** *For all $\varepsilon > 0$ there exists $k_\varepsilon$ such that for $k > k_\varepsilon$ and $c_k < c < 2k$, it is a.a.s. true that*

$$|d(v) - c| \geq \varepsilon\sqrt{k \log k}$$

*for at most $\varepsilon|K|$ vertices $v$ of $K$.*

*Proof.* Let $\varepsilon > 0$, and fix $k$, and $j \geq k+2$. From [5, Corollary 3 and Erratum], if $n_j$ denotes the number of vertices of degree $j$ in $K$, then a.a.s.

$$n_j = \frac{e^{-\mu}\mu^j}{j!}n + o(n), \tag{5}$$

where $\mu = \mu_{k,c}$ is the larger of the two positive solutions of the equation

$$\frac{\mu}{c} = e^{-\mu}\sum_{i\geq k-1}\frac{\mu^i}{i!}. \tag{6}$$

(The fact that there are two such solutions is known to be guaranteed by the fact that $c > c_k$.) Let $\varepsilon_1 > 0$, and suppose that $\mu = \Theta(k)$. Then, since the Poisson distribution is asymptotically normal with variance equal to its mean, we have for sufficiently large $k$

$$\sum_{|i-\mu|\geq\varepsilon_1\sqrt{k\log k}} e^{-\mu}\frac{\mu^i}{i!} < \varepsilon_1. \tag{7}$$

Also, by Lemma 1, we may assume that $c > k + \frac{1}{2}\sqrt{k \log k}$. Suppose that $c-2$ is substituted for $\mu$ in (6). It is then elementary to obtain that the right hand side of (6) is greater than $1 - 1/k$. Recalling also that $c < 2k$, this is greater than the left hand side of (6). On the other hand, if anything larger than $c$ is substituted for $\mu$ in (6) then the left hand side is greater than the right, since the right is equal to a probability strictly less then 1. So by continuity, $c - 2 < \mu_{k,c} < c$. Taking $\varepsilon_1$ slightly smaller than $\varepsilon$, the lemma follows from (7) and (5). ∎

# 4 Proof of Theorem 1

In this section, we denote by $K$ the $(k+2)$-core of $\mathcal{G}(n,p)$, where $pn = c$ for a constant $c > c_{k+2}$, and $k \geq 3$. Setting $R = S$ when $\delta_k(K) = 0$ and $R = S \cup \{x\}$ otherwise, the key inequalities in Theorem 2 and Corollary 2 are implied in the graph $K$ by

$$\sum_{v\in T}(d(v) - k) + k|R| - k\delta_k(K) \geq \omega(K - (R \cup T)) + \lambda_K(R,T). \tag{8}$$

and hence, noting that all vertices of $T$ have degree at least $k+2$, also by

$$2|T| + k|R| - k\delta_k(K) \geq \omega(K - (R \cup T)) + \lambda_K(R,T). \tag{9}$$

To prove Theorem 1, it suffices to show that there exists $k_0$ such that for $k \geq k_0$, a.a.s. for all disjoint subsets $R$ and $T$ of $V(K)$ with $|R| \geq \delta_k(K)$ and $|R \cup T| > \delta_k(K)$, either of the above bounds holds. The value of $k_0$ will not be optimized in our proof. We consider a number of cases according to the size of the set $R \cup T$. It is convenient throughout to let $s(n) = \log n / 2ec \log\log n$. From now on, $\lambda$ without any subscript denotes $\lambda_K$.

**Case 1** $|R \cup T| < s(n)$.

Let $Y$ be any set of at most $s(n)$ vertices of $K$. By Lemma 4, $K - Y$ contains a component with more than $|V(K)| - 2s(n)$ vertices a.a.s. Contract the other components of $K - Y$ to a set $X$ of single vertices to obtain a graph $J$ with $V(J) = X \cup Y$. Noting that the set of all vertices contributing to $J$ includes at most $3s(n)$ vertices of $G$, it follows by applying Lemma 3 that $J$ has no more edges than vertices. Hence

$$\lambda_J(X, Y) + \lambda(Y) \leq \lambda_J(X \cup Y) \leq |X| + |Y|.$$

Since $K$ is $(k+2)$-connected, each vertex of $X$ has degree at least $k+2$ in $J$. Also $X$ is an independent set in $J$, so we have $\lambda_J(X, Y) \geq (k+2)|X|$. It follows that $(k+1)|X| + \lambda(Y) \leq |Y|$. Since $|X| = \omega(K - Y) - 1$ we obtain

$$(k+1)(\omega(K - Y) - 1) + \lambda(Y) \leq |Y|. \tag{10}$$

Let $R, T \subset V(K)$ be disjoint sets with $0 < |R \cup T| < s(n)$ and let $Y = R \cup T$. Note that (10) implies $\omega(K - Y) + \lambda(Y) \leq |Y|$, and therefore

$$\begin{aligned} 2|T| + k|R| &\geq \omega(K - Y) + \lambda(Y) + k|R| + 2|T| - |Y| \\ &\geq \omega(K - Y) + \lambda(R, T) + (k-1)|R| + |T| \end{aligned}$$

since $|R| + |T| = |Y|$. This verifies (9) immediately if $\delta_k(K) = 0$, whilst if $\delta_k(K) = 1$, it follows from $|R| \geq 1$ and $|R| + |T| \geq 2$. This completes the analysis of Case 1.

For the rest of the proof, $\varepsilon_0$ denotes any positive constant which is sufficiently small, for instance $\gamma^4 e^{-7}/1000$.

**Case 2** $|R \cup T| \geq s(n)$, $|T| < \varepsilon_0 n$, and $|R| < 4\varepsilon_0 n$.

Let $Y = R \cup T$. In this case we estimate $\omega(K - Y)$ and $\lambda(R, T)$ separately. First we show that $\omega(K - Y) \leq |Y|/2$ a.a.s. provided that $k$ is large enough to ensure that $\gamma(k+2) \geq 4$ (this causes us to require $k_0 \geq 4/\gamma - 2$ in our proof). It suffices to show that if $X$ is the vertex set of any union of components of $K - Y$, then a.a.s. $|X| < |Y|/2$ or $|X| > n/2$. Suppose that $|Y|/2 \leq |X| \leq n/2$. Then Lemma 2 shows $\lambda(X, Y) \geq \gamma(k+2)|X|$. Let

$$I = \{y : s(n) \leq y \leq 5\varepsilon_0 n\} \quad \text{and} \quad I_y = \left\{x : \frac{y}{2} \leq x \leq \frac{n}{2}\right\}.$$

The expected number of pairs of sets $(X, Y)$ in $\mathcal{G}(n, p)$ satisfying the above requirements is at most (using $y \leq 2x$ and the theorem's hypothesis $k \geq 2c/3$ for the second step, and

$y \le 5\varepsilon_0 n \le \gamma^4 n/81e^7$ for the last)

$$
\begin{aligned}
\sum_{I_y \times I} \binom{n}{x+y}\binom{x+y}{x}\binom{xy}{\gamma(k+2)x} p^{\gamma(k+2)x} &\le \sum_{I_y \times I} \left(\frac{en}{x+y}\right)^{x+y} 2^{x+y} \left(\frac{exyc}{n\gamma(k+2)x}\right)^{\gamma(k+2)x} \\
&< \sum_{I_y \times I} \left(\frac{2en}{y}\right)^{3x} \left(\frac{3ey}{2n\gamma}\right)^{4x} \\
&= \sum_{I_y \times I} \left(\frac{81e^7 y}{2\gamma^4 n}\right)^{x} = o(1),
\end{aligned}
$$

where the sums are over $x \in I_y$ and $y \in I$. By Markov's Inequality, we conclude that $\omega(K - Y) \le |Y|/2$ a.a.s.

To finish verifying inequality (9) in Case 2, it remains to show that a.a.s.

$$
\lambda(R, T) \le \frac{3}{2}|T| + \left(k - \frac{1}{2}\right)|R| - k. \tag{11}
$$

Let $|R| = \sigma n$ and $|T| = \tau n$, where $\sigma$ and $\tau$ are allowed to depend on $n$. Then the number of ways of choosing the sets $R$ and $T$ is bounded above by

$$
\binom{n}{\sigma n}\binom{n}{\tau n} < \left(\frac{e}{\sigma}\right)^{\sigma n}\left(\frac{e}{\tau}\right)^{\tau n}.
$$

For any $\rho \ge 0$, the probability that there are at least $\rho n$ edges between $R$ and $T$ is at most

$$
\binom{\sigma\tau n^2}{\rho n}\left(\frac{c}{n}\right)^{\rho n} \le \left(\frac{e\sigma\tau c}{\rho}\right)^{\rho n}
$$

Now set $\rho = \rho_\varepsilon = \big((k - \frac{1}{2})\sigma + \frac{3}{2}\tau\big)(1 - \varepsilon)$ for any $\varepsilon > 0$. Then for large enough $n$, the right hand side of (11) is greater than $\rho_\varepsilon n$. Hence, multiplying the bound on the number of choices of $S$ and $T$ by the probability of at least $\rho_\varepsilon n$ edges, the expected number of sets $S$ and $R$ that fail to satisfy (11) is at most

$$
\left(\left(\frac{e\sigma\tau c}{\rho_\varepsilon}\right)^{\rho_\varepsilon}\left(\frac{e}{\sigma}\right)^{\sigma}\left(\frac{e}{\tau}\right)^{\tau}\right)^{n}.
$$

So by the first moment principle, the bound (11) is true a.a.s. if for some fixed $\varepsilon, \varepsilon' > 0$

$$
\left(\frac{ec\sigma\tau}{\rho_\varepsilon}\right)^{\rho_\varepsilon} \le \left(\frac{\sigma}{e}\right)^{\sigma}\left(\frac{\tau}{e}\right)^{\tau}(1 - \varepsilon').
$$

For this it suffices that

$$
\frac{ec}{\rho_\varepsilon} \le \min\left\{\left(\frac{\sigma}{e}\right)^{1/(k-\frac{1}{2})}, \left(\frac{\tau}{e}\right)^{2/3}\right\}(1 - \varepsilon') \tag{12}
$$

since $\sigma/\rho_\varepsilon < 1/(k - \frac{1}{2})$, $\tau/\rho < 2/3$, and $\sigma, \tau < e$.

Since $\tau < e^{-9}$ and $c < 2k-1$ for large enough $k$,

$$\frac{ec}{\rho_\varepsilon} = \frac{ec/(1-\varepsilon)}{(k-\frac{1}{2})/\tau + 3/2\sigma} < \frac{e\tau c}{(1-\varepsilon)(k-\frac{1}{2})} < \frac{2e\tau}{1-\varepsilon} < \frac{2\tau^{2/3}}{e^2(1-\varepsilon)} < \left(\frac{\tau}{e}\right)^{2/3}(1-\varepsilon') \tag{13}$$

for $\varepsilon, \varepsilon'$ sufficiently small. We will prove by contradiction that

$$ec/\rho_\varepsilon \le (\sigma/e)^{1/(k-\frac{1}{2})}(1-\varepsilon'). \tag{14}$$

If it is false,

$$\left(\frac{\sigma}{e}\right)^{1/(k-\frac{1}{2})}(1-\varepsilon') < \frac{e\tau c}{(1-\varepsilon)(k-\frac{1}{2})} < e^{-7}$$

for $\varepsilon$ sufficiently small. It follows, for $\varepsilon'$ sufficiently small, that $\sigma < e^{-4k}$. Using this, as well as $2k-1 \ge 5$ and $c < 2k-1 < e^{2k-1}$, we have

$$\frac{ec}{(k-\frac{1}{2})/\tau + 3/2\sigma} < \frac{2e\sigma c}{3} < \frac{2c\sigma^{1/2}}{3e^{2k}} < \frac{2\sigma^{1/2}}{e} < \frac{2}{3}\left(\frac{\sigma}{e}\right)^{1/(k-\frac{1}{2})}$$

which gives (14) in any case via (13).

**Case 3** $|R \cup T| \ge s(n)$ and $|T| < \varepsilon_0 n$ and $|R| \ge 4\varepsilon_0 n$.

To verify (9) it is enough to show $\lambda(R,T) \le \frac{3}{4}k|R|$, because $\omega\big(K - (R \cup T)\big) \le n < \frac{1}{5}k|R| < \frac{1}{4}k|R| - k$ for large enough $k$ (as $\varepsilon_0$ is fixed). We have $2c\varepsilon_0 n \le \frac{1}{2}c|R| \le \frac{3}{4}k|R|$ since we the bound on $c$ in the theorem gives $c < 3k/2$ for large $k$. Thus (9) follows if $\lambda(R,T) \le 2c\varepsilon_0 n$ holds. So, this case is covered by showing that a.a.s. all sets of at most $\varepsilon_0 n$ vertices (in particular, $T$) have total degree at most $2c\varepsilon_0 n$.

It is well known that in the random graph $\mathcal{G}(n, c/n)$, the vertex degrees are have asymptotically Poisson distribution with mean $c$: the number of vertices of degree $j$ is a.a.s. asymptotic to $e^{-c}c^j/j!$ as $n \to \infty$, for each fixed $j$. It follows that the sum of the degrees of those vertices of degree less than $3c/2$ is a.a.s. asymptotic to

$$n \sum_{j<3c/2} j\frac{e^{-c}c^j}{j!} = n\left(c - \sum_{j \ge 3c/2} j\frac{e^{-c}c^j}{j!}\right).$$

By elementary considerations and using $j! \ge (j/e)^j$, for large enough $c$ (i.e. large enough $k$, considering the observation after Lemma 1) we have

$$\sum_{j \ge 3c/2} j\frac{e^{-c}c^j}{j!} \le \frac{3c}{2}e^{c/2}\left(\frac{2}{3}\right)^{3c/2} \le \frac{3c}{2}\left(\frac{8}{9}\right)^{c/2} < \varepsilon_0.$$

Thus, since $\mathcal{G}(n, c/n)$ a.a.s. has total degree $cn + o(n)$, the vertices of degree at least $3c/2$ a.a.s. have total degree less than $\varepsilon_0 n + o(n)$. Of the other vertices, any set of at most $\varepsilon_0 n$ have total degree at most $(3c/2)\varepsilon_0 n$. So we conclude that any set of at most $\varepsilon_0 n$ vertices has total degree at most

$$\varepsilon_0 n + o(n) + \frac{3c}{2}\varepsilon_0 n < 2c\varepsilon_0 n,$$

as required.

**Case 4** $|T| \geq \varepsilon_0 n$.

By Lemma 1, $c_k = k + \sqrt{k \log k} + o(\sqrt{k \log k})$. Then by Lemma 5, for all $\varepsilon > 0$, if $k$ is sufficiently large, we have a.a.s.

$$\sum_{v \in T} d(v) > \big(k + (1-\varepsilon)\sqrt{k \log k}\,\big)|T|.$$

So, using the fact that $\varepsilon_0$ is an absolute constant and hence that $n = O(|T|)$, we may deduce that (8) holds a.a.s. if we show, for some $\varepsilon > 0$, that a.a.s.

$$(1-\varepsilon)\sqrt{k \log k}\,|T| + k|R| \geq \lambda(R, T). \tag{15}$$

We will prove this by considering the cases $|R| < \eta n$ and $|R| \geq \eta n$ separately, where $\eta = \frac{1}{4}\varepsilon_0$. For $|R| < \eta n$, we will use

$$\lambda(R, T) \leq \sum_{v \in R} d(v).$$

For this, we may assume that if $|R| = \sigma n$ then $R$ contains the $\sigma n$ vertices of largest degree in $G$ (and that they have the same degrees in $K$). Using the argument about the degrees of $G \in \mathcal{G}(n,p)$ as in Case 3, it is straightforward to show that a.a.s. these vertices have total degree at most

$$\begin{aligned} c\sigma n + O(\sqrt{c}\,n) &< \big(k + 2\sqrt{k \log k}\big)|R| + O(\sqrt{k}\,n) \\ &< (1-\varepsilon)\sqrt{k \log k}\,|T| + k|R| \end{aligned}$$

for $k$ sufficiently large and some $\varepsilon > 0$, since $|R| < \eta n \leq \frac{1}{4}|T|$ and $|T| \geq \varepsilon_0 n$. This gives (15).

It only remains to treat those sets $R$ for which $|R| \geq \eta n$. Then using the same argument as with $T$ in Case 3, the sum of degrees of vertices in $R$ is a.a.s. at most $\big(k+(1+\eta)\sqrt{k \log k}\big)|R|$.

For a set $R$ of this size in $\mathcal{G}(n,p)$, the expected value of $\lambda(R)$ is

$$\binom{\eta n}{2}\frac{c}{n} \sim \frac{1}{2}cn\eta^2 > \frac{1}{2}kn\eta^2$$

since $c > k$. Moreover, $\lambda(R)$ is binomially distributed. So by Chernoff's inequality (see for example [9, Theorem 2.1]),

$$\mathbf{P}\Big(\lambda(R) \leq \frac{kn\eta^2}{4}\Big) \leq e^{-kn\eta^2/16} = o(2^{-n})$$

for sufficiently large $k$ (recall that $\eta$ is an absolute constant). Hence, a.a.s. every set $R$ that is this large induces a subgraph of at least $\frac{1}{4}kn\eta^2 \geq \frac{1}{4}|R|k\eta^2$ edges. Provided $R \subseteq V(K)$, it contains exactly the same number of edges in $K$ as in $\mathcal{G}(n,p)$. Hence we have that a.a.s.

for all such $R$ and $T$,

$$\begin{aligned}\lambda(R,T) &\leq \sum_{v\in R} d(v) - \frac{1}{2}|R|k\eta^2 \\ &\leq \big(k + (1+\eta)\sqrt{k\log k}\big)|R| - \frac{1}{2}|R|k\eta^2 \\ &\leq k|R|\end{aligned}$$

for large enough $k$. This gives (15), as required. ∎

# 5 Proof of Lemma 1

A weakened version of the main result in Pittel, Spencer and Wormald [11] is that if $c$ is fixed, $\mathcal{G}(n, c/n)$ a.a.s. has no $k$-core if $c < c_k$, a.a.s. has one if $c > c_k$, where $c_k$ is defined in (1). A little calculation shows that $c_k$ and $\lambda_k$ ($\lambda_k$ is also defined in (1)) satisfy

$$c_k\pi_k(\lambda_k) = \lambda_k \tag{16}$$

$$c_k = (k-2)!e^{\lambda_k}\lambda_k^{-(k-2)} \tag{17}$$

with $\pi_k$ defined in (2). Substituting (17) and (2) into (16) gives

$$\lambda_k = \sum_{j\geq 0} \frac{{\lambda_k}^{j+1}}{[k+j-1]_{j+1}}$$

(where square brackets denote falling factorials) and so, multiplying by $(k-1)/\lambda_k$, we obtain

$$k-1 = \sum_{j\geq 1} \frac{{\lambda_k}^{j}}{[k+j-1]_{j}}.$$

Since the right hand side is an increasing function of $\lambda_k$, the value of $\lambda_k$ is uniquely determined. Moreover, since $(k+j)/\lambda_k$ is exactly the ratio of the $j$th to the $(j+1)$th term in the summation, the largest term in the summation occurs for $k+j \approx \lambda_k$ and from elementary considerations it is easy to see that $\lambda_k = k + O(\sqrt{k\log k})$. Thus, putting

$$\lambda_k = (k-2)(1+t), \tag{18}$$

we know that $t = o(1)$. In addition, rewriting (2) as

$$\pi_k = 1 - \sum_{j\leq k-2} \frac{e^{-\lambda}\lambda^j}{j!}, \tag{19}$$

we now see that $\pi_k = 1 - o(1)$, and hence also $c_k \sim k$.

To get a slightly better bound on $t$ straight away, substitute (17) into (16), use Stirling's formula with its correction term due to Robbins: $j! = (j/e)^j \sqrt{2\pi j}(1 + O(1/j))$, and take logarithms to give

$$\log \pi_k = \frac{1}{2} \log\Big(\frac{k-2}{2\pi}\Big) + (k-1)\log(1+t) - (k-2)t + O\Big(\frac{1}{k}\Big). \tag{20}$$

Recalling from above that $\log \pi_k = o(1)$ and $t = o(1)$, we may expand $\log(1+t)$ to show that

$$t \sim \left(\frac{q_k}{k}\right)^{1/2} \tag{21}$$

where $q_k = \log k - \log(2\pi)$.

Taking out a factor of $1/c_k$ from the terms in the summation in (19), using (17) we obtain

$$\pi_k = 1 - \frac{1}{c_k} \sum_{m=0}^{k-2} (1+t)^{-m} \left(\frac{k-1}{k-2}\right)^m \prod_{j=1}^{m} \left(1 - \frac{j}{k-2}\right).$$

The terms in the summation are monotonically decreasing. Since $(1+t)^{-m} = \exp(-mt + O(mt^2))$, we see that, for any $\varepsilon > 0$, the terms for $m > k^{1/2+\varepsilon}$ sum to $o(1/k)$. For $m = O(k^{1/2+\varepsilon})$, we see after expanding that the product over $j$ is

$$e^{-m^2/2k + O(m/k + m^3/k^2)} = 1 + O\Big(\frac{m^2}{k} + \frac{m^3}{k^2}\Big).$$

Putting $r = \log(1+t)$ and recalling $c \sim k$, we now have

$$\pi_k \;=\; 1 \;-\; \frac{1}{c_k} \sum_{m=0}^{k^{1/2+\varepsilon}} e^{-mr}\Big(1 + O\Big(\frac{m^2}{k} + \frac{m^3}{k^2}\Big)\Big) \;+\; o\Big(\frac{1}{k^2}\Big).$$

To estimate the first error term we approximate the summation by an integral, so that term becomes

$$O(1) \cdot \sum_{m=0}^{k^{1/2+\varepsilon}} e^{-mr} \frac{m^2}{k} \;=\; O(k^{-1}) \int_0^\infty e^{-rx} x^2 \, dx,$$

which is $O(r^{-2}k^{-1}) = O(t^{-2}k^{-1}) = O(1/\log k)$ using (21). The other error term is similarly $O(1/\log k)$. The main term in the summation is a truncated geometric series with the truncated terms negligible, so we have

$$\begin{aligned} \pi_k &= 1 - \frac{1}{c_k}\left(o(1) + \sum_{m=0}^{\infty} e^{-mr}\right) = 1 - \frac{1}{c_k(1-e^{-r})} + O\Big(\frac{1}{k \log k}\Big) \\ &= 1 - \frac{1}{c_k(1-(t+1)^{-1})} + O\Big(\frac{1}{k \log k}\Big) = 1 - \frac{t+1}{c_k t} + O\Big(\frac{1}{k \log k}\Big). \end{aligned}$$

Using this with (16) shows that

$$c_k = \lambda_k + t^{-1} + 1 + O\Big(\frac{1}{\log k}\Big). \tag{22}$$

We thus continue with

$$\pi_k = 1 - \frac{t+1}{\lambda_k t} + O\Big(\frac{1}{k\log k}\Big) = 1 - \frac{1}{kt} + O\Big(\frac{1}{k\log k}\Big).$$

So we may substitute $\log \pi_k = -1/(kt) + O\big(1/(k\log k)\big)$ in the left side of (20), and $t = (1+x)(q_k/k)^{1/2}$ into the right side. We know that $x = o(1)$ from (21), and we may expand $\log(1+t)$ as $t - \frac{1}{2}t^2 + \frac{1}{3}t^3 + O(t^4)$. The upshot is that

$$t = \Big(\frac{q_k}{k}\Big)^{1/2} + \frac{q_k+3}{3k} + O\Big(\frac{1}{k\log k}\Big).$$

This determines $t$, and since $\lambda_k \sim k$, we have from (18) that $\lambda_k = k + kt - 2 + O(1/\log k)$. Now using (22) and the formula for $t$ immediately above (which in particular gives $1/t = \sqrt{k/q_k} - 1/3 + O(1/\log k)$), we obtain Lemma 1. ∎

**Acknowledgement.** We would like to thank Mike Molloy and Siu On Chan for some comments helping us to correct and shorten our proofs.

## References

[1] B. Bollobás, *Extremal Graph Theory*, London Math. Soc. Monographs No. 11, Academic, London (1978).

[2] B. Bollobás, C. Cooper, T. Fenner, and A. Frieze, Edge disjoint Hamilton cycles in sparse random graphs of minimum degree at least $k$, *J. Graph Theory* **34** (2000), no. 1, 42–59.

[3] B. Bollobás, J.H. Kim, and J. Verstraëte, Regular subgraphs of random graphs, *Random Structures & Algorithms*, **29** (2006), 1–13.

[4] I. Benjamini, G. Kozma, and N. Wormald, The mixing time of the giant component of a random graph, *Preprint.*

[5] J. Cain and N. Wormald, Encores on cores, *Electronic Journal of Combinatorics* **13** (2006), RP 81.

[6] P. Flajolet, D. Knuth, and B. Pittel, The first cycles in an evolving graph. Graph theory and combinatorics (Cambridge, 1988). *Discrete Math.* **75** (1989), no. 1-3, 167–215.

[7] S. Janson, Cycles and unicyclic components in random graphs, *Combin. Probab. Comput.* **12** (2003), no. 1, 27–52.

[8] S. Janson and M. Luczak, A simple solution to the $k$-core problem, Tech. Report 2005:31, Uppsala.

[9] S. Janson, T. Łuczak, and A. Ruciński, *Random Graphs*, Wiley, New York, 2000.

[10] J.H. Kim, Poisson cloning model for random graphs, *International Congress of Mathematicians*, Vol. III, 873–897, Eur. Math. Soc., Zürich, 2006.

[11] B. Pittel, J. Spencer, and N. Wormald, Sudden emergence of a giant $k$-core in a random graph, *J. Combinatorial Theory, Series B* **67** (1996), 111–151.

[12] M. Pretti and M. Weigt, Sudden emergence of $q$-regular subgraphs in random graphs, *Europhys. Lett.* 75, 8 (2006).

[13] W. Tutte, A short proof of the factor theorem for finite graphs, *Canadian J. Math.* **6** (1954), 347–352.