On the threshold for k-regular subgraphs of random graphs

Paweł Prałat
Department of Mathematics and Statistics
Dalhousie University
Halifax NS, Canada

Jacques Verstraëte
Department of Mathematics and Statistics
McGill University
Montreal QC, Canada

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 $(k+2)$-core of a random graph $\mathcal{G}(n,p)$ asymptotically almost surely has a spanning k-regular subgraph. Thus the threshold for the appearance of a k-regular subgraph of a random graph is at most the threshold for the $(k+2)$-core. In particular, this pins down the point of appearance of a k-regular subgraph in $\mathcal{G}(n,p)$ to a window for p of width roughly $2/n$ for large n and moderately large k.

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 [6] and Flajolet, Knuth and Pittel [5]. By analysing the vertex deletion algorithm for the Erdős-Rényi model $\mathcal{G}(n,p)$ of random graphs, Pittel, Spencer and Wormald [12] proved that for fixed $k \geq 3$, there exists a constant c_k such that c_k/n is a sharp threshold for a k-core in $\mathcal{G}(n,p)$. (When discussing thresholds of k-cores and k-regular subgraphs, we mean thresholds for nonempty k-cores and nonempty k-regular subgraphs.) Here

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

*Supported by the Canada Research Chairs Program and NSERC

where $\pi_k(\lambda)$ is defined by

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

and λ_k is the positive number minimising the right hand side of (1). Recently, a number of simpler proofs establishing the threshold c_k/n for the k-core have been published (see Kim [9], Cain and Wormald [4], and Janson and Luczak [7]).

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 is contained in the k-core, 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 [2] 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 $c \geq c_4$, 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 [13] 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 [2, Conjecture 1.3]. For $k = 3$, these thresholds differ, as shown using the first moment method in [2].

In this paper, we improve the window of the threshold for k-regular subgraphs in $\mathcal{G}(n,p)$ by proving Theorem 1 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 $k \geq k_0$, the $(k+2)$-core of a random graph $\mathcal{G}(n,p)$ is nonempty and contains a k-factor or is k-factor-critical a.a.s.*

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 [1] 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 [12]. 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. It follows immediately from this lemma that

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

Hence, combining the lemma and Theorem 1, 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 questions remain: (1) to determine whether there is a sharp threshold for the appearance of a k-regular subgraph, and (2) whether 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.

Throughout the paper, we denote by $\mathcal{G}(n,p)$ the Erdős-Rényi model of random graphs. If G is a graph with vertex set $V(G)$, then $\lambda(S,T)$ 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 $\lambda(S)$ instead. The number of components of a graph G is denoted by $\omega(G)$.

2 Factors of Graphs

In this section, we allow graphs to contain multiple edges. Let G be a graph 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. In particular, a *1-factor* of G is a perfect matching of G. Tutte's 1-Factor Theorem gives the following necessary and sufficient condition for a graph G to have a 1-factor.

Theorem 2. *Let G be a graph, and let $\mathrm{o}(G)$ denote the number of components of G with an odd number of vertices. Then G has a 1-factor if and only if*

$$\mathrm{o}(G - X) \leq |X| \tag{3}$$

for every set $X \subseteq V(G)$.

Using Tutte's 1-Factor Theorem applied to a new graph, $\phi(G)$, a necessary and sufficient condition can be found for a graph to have a k-factor (this is a special case of Tutte's f-factor theorem; see Lovász and Plummer [10] for details). To construct $\phi(G)$, let $V(G) = \{v_1, v_2, \ldots, v_n\}$, and let $V(\phi(G)) = U \cup V$, where U and V are disjoint sets and V is partitioned into independent sets $(V_1, V_2, \ldots, V_n)$ with $|V_i| = d(v_i)$ and U is partitioned into sets $(U_1, U_2, \ldots, U_n)$ such that $|U_i| = k$. Then $\phi(G)$ consists of all edges between U_i and V_i for $i \in \{1, 2, \ldots, n\}$, together with a matching on V such that when we contract all the independent sets V_i in $\phi(G) - U$ to single vertices, we obtain G. An example is shown below in Figure 1, where G is a quadrilateral, $k = 2$, and a 1-factor in $\phi(G)$ corresponds to a 2-factor in G.

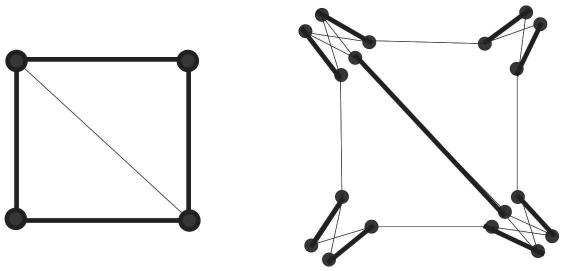

Figure 1 : 2-factor in G and corresponding 1-factor in $\phi(G)$.

It is not hard to see that G has a k-factor if and only if $\phi(G)$ has a 1-factor. We will show that the following condition on G is enough to guarantee a k-factor in G. It is convenient to define

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

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

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

Then G has a k-factor or is k-factor critical according as $\delta_k(G) = 0$ or $\delta_k(G) = 1$.

Proof. We first consider the case $\delta_k(G) = 0$. Let $H = \phi(G)$. To show that G has a k-factor, it is sufficient to show that for all $X \subseteq V(H)$, $\mathrm{o}(H - X) \leq |X|$, by (3). If $X = \emptyset$, then this follows from the fact that H is connected and

$$|V(H)| = k|V(G)| + \sum_{i=1}^{n} d(v_i) \equiv k|V(G)| \equiv 0 \bmod 2.$$

In what remains, we verify (3) for $X \neq \emptyset$. Suppose that for some i, $0 < |X \cap U_i| < |U_i|$. Since U_i and V_i form a complete bipartite graph, we may delete one of the vertices of U_i from X, and the number of components of $H - X$ does not change. Then the right side of (3) decreases, and the left side decreases by at most 1. Hence we assume $X \cap U_i = \emptyset$ or $X \cap U_i = U_i$ for all i.

Define the following sets of vertices of G:

$$S = \{v_i \in V(G) : X \cap U_i = U_i\}, \quad T = \{v_i \in V(G) : X \cap V_i = V_i\},$$

and the following sets of vertices of H:

$$Y = \bigcup_{i=1}^{n} \{V_i - X : v_i \in S\}, \quad X_0 = \bigcup_{i=1}^{n} \{X \cap V_i : v_i \notin T\}, \quad X_1 = \bigcup_{i=1}^{n} \{V_i : v_i \in T\}.$$

Suppose that for some i, $v_i \in S \cap T$ (that is, $X \cap U_i = U_i$ and $X \cap V_i = V_i$). But then we may delete all the vertices of U_i from X, and both the left side and the right side of (3) decreases by k. Hence we assume $S \cap T = \emptyset$.

For convenience, given a subgraph F of $\phi(G)$, we write $\phi^{-1}(F)$ for the subgraph of G obtained from F by deleting all vertices of U and contracting all the sets $V_i \cap V(F)$. For each component F of $H - X$ containing a vertex of $U = \bigcup_{i \leq n} U_i$, either F is an isolated vertex in U, or $\phi^{-1}(F) - S$ is a component of $G - (S \cup T)$. So at most $k|T| + \omega\big(G - (S \cup T)\big)$ components of $H - X$ contain a vertex of U. Now let F be a component of $H - X$ containing no vertices of U. Then $|V(F)| \leq 2$, so the only components of $H - X$ containing no vertices of U and contributing to $\mathrm{o}(H - X)$ are isolated vertices – and these are vertices of Y which are adjacent to a vertex of $X = X_0 \cup X_1$. So the number of these isolated vertices is:

$$\lambda_H(Y, X_0 \cup X_1) = \lambda_H(Y, X_0) + \lambda_H(Y, X_1) = \lambda_H(Y, X_0) + \lambda(S, T) \leq |X_0| + \lambda(S, T)$$

where $\lambda_H(Y, X)$ is the number of edges between X and Y in H. It now follows using (4) that

$$\begin{aligned} \mathrm{o}(H - X) &\leq \omega(G - (S \cup T)) + k|T| + \lambda(S, T) + |X_0| \\ &\leq \sum_{v \in T} d(v) + k|S| + |X_0| \;=\; |X_1| + k|S| + |X_0| \;=\; |X|. \end{aligned}$$

So we have shown $|X| \geq \mathrm{o}(H - X)$ for every set $X \subseteq V(H)$, as required. This completes the proof for $\delta_k(G) = 0$.

Finally, suppose $\delta_k(G) = 1$ and let $v \in V(G)$. We show that $G - \{v\}$ has a k-factor. Let G' be the graph obtained by adding a vertex v' to G and joining v to v' with k parallel edges. Then G' is connected, and $\delta_k(G') = 0$. Furthermore, it is straightforward to check that (4) is satisfied in G'. By the first part of the proof, G' has a k-factor. Deleting both v and v' from this k-factor of G', we get a k-factor in $G - \{v\}$, as required. ∎

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, ∂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 [3, Lemma 5.3]. (That lemma concerns graphs with a given degree sequence, all degrees between 3 and $n^{0.02}$. See the proof of [3, Theorem 4.2] to find the connection with the following.)

Lemma 3. *There is a positive constant γ 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, γ denotes the constant appearing in Lemma 3.

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. We first show that a.a.s. there are no sets of $2y$, $y \le s(n)$ vertices in $\mathcal{G}(n,p)$ inducing at least $2y+1$ edges: the expected number of subgraphs of $2y$ vertices with at least $2y+1$ edges, for some $y \le s(n)$, is at most

$$\begin{aligned}\sum_{y \le s(n)} \binom{n}{2y} \binom{\binom{2y}{2}}{2y+1} p^{2y+1} &< \sum_{y \le s(n)} \frac{(\log n)^{2y+1}}{n} \\ &< n^{-1/2}.\end{aligned}$$

So the claim follows from Markov's inequality. Thus, we may assume that all sets of $2y \le 2s(n)$ vertices in K induce at most $2y$ edges. We may also assume that a.a.s. the property in Lemma 3 holds.

It suffices to show that if X is the vertex set of a union of components of $K - Y$ and $|X| \le \frac{1}{2}|V(K)|$, then $|X| < s(n)$. Suppose $|X| \ge s(n)$. From the property in Lemma 3,

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

Suppose $|Y| = y$. By averaging, some $Z \subset X$ of size $|Y| \le s(n)$ satisfies

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

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

In fact Łuczak [11] showed that the k-core is k-connected a.a.s., as stated in the next lemma.

Lemma 5. *For $k \ge 3$ and $c > c_k$, the k-core of $\mathcal{G}(n, c/n)$ is k-connected a.a.s.*

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

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

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

Proof. Let $\varepsilon > 0$, and fix k, and $j \ge k+2$. From [4, Corollary 3 and Erratum], if n_j denotes the number of vertices of degree j in $\mathcal{K}$, 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 μ 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 anthing larger than c is substituted for μ 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 ε_1 slightly smaller than ε, 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 $pn = c$ and $c_{k+2} < c$, for $k \geq 3$. To prove Theorem 1, we show that there exists k_0 such that for $k \geq k_0$, (4) holds in K a.a.s. The value of k_0 will not be optimized in the proof to follow. To prove the theorem, we consider a number of cases according to the sizes of the sets S and T in the lemma, where $S \cup T \neq \emptyset$. It is convenient throughout to let $s(n) = \log n / 2ec \log\log n$.

Case 1 $|S| + |T| < s(n)$.

Let $Y = S \cup T$. By Lemma 4, $K - Y$ contains a component with more than $|V(K)| - 2s(n)$ vertices a.a.s. Let $\omega_{\rm s}(K-Y)$ denote the number of components of $K - Y$ of size less than $s(n)$, and let X be the set of vertices in these components. As in the proof of Lemma 4, $\lambda(X) \leq |X|$ and $\lambda(X \cup Y) \leq |X| + |Y|$ a.a.s. However, every vertex of X has degree at least $k+2$, by definition of K, so $\lambda(X) + \lambda(X, Y) \geq (k+1)|X|$ a.a.s. since $\lambda(X) \leq |X|$ holds a.a.s. It follows that

$$\begin{aligned} |X| + |Y| \;\geq\; \lambda(X \cup Y) &= \lambda(X) + \lambda(X,Y) + \lambda(Y) \\ &\geq k|X| + |X| + \lambda(S,T) \\ &\geq k \cdot \omega_{\rm s}(K-Y) + |X| + \lambda(S,T). \end{aligned}$$

Therefore, since $\omega_{\rm s}(K-Y) = \omega(K-Y) - 1$,

$$k \cdot \omega\big(K - (S \cup T)\big) - k + \lambda(S,T) + k|T| \quad \leq \quad |S| + (k+2)|T| - |T|\,.$$

Since $k \geq 3$, and $\omega\big(G - (S \cup T)\big) \geq 1$,

$$\begin{aligned}
\omega\big(K - (S \cup T)\big) + \lambda(S,T) + k|T| &\leq k \cdot \omega\big(K - (S \cup T)\big) - k + 1 + \lambda(S,T) + k|T| \\
&\leq |S| + (k+2)|T| - |T| \\
&\leq k|S| + \sum_{v \in T} d(v) - (k-1)|S| - |T| + 1 \\
&\leq k|S| + \sum_{v \in T} d(v) - 2|S| - |T| + 1 \,.
\end{aligned}$$

If $S \neq \emptyset$ or $|T| > 1$, then this is less than $k|S| + \sum_{v \in T} d(v)$, as required. If $S = \emptyset$ and $|T| = 1$, then $\omega(K - T) = 1$ a.a.s. (see Lemma 5), which implies

$$\omega(K - T) + k|T| = 1 + k < 2 + k \leq \sum_{v \in T} d(v).$$

Therefore (4) holds a.a.s. in K, so K has a k-factor or is k-factor critical, by Lemma 2.

For the rest of the proof, ε_0 is an absolute constant; we will take $\varepsilon_0 = e^{-9}$ for definiteness.

Case 2 $|S| + |T| \geq s(n)$, $|T| < \varepsilon_0 n$, and $|S| < 4\varepsilon_0 n$.

Let $Y = S \cup T$. In this case we estimate $\omega(K - Y)$ and $\lambda(S,T)$ separately. First we show that $\omega(K - Y) \leq |Y|/2$ a.a.s. provided that k is large enough to ensure that $\gamma(k+2) \geq 4$ (so this tells us $k_0 \geq 4/(\gamma - 2)$ is required 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 3 shows $\lambda(X,Y) \geq \gamma(k+2)|X|$. Let

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

The expected number of pairs of sets (X,Y) in $\mathcal{G}(n,p)$ satisfying the above requirements is

$$\begin{aligned}
\sum_{I_y\times I}\binom{n}{x+y}\binom{x+y}{x}\binom{xy}{\gamma(k+2)x}p^{\gamma(k+2)x} &< \sum_{I_y\times I} e^{3x+\gamma(k+2)x}2^{x+y}\left(\frac{n}{x}\right)^{3x}\left(\frac{y}{n}\right)^{\gamma(k+2)x}\\
&< \sum_{I_y\times I} e^{3\gamma(k+2)x/2}2^{3x}\left(\frac{y^2}{xn}\right)^{\gamma(k+2)x/2}\\
&< \sum_{I_y\times I} e^{-\gamma(k+2)x}e^{3x} \qquad \text{(since } y\le 5\varepsilon_0 n\text{)}\\
&\le \sum_{I_y\times I} e^{-x} \qquad \text{(since } \gamma(k+2)\ge 4\text{)}\\
&< \int_{I_y\times I} e^{-(x-1)}dxdy\\
&< \int_I e^{-(y-2)/2}dy\\
&< e^{-(s(n)-2)/2} \;=\; o(1)\,,
\end{aligned}$$

where the sums are over $(x,y)\in I_y\times I$. So in fact the expected number of sets X and Y as described above is $o(1)$. By Markov's Inequality, we conclude that $\omega(K-Y)\le |Y|/2$ a.a.s.

To finish verifying inequality (4) in case 2, it remains to show that a.a.s.

$$\lambda(S,T) < \frac{3}{2}|T| + \left(k-\frac{1}{2}\right)|S|. \tag{8}$$

Let $|S|=\sigma n$ and $|T|=\tau n$ and let $\rho=(k-\frac{1}{2})\sigma+\frac{3}{2}\tau$ (here σ and τ are allowed to depend on n). Then the number of ways of choosing the sets S and T is bounded above by

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

The probability that there are at least ρn edges between S and T is at most

$$\begin{aligned}
\binom{\sigma\tau n^2}{\rho n}\left(\frac{c}{n}\right)^{\rho n} &\le \left(\frac{e\sigma\tau c}{\rho}\right)^{\rho n}\\
&= \left(\frac{ec}{(k-1/2)/\tau+3/2\sigma}\right)^{(k-1/2)\sigma n+3\tau n/2}.
\end{aligned}$$

Multiplying by the bound on the number of choices of S and T, the bound (8) is true a.a.s. if

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

for which it suffices that

$$\frac{ec}{(k-1/2)/\tau+3/2\sigma} < \min\left\{\left(\frac{\sigma}{e}\right)^{1/(k-1/2)}, \left(\frac{\tau}{e}\right)^{2/3}\right\}. \tag{9}$$

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

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

To prove that the expression on the left in (9) is less than $(\sigma/e)^{1/(k-1/2)}$, we can assume

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

It follows that $\sigma < e^{-4k}$ and therefore, since $2k - 1 \geq 2$,

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

Case 3 $|S| + |T| \geq s(n)$ and $|T| < \varepsilon_0 n$ and $|S| \geq 4\varepsilon_0 n$.

To prove that the requirements of Lemma 2 are satisfied a.a.s. for k sufficiently large, it is enough to show that a.a.s. for every pair of sets (S, T) under consideration, $\lambda(S, T) \leq \frac{3}{4}k|S|$. This is because the sum of vertex degrees in T is at least $k|T|$, and also $\omega\big(G - (S \cup T)\big) \leq n < \frac{1}{4}k|S|$ for large enough k (as ε_0 is fixed). We may assume that $c \sim c_k$, and so if k is sufficiently large, Lemma 1 shows that $c < \frac{3}{2}k$. Thus it suffices to show $\lambda(S, T) \leq 2c\varepsilon_0 n$, since this is at most $\frac{1}{2}c|S|$. This follows immediately once we show 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 $e^{-c}c^j/j!$. It follows that the sum of the degrees of those vertices of degree less than $3c/2$ is a.a.s. asymptotic to

$$n \sum_{j<3c/2} j\frac{e^{-c}c^j}{j!}.$$

Since the Poisson distribution is asymptotically normal with mean c, for large enough c (i.e. large enough k) we have

$$\sum_{j<3c/2} j\frac{e^{-c}c^j}{j!} > c - \varepsilon_0.$$

Since the sum of vertex degrees is a.a.s. asymptotic to cn, the ones of degree at least $3c/2$ a.a.s. have total degree less than $\varepsilon_0 n$. Assuming this is true, there are at most $2\varepsilon_0 n/3c$ such vertices, and any set of at most $\varepsilon_0 n$ vertices thus has total degree at most

$$\varepsilon_0 n + \frac{3c}{2}\varepsilon_0 n < 2c\varepsilon_0 n$$

as required.

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

By Lemmas 1 and 6, for all $\varepsilon > 0$, if k is sufficiently large, then a.a.s.

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

So by Lemma 2 and the fact that $n = O(|T|)$, it is enough to show, for some $\varepsilon > 0$, that a.a.s.

$$(1-\varepsilon)\sqrt{k\log k}\,|T| + k|S| \geq \lambda(S,T). \tag{10}$$

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

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

For this, we may assume that if $|S| = \sigma n$ then S contains the σn 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

$$\begin{aligned} c\sigma n + O(\sqrt{c}\,n) &< \big(k + 2\sqrt{k\log k}\big)|S| + O(\sqrt{k}\,n) \\ &< k|S| + 3\sqrt{k\log k}\,|S| \end{aligned}$$

for k sufficiently large, which is at most

$$(1-\varepsilon)\sqrt{k\log k}\,|T| + k|S|$$

since $|S| < \eta n \leq \frac{1}{4}|T|$. This gives (10).

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

For a set S of this size in $\mathcal{G}(n,p)$, the expected value of $\lambda(S)$ 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, $\lambda(S)$ is binomially distributed. So by Chernoff's inequality (see for example [8, Theorem 2.1]),

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

for sufficiently large k (recall that η is an absolute constant). Hence, a.a.s. every set S that is this large induces a subgraph of at least $\frac{1}{4}kn\eta^2 \geq \frac{1}{4}|S|k\eta^2$ edges. Provided $S \subseteq V(K)$, 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 S and T,

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

for large enough k. This gives (10), as required. ∎

5 Proof of Lemma 1

A weakened version of the main result in Pittel, Spencer and Wormald [12] 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 λ_k (λ_k is also defined in (1)) satisfy

$$c_k \pi_k(\lambda_k) = \lambda_k \tag{11}$$

$$c_k = (k-2)! e^{\lambda_k} \lambda_k^{-(k-2)} \tag{12}$$

with π_k defined in (2). Substituting (12) and (2) into (11) gives

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

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

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

Since the right hand side is an increasing function of λ_k, the value of λ_k is uniquely determined. Moreover, since $(k+j)/\lambda_k$ is exactly the ratio of the jth 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

$$\lambda_k = (k-2)(1+t), \tag{13}$$

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

$$\pi_k = 1 - \sum_{j \le k-2} \frac{e^{-\lambda} \lambda^j}{j!}, \tag{14}$$

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 (12) into (11), 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

$$\log \pi_k = \frac{1}{2} \log\Big(\frac{k-2}{2\pi}\Big) + (k-1)\log(1+t) - (k-2)t + O\Big(\frac{1}{k}\Big). \tag{15}$$

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

$$t \sim \left(\frac{q_k}{k}\right)^{1/2} \tag{16}$$

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

Taking out a factor of $1/c_k$ from the terms in the summation in (14), using (12) 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

$$\begin{aligned} O(1)\cdot\sum_{m=0}^{k^{1/2+\varepsilon}} e^{-mr}\frac{m^2}{k} &= O(k^{-1})\int_0^\infty e^{-rx}x^2\,dx \\ &= O(r^{-2}k^{-1}) \\ &= O(t^{-2}k^{-1}) \\ &= O\Big(\frac{1}{\log k}\Big) \end{aligned}$$

using (16). 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 (11) shows that

$$c_k = \lambda_k + t^{-1} + 1 + O\Big(\frac{1}{\log k}\Big). \tag{17}$$

So we may continue with

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

Hence

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

Next, substitute this in the left side of (15), and $t = (1+x)(q_k/k)^{1/2}$ into the right side. We know that $x = o(1)$ from (16), 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 (13) that $\lambda_k = k + kt - 2 + O(1/\log k)$. Now using (17) and the formula for t immediately above (which in particular gives $1/t = (q_k/k)^{1/2} - 1/3$), we obtain Lemma 1. ∎

References

[1] 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.

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

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

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

[5] 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.

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

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

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

[9] J. H. Kim, Poisson cloning model for random graphs. *Preprint*.

[10] L. Lovász and M.D. Plummer, Matching theory. North-Holland Mathematics Studies, 121. *Annals of Discrete Mathematics*, 29. Budapest, 1986, xxvii+544 pp.

[11] T. Łuczak, Size and connectivity of the k-core of a random graph. *Discrete Math.* **91** (1991), no. 1, 61–68.

[12] 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.

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