$(k+1)$-cores have k-factors

Siu On Chan* Michael Molloy*

June 11, 2012

Abstract

We prove that the threshold for the appearance of a k-regular subgraph in $G_{n,p}$ is at most the threshold for the appearance of a non-empty $(k+1)$-core. This improves a result of Pralat, Verstraete and Wormald [5] and proves a conjecture of Bollobás, Kim and Verstraete [3].

1 Introduction

This paper concerns k-regular subgraphs of random graphs. A natural starting point for such a study is with the k-core; i.e. the unique maximal subgraph with minimum degree at least k. Pittel, Spencer and Wormald [4] determined the threshold $c_k = k + \sqrt{k \log k} + o(\sqrt{k})$ for the appearance of a non-empty k-core in $G_{n,c/n}$, the random graph with n vertices where each of the $\binom{n}{2}$ possible edges appears independently with probability c/n. So for $c < c_k$, w.h.p.[1] $G_{n,c/n}$ has no non-empty k-core and hence w.h.p. has no k-regular subgraph.

In [3], Bollobás, Kim and Verstraete proved that $G_{n,c/n}$ w.h.p. contains a k-regular subgraph if $c > \rho_k$ for a specific function $\rho_k = 4k + o(k)$; note that $\rho_k \approx 4c_k$. They also proved that the threshold for the appearance of a 3-regular subgraph is strictly larger than c_3 and conjectured that the threshold for the appearance of a k-regular subgraph is strictly larger than c_k for all $k \geq 4$. Pretti and Weigt [6] used techniques from statistical physics to predict the contrary: for every $k \geq 4$, the threshold for the appearance of a k-regular subgraph is equal to c_k. Those conflicting conjectures remain unresolved.

Bollobás, Kim and Verstraete also conjectured that if $c > c_{k+1}$ then w.h.p. the $(k+1)$-core of $G_{n,c/n}$ has a k-regular subgraph (see Conjecture 1.3 from [3]). We prove this conjecture here for k sufficiently large.

*Dept of Computer Science, University of Toronto, {siuon,molloy}@cs.toronto.edu.

[1] A property holds *with high probability (w.h.p.)* if it holds with probability tending to 1 as $n \to \infty$.

A *k-factor* of a graph G is a spanning k-regular subgraph; note that if G has a k-factor, then $k \times |V(G)|$ must be even. G is said to be *k-factor-critical* if for every $v \in V(G)$, $G - v$ has a k-factor. Suppose $c_{k+2} < c < c_{k+2} + 2\sqrt{k \log k}$ and let C denote the $(k+2)$-core of $G_{n,c/n}$. Pralat, Verstraete and Wormald [5] proved that if k is sufficiently large then w.h.p.: (i) if $k \times |V(C)|$ is even then C contains a k-factor; (ii) if $k \times |V(C)|$ is odd then C is k-factor-critical. We extend this result to the $(k+1)$-core:

Theorem 1.1. *There is an absolute constant k_0 such that for all $k \geq k_0$, and for any $c_{k+1} < c < c_{k+1} + 2\sqrt{k \log k}$, w.h.p. the $(k+1)$-core, K, of $G_{n,c/n}$ satisfies:*

(a) if $k \times |V(K)|$ is even then K has a k-factor;

(b) if $k \times |V(K)|$ is odd then K is k-factor-critical.

This result is best possible (for large k) in that, as observed in [5], for every $c > c_k$ w.h.p. the k-core of $G_{n,c/n}$ neither contains a k-factor nor is k-factor-critical, because it w.h.p. contains many vertices of degree greater than k whose neighbours all have degree exactly k.

By monotonicity, Theorem 1.1 implies that for any $c > c_{k+1}$, w.h.p. the $(k+1)$-core of $G_{n,c/n}$ contains a k-regular subgraph, although for very large c we do not guarantee an actual k-factor. This proves the aforementioned conjecture from [3] for $k \geq k_0$. It also establishes that the threshold for the appearance of a k-regular subgraph is at most the threshold for the appearance of a $(k+1)$-core, for large k. It was remarked in [3] that perhaps w.h.p. the $(k+1)$-core of the random graph will contain a k-factor (so long as its size times k is even); Theorem 1.1 confirms this for large k with c not much greater than c_k.

Our proof makes use of Tutte's Factor Theorem [7] (see also Exercise 3.3.29 of [8]). We state it here, in terms of k-factors; Tutte's actual statement applies to more general factors. For disjoint $X, Y \subseteq V(G)$, we use $e(X, Y)$ to denote the number of edges with one endpoint in X and the other in Y. And we use $q(X, Y)$ to denote the number of components Q of $G - (X \cup Y)$ such that $k|Q|$ and $e(Q, Y)$ have different parities.

Theorem 1.2 (Tutte [7]). *A graph G has a k-factor iff for every disjoint pair of sets $R, W \subseteq V(G)$,*

$$k|R| \geq q(R, W) + k|W| - \sum_{v \in W} \deg_{G-R}(v).$$

Rearranging, we see that the condition of Theorem 1.2 is equivalent to:

$$\sum_{v \in W} \deg_G(v) + k|R| \geq q(R, W) + k|W| + e(R, W). \tag{1}$$

To prove Theorem 1.1(a), we will prove that w.h.p. K satisfies a stronger condition. Using $\omega(H)$ to denote the number of components of a subgraph H, we will show that for every disjoint pair of sets $S, T \subseteq V(K)$ with $S \cup T \neq \emptyset$,

$$\sum_{v \in T} \deg_K(v) + k|S| \geq \omega(K - S \cup T) + k|T| + e(S, T). \tag{2}$$

By Theorem 1.2 with $R := S, W := T$, this will suffice to prove Theorem 1.1(a) since $\omega(K - R \cup W) \geq q(R, W)$. (The case $S = T = \emptyset$ for (1) follows w.h.p. from the connectivity of K; see Lemma 2.1 below.)

For part (b), it would suffice to prove that for every disjoint pair of sets $S, T \subseteq V(K)$ with $|S| \geq 1$ and $|S \cup T| \geq 2$, we have:

$$\sum_{v \in T} \deg_K(v) + k|S| \geq \omega(K - S \cup T) + k|T| + e(S, T) + k. \tag{3}$$

It is straightforward to show that if S, T satisfy (3) then for any $x \in S$, (2) holds upon substituting $K := K - x, S := S - x$ (the quick argument appears in the proof of Corollary 2 of [5]). It follows that if (3) were to hold for all S, T with $|S| \geq 1$ and $|S \cup T| \geq 2$ then this would establish Theorem 1.1(b). This was indeed the case in [5] (their (4) is equivalent to our (3)). However, there is a case in our setting where (3) does not hold: Consider a vertex x whose neighbours all have degree $k + 1$ in K. In $K - x$, they all have degree k, and this forces all of their edges into any k-factor. It is easy to verify that $S = \{x\}$ and $T = N(x)$ will violate (3); equivalently, $S = \emptyset$ and $T = N(x)$ will violate (2) when K is replaced by $K - x$. Fortunately $R = \emptyset$ and $W = N(x)$ does not violate (1), with $G = K - x$, and so we can cover this case by dealing directly with (1).

Our proof follows the same outline as that of [5], where they proved the analogue of Theorem 1.1 for the $(k + 2)$-core. Their proof covered four separate cases for the sizes of S, T; our proof follows the same cases (see the proof Lemma 3.1 below). Our analysis of Case 1 is similar to theirs, but we require a somewhat different argument for the setting of this paper. Case 2 is where the main new ideas of this paper are required in order to deal with the $(k+1)$-core rather than the $(k+2)$-core. Their arguments for Cases 3 and 4 apply to the setting of this paper, so we do not need any new ideas there; we combine them into our Case 3. The reader who is already familiar with [5] may want to focus on Case 2 of Lemma 3.1 (in particular, Case 2b).

2 Preliminary Lemmas

Here we will prove that w.h.p. the $(k + 1)$-core, K, of a random graph satisfies certain properties. In the next section, we will prove that any graph with minimum degree $k + 1$ and satisfying those properties either has a k-factor or is k-factor-critical. This proves Theorem 1.1.

The statements of these properties require constants: $\gamma > 0$, which is independent of c, k and comes from Lemma 2.1 below; and $\epsilon_0 > 0$, also independent of c, k, chosen so that $\epsilon_0 < \frac{\gamma^2}{10^5}$. We define:

$$s(n) = \log n/(2ec \log\log n).$$

We now list our properties. Given a set of vertices S, we define $N(S)$ to be the set of vertices not in S that are adjacent to at least one vertex in S.

(P1) For every $Y \subseteq V(K)$ with $|Y| \leq \frac{1}{2}|V(K)|$, we have:

$$e(Y, K - Y) \geq \gamma(k+1)|Y|.$$

(P2) Every $Y \subseteq V(K)$ with $|Y| \leq 4s(n)$ has at most $|Y|$ edges.

(P3) For every $Y \subseteq V(K)$ with $|Y| \leq s(n)$, $K - Y$ contains a component with more than $|V(K)| - 2s(n)$ vertices.

(P4) For every disjoint pair of sets $X, Y \subseteq V(K)$ with $|X| \geq \frac{1}{200}|Y|$ and $|Y| \leq \epsilon_0 n$ we have:

$$e(X, Y) \leq \frac{1}{2}\gamma k|X|.$$

(P5) For every disjoint pair of sets $S, T \subseteq V(K)$ with $s(n) \leq |S| + |T| \leq \epsilon_0 n$ we have:

$$e(S, T) < \frac{101}{100}|T| + \frac{k}{2}|S|.$$

(P6) For every disjoint pair of sets $S, X \subseteq V(K)$ with $|S| \leq \frac{\epsilon_0}{20k}n$ and $|X| \geq |S|$ we have:

$$e(X, (S \cup N(S)) \backslash X) \leq \frac{1}{2}\gamma k|X|.$$

(P7) For every $S \subseteq V(K)$ with $|S| \leq \frac{\epsilon_0}{20k}n$ we have:

$$e(S, N(S)) \leq |N(S)| + \frac{k}{4}|S|.$$

(P8) For every disjoint pair of sets $S, T \subseteq V(K)$ with $|T| < \frac{1}{10}\epsilon_0 n$ and $|S| > \frac{9}{10}\epsilon_0 n$, we have:

$$e(S, T) < \frac{3}{4}k|S|.$$

(P9) For every disjoint pair of sets $S, T \subseteq V(K)$ with $|T| \geq \frac{1}{10}\epsilon_0 n$, we have:

$$e(S, T) \leq k|S| + (1 - \epsilon)\sqrt{k \log k}|T|, \qquad \text{and} \qquad \sum_{v \in T} d(v) > (k + (1 - \frac{\epsilon}{2}\sqrt{k \log k}))|T|.$$

In proving that these properties hold w.h.p. for the $(k+1)$-core, we will often use the following well-known bound which follows easily from Stirling's Inequality:

$$\binom{a}{b} \leq \left(\frac{ea}{b}\right)^b.$$

Recall that the hypothesis of Theorem 1.1 requires $c < c_k + 2\sqrt{k \log k} < 2k$ (for k sufficiently large).

Our first three lemmas below come from [5] (with minor rephrasing).

Lemma 2.1. *(Lemma 2 of [5].) Consider any $k \geq 3$ and $c > c_{k+1}$. There is a constant $\gamma > 0$ (independent of c, k) such that w.h.p. the $(k+1)$-core K of $G_{n,c/n}$ satisfies (P1).*

A standard first moment argument nearly identical to the proof of Lemma 3 of [5] yields:

Lemma 2.2. *For any $c > c_{k+1}$, w.h.p. the $(k+1)$-core K of $G_{n,c/n}$ satisfies (P2).*

(In fact, the argument shows that the entire graph $G_{n,c/n}$ satisfies (P2).) Lemma 4 of [5] says:

Lemma 2.3. *(Lemma 4 of [5].) If k is sufficiently large then for every $c > c_{k+1}$, w.h.p. the $(k+1)$-core K of $G_{n,c/n}$ satisfies (P3).*

We include their brief proof:

Proof: Let X be the union of the vertex sets of some components of $K - Y$, such that $|X| > s(n)$. We will show that if K satisfies (P1) and (P2) then $|X| > \frac{1}{2}|V(K)|$; this implies the lemma.

Consider any $Z \subseteq X$ where $|Z| = |Y|$. Thus $|Y \cup Z| \leq 2s(n)$ and so by (P2) we can assume $e(Y, Z) \leq |Y \cup Z| = 2|Z|$. Averaging over all such $Z \subseteq X$ yields $e(Y, X) \leq 2|X| < \gamma k|X|$, for k sufficiently large (since γ does not depend on k). Since $e(X, K - X) = e(Y, X)$, (P1) implies $|X| > \frac{1}{2}|V(K)|$ as required. $\square$

Lemma 2.4. *For any $2k > c > c_{k+1}$, w.h.p. the $(k+1)$-core K of $G_{n,c/n}$ satisfies (P4).*

Proof This argument is very similar to one in [5] where they prove that a similar bound holds w.h.p.

In fact, we will prove the stronger statement that for every $2k > c > 0$, w.h.p. every disjoint pair of sets X, Y in $G_{n,c/n}$ with $|X| \geq \frac{1}{200}|Y|$ and $|Y| \leq \epsilon_0 n$ satisfies:

$$e(X, Y) \leq \frac{1}{2}\gamma k|X|. \tag{4}$$

Clearly (4) holds for $X = \emptyset$, so we can assume $|X| \geq 1$.

Let $xn = |X|$, and $yn = |Y|$. For any fixed x, y, the expected number of sets X, Y in

$G_{n,c/n}$ that violate (4) is at most:

$$
\begin{aligned}
& \binom{n}{yn}\binom{n}{xn}\binom{(yn)(xn)}{\frac{1}{2}\gamma kxn}\left(\frac{c}{n}\right)^{\frac{1}{2}\gamma kxn} \\
< & \left(\frac{e}{y}\right)^{yn}\left(\frac{e}{x}\right)^{xn}\left(\frac{exyn^2c}{\frac{1}{2}\gamma kxn^2}\right)^{\frac{1}{2}\gamma kxn} \\
< & \left(\frac{e}{y/200}\right)^{201xn}\left(\frac{4ey}{\gamma}\right)^{\frac{1}{2}\gamma kxn} \qquad \text{since } x > \frac{y}{200}, \frac{e}{y/200} > 1 \text{ and } c < 2k \\
< & \left(\frac{3200e^3y}{\gamma^2}\right)^{\frac{1}{4}\gamma kxn} \qquad \text{if } k \text{ is large enough that } 201 < \frac{1}{4}\gamma k \\
< & \left(\frac{1}{2}\right)^{xn} \qquad \text{since } y \le \epsilon_0 < \frac{1}{2}\left(\frac{\gamma^2}{3200e^3}\right) \text{ and } \frac{1}{4}\gamma k > 1.
\end{aligned}
$$

If $|X| = i$ then there are at most $200i$ choices for y, since $s(n) < |Y| \le 200|X|$. Therefore, summing over i we find that the expected number of pairs X, Y violating (4) with $|X| \ge \log n$ is less than:

$$\sum_{i \ge \log n} 200i\left(\frac{1}{2}\right)^i = o(1).$$

For $|X| < \log n$ we have $|Y| < 200\log n$; i.e. $y < \frac{200\log n}{n}$. Thus $\left(\frac{3200e^2y}{\gamma^2}\right)^{\frac{1}{4}\gamma kxn} < \frac{1}{n^3}$ (since we can assume $xn = |X| \ge 1$ and we can choose k_0 to be large enough that, for $k \ge k_0$, $\frac{1}{4}\gamma k \ge 4$). There are fewer than n^2 choices for x, y and so the expected number of pairs X, Y with $|X| < \log n$ that violate (4) is $o(1)$. □

Lemma 2.5. *For any $2k > c > c_{k+1}$, w.h.p. the $(k+1)$-core K of $G_{n,c/n}$ satisfies (P5).*

Proof This argument is very similar to one in [5] where they prove that a similar bound holds w.h.p.

In fact, we will prove the stronger statement that for every $2k > c > 0$, w.h.p. every disjoint pair of sets S, T in $G_{n,c/n}$ with $s(n) \le |S| + |T| \le \epsilon_0 n$ satisfies:

$$e(S,T) < \frac{101}{100}|T| + \frac{k}{2}|S|. \tag{5}$$

Let $\sigma n = |S|$ and $\tau n = |T|$. For any choice of σ, τ, the expected number of such sets S, T

in $G_{n,c/n}$ violating (5) is at most:

$$\binom{n}{\sigma n}\binom{n}{\tau n}\binom{(\sigma n)(\tau n)}{\frac{101}{100}\tau n+\frac{k}{2}\sigma n}\left(\frac{c}{n}\right)^{\frac{101}{100}\tau n+\frac{k}{2}\sigma n} < \left(\frac{e}{\sigma}\right)^{\sigma n}\left(\frac{e}{\tau}\right)^{\tau n}\left(\frac{e\sigma\tau n^2 c}{(\frac{101}{100}\tau n+\frac{k}{2}\sigma n)n}\right)^{\frac{101}{100}\tau n+\frac{k}{2}\sigma n}$$
$$= \left(\frac{e}{\sigma}\right)^{\sigma n}\left(\frac{e}{\tau}\right)^{\tau n}\left(\frac{e\sigma\tau c}{\frac{101}{100}\tau+\frac{k}{2}\sigma}\right)^{\frac{101}{100}\tau n+\frac{k}{2}\sigma n}.$$

Since $c < 2k$ and $\tau < \epsilon_0 < (16e^3)^{-100}$, we have:

$$\frac{e\sigma\tau c}{\frac{101}{100}\tau+\frac{k}{2}\sigma} < \frac{e\sigma\tau c}{\frac{k}{2}\sigma} < 4e\tau < \left(\frac{\tau}{2e}\right)^{\frac{100}{101}}.$$

Furthermore, if $\sigma > e^{-k/3}$ then for k sufficiently large we have:

$$\frac{e\sigma\tau c}{\frac{101}{100}\tau+\frac{k}{2}\sigma} < \left(\frac{\tau}{2e}\right)^{\frac{100}{101}} < e^{-1} < \left(\frac{\sigma}{2e}\right)^{\frac{2}{k}},$$

while if $\sigma \le e^{-k/3}$ then for k sufficiently large we have:

$$\frac{e\sigma\tau c}{\frac{101}{100}\tau+\frac{k}{2}\sigma} < \frac{e\sigma\tau c}{\frac{101}{100}\tau} < ec\sigma^{1/2}\sigma^{1/2} < e(2k)e^{-k/6}\sigma^{1/2} < \sigma^{1/2} < \left(\frac{\sigma}{2e}\right)^{\frac{2}{k}}.$$

This implies that the expected number of pairs S, T with $|S| = \sigma n, |T| = \tau n$ is at most

$$\left(\frac{e}{\sigma}\right)^{\sigma n}\left(\frac{e}{\tau}\right)^{\tau n}\left(\frac{\tau}{2e}\right)^{\frac{100}{101}\frac{101}{100}\tau n}\left(\frac{\sigma}{2e}\right)^{\frac{2}{k}\frac{k}{2}\sigma n} = \left(\frac{1}{2}\right)^{(\sigma+\tau)n} = \left(\frac{1}{2}\right)^{|S|+|T|}.$$

For each choice of $y = |S| + |T|$, there are y choices for $|S|, |T|$. So the expected number of sets S, T violating (5) is at most:

$$\sum_{y=s(n)}^{n} y(\frac{1}{2})^y = o(1).$$

□

Properties (P6) and (P7) are key to the main new ideas required for this paper. Before establishing that they hold w.h.p. we begin with a technical lemma.

Lemma 2.6. *For any $2k > c > c_{k+1}$, w.h.p. every set S in $G_{n,c/n}$ of size at most $\frac{\epsilon_0}{20k}n$ satisfies:*

(a) $\left(\frac{25|S\cup N(S)|}{\gamma n}\right)^{\frac{1}{2}\gamma k}\frac{n^2}{|S|^2} < \frac{1}{2e^2}$.

(b) $\left(\frac{25|N(S)|}{n}\right)^{k/4} \frac{n}{|S|} < \frac{1}{2e}$.

Proof Note that if (a) holds then $\frac{25|S\cup N(S)|}{\gamma n} < 1$ and so

$$\left(\frac{25|N(S)|}{n}\right)^{k/4} \frac{n}{|S|} < \left(\frac{25|S\cup N(S)|}{\gamma n}\right)^{k/4} \frac{n}{|S|} < \left(\frac{25|S\cup N(S)|}{\gamma n}\right)^{\frac{1}{2}\gamma k} \frac{n^2}{|S|^2}\frac{|S|}{n} < \frac{1}{2e^2}\times\frac{\epsilon_0}{20k} < \frac{1}{2e};$$

i.e. (a) implies (b). So we will prove (a).

Let S^* be the $|S|$ vertices of largest degree in $G_{n,c/n}$ and let D be the sum of their degrees. Clearly $|N(S)| \leq D$. For $i \geq 0$, the expected number of vertices of degree i in $G_{n,c/n}$ is $\left(\frac{c^i}{i!}e^{-c} + o(1)\right)n$. Standard methods (eg Lemma 3.10 of [2]) show that this number is concentrated enough that: w.h.p. for all i such that $\frac{c^i}{i!}n \geq \sqrt{n}$ we have (i) at most $\frac{c^i}{i!}n$ vertices have degree i and (ii) at most $\sum_{j\geq i}\frac{c^j}{j!}n$ vertices have degree at least i. Also, it is well-known that the maximum degree in $G_{n,c/n}$ is w.h.p. less than $\log n$ (see e.g. Exercise 3.5 of [2]). We will assume that these w.h.p. properties hold, and show that for every choice of $|S|$, the bound in (a) holds. This establishes our lemma.

Case 1: $|S| \leq n^{2/3}$. Since the maximum degree is less than $\log n$, we have $D \leq |S|\log n$ and so

$$\left(\frac{25|S\cup N(S)|}{\gamma n}\right)^{\frac{1}{2}\gamma k} \frac{n^2}{|S|^2} \leq \left(\frac{25|S|(\log n+1)}{\gamma n}\right)^{\frac{1}{2}\gamma k} \frac{n^2}{|S|^2}.$$

That product increases with $|S|$ and so is at most:

$$\left(\frac{25n^{2/3}(\log n+1)}{\gamma n}\right)^{\frac{1}{2}\gamma k} \frac{n^2}{(n^{2/3})^2} = o(1).$$

For the next two cases, we define I to be the largest integer such that $\frac{c^I}{I!}n \geq |S|$, and $i^* \geq I$ to be the largest integer for which $\frac{c^{i^*}}{i^*!}n \geq \sqrt{n}$. It is easily verified that $\sum_{i>i^*}\frac{c^i}{i!}n < 2\frac{c^{i^*+1}}{(i^*+1)!}n < 2\sqrt{n}$, and so fewer than $2\sqrt{n}$ vertices have degree greater than i^*. Since those vertices all have degree at most $\log n$, we have:

$$D < \sum_{i=I}^{i^*} i\frac{c^i}{i!}n + 2\sqrt{n}\log n.$$

Case 2: $|S| > n^{2/3}$ and $I \geq 4c$. Since $I \geq 4c$, it is easily verified that $\frac{c^I}{I!}n + \sum_{i\geq I} i\frac{c^i}{i!}n < 2I\frac{c^I}{I!}n$. Also, $\frac{c^I}{I!}n \geq |S| > n^{2/3} > 2\sqrt{n}\log n$. So our bound on D above, and the fact that we can take $c > 2$, yields $|S| + D < 2I\frac{c^I}{I!}n + 2\sqrt{n}\log n < 3I\frac{c^I}{I!}n < 3I^2\frac{c^{I+1}}{(I+1)!}n < 3I^2|S|$, by our choice of I. In the next line, we will use the fact, from the previous sentence, that

$|S \cup N(S)| \leq |S| + D$ is at most $3I\frac{c^I}{I!}n$ and at most $3I^2|S|$:

$$\left(\frac{25|S \cup N(S)|}{\gamma n}\right)^{\frac{1}{2}\gamma k} \frac{n^2}{|S|^2} < \left(\frac{25 \times 3I^2}{\gamma}\right)^2 \left(\frac{25 \times 3I\frac{c^I}{I!}}{\gamma}\right)^{\frac{1}{2}\gamma k - 2}.$$

This product is easily seen to decrease as $I \geq 4c$ increases, and so it is at most

$$\left(\frac{25 \times 48c^2}{\gamma}\right)^2 \left(\frac{25 \times 12c\frac{c^{4c}}{(4c)!}}{\gamma}\right)^{\frac{1}{2}\gamma k - 2} < \frac{1}{2e^2},$$

for k (and hence c) sufficiently large.

Case 3: $I < 4c$. $D/|S|$ is monotone decreasing as $|S|$ increases, since increasing $|S|$ adds to S^* vertices of degree at most that of all those already in S^*. Therefore, $|S \cup N(S)|/|S|$ is at most the value $(|S| + D)/|S|$ at $I = 4c$ and $|S| = \frac{c^I}{I!}n$. Applying the analysis from Case 2, at that value of $|S|$ we have $|S| + D \leq 3I\frac{c^I}{I!}n = 3I|S| = 12c|S|$. Therefore, using the facts that $c < 2k$ and $|S| < \frac{\epsilon_0}{20k}n < \frac{\epsilon_0}{10c}n < \frac{\gamma}{25 \times 24c}n$ we have:

$$\left(\frac{25|S \cup N(S)|}{\gamma n}\right)^{\frac{1}{2}\gamma k} \frac{n^2}{|S|^2} < \left(\frac{25 \times 12c}{\gamma}\right)^2 \left(\frac{25 \times 12c|S|}{\gamma n}\right)^{\frac{1}{2}\gamma k - 2} < \left(\frac{25 \times 24k}{\gamma}\right)^2 \left(\frac{1}{2}\right)^{\frac{1}{2}\gamma k - 2} < \frac{1}{2e^2},$$

for k sufficiently large. $\square$

Lemma 2.7. *For any $2k > c > c_{k+1}$, w.h.p. the $(k+1)$-core K of $G_{n,c/n}$ satisfies (P6).*

Proof In fact, we will prove the stronger statement that for every $2k > c > 0$, w.h.p. every disjoint pair of sets S, X in $G_{n,c/n}$ with $|S| \leq \frac{\epsilon_0}{20k}n$ and $|X| \geq |S|$ satisfies:

$$e(X, (S \cup N(S)) \backslash X) \leq \frac{1}{2}\gamma k|X|. \tag{6}$$

We fix $|S| = \sigma n \leq \frac{\epsilon_0}{20k}n$, $|X| = xn \geq \sigma n$ and we let ρ be the solution to $\left(\frac{25\rho}{\gamma}\right)^{\frac{1}{2}\gamma k} \frac{1}{\sigma^2} = \frac{1}{2e^2}$. By Lemma 2.6(a), w.h.p. K is such that for every choice of S we have $|S \cup N(S)| \leq \rho n$.

We will bound the expected number of pairs S, X with $|S \cup N(S)| \leq \rho n$ and $e(X, (S \cup N(S)) \backslash X) > \frac{1}{2}\gamma k|X|$. We first choose S, X, then expose $N(S)$. We assume that $|S \cup N(S)| \leq \rho n$ and bound the probability, under this assumption, that $e(X, (S \cup N(S)) \backslash X) > \frac{1}{2}\gamma k|X|$. This yields a bound of at most:

$$\begin{aligned}
& \binom{n}{\sigma n}\binom{n}{xn}\binom{(\rho n)(xn)}{\frac{1}{2}\gamma kxn}\left(\frac{c}{n}\right)^{\frac{1}{2}\gamma kxn} \\
< \ & \left(\frac{e}{\sigma}\right)^{\sigma n}\left(\frac{e}{x}\right)^{xn}\left(\frac{ec\rho}{\frac{1}{2}\gamma k}\right)^{\frac{1}{2}\gamma kxn} \\
< \ & \left(\frac{e}{\sigma}\right)^{2xn}\left(\frac{25\rho}{\gamma}\right)^{\frac{1}{2}\gamma kxn} \qquad \text{since } \sigma \leq x \text{ and } c < 2k \\
= \ & \left(\frac{1}{2}\right)^{xn} \qquad \text{since } \left(\frac{25\rho}{\gamma}\right)^{\frac{1}{2}\gamma k} = \frac{\sigma^2}{2e^2}.
\end{aligned}$$

The sum of $\left(\frac{1}{2}\right)^{xn}$ over all $i = |X|, j = |S|$ with $|X| \geq \max(|S|, \sqrt{\log n})$ is at most

$$\sum_{i \geq \sqrt{\log n}} \sum_{j \leq i} \left(\frac{1}{2}\right)^i = \sum_{i \geq \sqrt{\log n}} i\left(\frac{1}{2}\right)^i = o(1).$$

Therefore, w.h.p. there are no sets S, X violating (6) with $|X| \geq \sqrt{\log n}$ and with $|S \cup N(S)| \leq \rho n$. By Lemma 2.6(a), this implies that w.h.p. there are no S, X violating (6) with $|X| \geq \sqrt{\log n}$.

For the case where $|S| \leq |X| < \sqrt{\log n}$, let H be the subgraph induced by X, S and the endpoints in $S \cup N(S)$ of more than $\frac{1}{2}\gamma k|X|$ edges from X. It is straightforward to show that H has more edges than vertices: Indeed, if ℓ is the number of vertices of H in $N(S)$, then $|E(H)| > \ell + \frac{1}{2}\gamma k|X| \geq \ell + 2|X|$ and $|V(H)| = \ell + |S| + |X| \leq \ell + 2|X|$. But H has at most $|S| + |X| + \frac{1}{2}\gamma k|X| = O(\sqrt{\log n})$ vertices. So by Lemma 2.2, w.h.p. no such H exists. □

Lemma 2.8. *For any $c > c_{k+1}$, w.h.p. the $(k+1)$-core K of $G_{n,c/n}$ satisfies (P7).*

Proof In fact, we will prove the stronger statement that for every $c > 0$, w.h.p. every set S in $G_{n,c/n}$ with $|S| \leq \frac{\epsilon_0}{20k}n$ satisfies:

$$e(S, N(S)) \leq |N(S)| + \frac{k}{4}|S|. \tag{7}$$

Consider any set S of at most $\frac{\epsilon_0}{20k}n$ vertices. Let $\sigma = |S|/n$. Let ν be the solution to $(25\nu)^{k/4}/\sigma = \frac{1}{2e}$. By Lemma 2.6(b), w.h.p. K is such that we must have $|N(S)| \leq \nu n$.

We expose the edges from S to $N(S)$ as follows[2]: First, for every $v \notin S$, we test the presence of an edge from v to each of the vertices in S, one-at-a-time, and stop testing as

[2]We are grateful to an anonymous referee for suggesting this approach, which is simpler and more elegant than our original proof.

soon as we discover the first edge. This determines the vertices of $N(S)$. Next, for each $u \in N(S)$, we test the presence of an edge from u to each vertex in S for which the test was not carried out during the first step. The total number of edge-tests carried out in the second step is less than $|S| \times |N(S)|$. Note that S violates (7) iff more than $\frac{k}{4}|S|$ edges are exposed during the second step. So the expected number of sets S of size σn that violate (7) and for which $|N(S)| \leq \nu n$ is at most:

$$\begin{aligned} & \binom{n}{\sigma n} \times \binom{\sigma n \times \nu n}{\frac{k}{4}\sigma n} \left(\frac{c}{n}\right)^{\frac{k}{4}\sigma n} \\ < \quad & \left(\frac{e}{\sigma}\right)^{\sigma n} \left(\frac{ec\sigma\nu n^2}{\frac{k}{4}\sigma n^2}\right)^{\frac{k}{4}\sigma n} \\ < \quad & \left(\frac{e}{\sigma}\right)^{\sigma n} (25\nu)^{\frac{k}{4}\sigma n} \qquad \text{since } c < 2k \\ = \quad & \left(\frac{1}{2}\right)^{\sigma n} \qquad \text{by our choice of } \nu. \end{aligned}$$

Summing over all $i = |S|$, we obtain that the expected number of sets S of size greater than $\log n$ that violate (7) is at most $\sum_{i \geq \log n} 2^{-i} = o(1)$. For the case where $|S| < \log n$, we use the well known fact that w.h.p. the maximum degree in $G_{n,c/n}$ is less than $\log n$ (see e.g. Exercise 3.5 in [2]), and so $|N(S)| \leq |S| \log n$. Thus, replacing νn above by $|S| \log n$ and applying $|S| \leq \log n$ and $k \geq 10$ we obtain a smaller bound of

$$\left(\frac{en}{|S|}\right)^{|S|} \left(\frac{25|S|\log n}{n}\right)^{\frac{k}{4}|S|} < (25e\log n)^{\frac{k}{4}|S|} \left(\frac{|S|}{n}\right)^{(\frac{k}{4}-1)|S|} < \left(\frac{\log^2 n |S|}{n}\right)^{(\frac{k}{4}-1)|S|} < \frac{1}{n^2}.$$

Multiplying this by the $\log n$ choices for $|S| \leq \log n$ yields that the expected number of violating sets of size at most $\log n$ is $o(1)$. This establishes the lemma. $\square$

Our final two lemmas follow from the arguments of [5]:

Lemma 2.9. *For any $2k > c > c_{k+1}$, w.h.p. the $(k+1)$-core K of $G_{n,c/n}$ satisfies (P8).*

Proof sketch: The same analysis as in Case 3 of [5] shows that w.h.p. every such S, T satisfies $e(S,T) \leq \frac{3}{4}k|S|$. Indeed, they use a straightforward bound on the tail of the degree sequence to show that w.h.p. $\sum \deg(v)$ over all $v \in G_{n,c/n}$ with $\deg(v) > \frac{3}{2}c$ is less than $\epsilon_0 n$, and trivially, $\sum \deg(v)$ over all $v \in T$ with $\deg(v) \leq \frac{3}{2}c$ is at most $\frac{3}{2}c|T| < \frac{3}{20}c\epsilon_0 n$. So, using $c < 2k$ and $|S| > \frac{9}{10}\epsilon_0 n$, we obtain:

$$e(S,T) \leq \sum_{v \in T} \deg(v) < \epsilon_0 n + \frac{3}{20}c\epsilon_0 n < \frac{1}{5}c\epsilon_0 n < \frac{3}{4}k|S|.$$

Lemma 2.10. *For any $c_{k+1} + 2\sqrt{k \log k} > c > c_{k+1}$, w.h.p. the $(k+1)$-core K of $G_{n,c/n}$ satisfies (P9).*

Proof sketch: Since $|T| \geq \frac{1}{10}\epsilon_0 n$ then the same argument that yielded (14) from [5] (the only difference is a trivial reworking of a few constants) yields that there exists $\epsilon > 0$ such that w.h.p. $e(S,T) \leq k|S| + (1-\epsilon)\sqrt{k \log k}|T|$ for every such S, T. The degree sequence analysis preceding (14) in [5] (after replacing ϵ by $\frac{\epsilon}{2}$) yields that for k sufficiently large, we w.h.p. have $\sum_{v \in T} d(v) > (k + (1 - \frac{\epsilon}{2}\sqrt{k \log k}))|T|$ for every such T.

3 Proof of the main theorem

We consider any $k \geq k_0$ and $c_{k+1} < c < c_{k+1} + 2\sqrt{k \log k}$. We let K be the $(k+1)$-core of $G_{n,c/n}$. By Lemmas 2.1 to 2.10, we can assume that K satisfies properties (P1),...,(P9) from the previous section. Our main lemma is to show that (2) holds and that (3) holds except for a special case:

Lemma 3.1. *Suppose that a graph K has minimum degree at least $k+1$ and satisfies (P1),...,(P9).*

(a) Inequality (2) holds for every disjoint $S, T \subseteq V(K)$ with $S \cup T \neq \emptyset$; and

(b) Inequality (3) holds for every disjoint $S, T \subseteq V(K)$ with

(i) $|S| \geq 2$; or

(ii) $|S| = 1$ and $K - S \cup T$ is disconnected.

This proves our main theorem as follows:

Proof of Theorem 1.1: By Lemmas 2.1,...,2.10, w.h.p. K satisfies (P1),...,(P9).

Part (a): (2) implies (1) for $G = K, R = S, W = T$. So Lemma 3.1(a) implies that (1) holds for $G = K$, for all $R, W \subseteq V(K)$ with $R \cup W \neq \emptyset$. (P1) implies that K is connected, which implies (1) for $R \cup W = \emptyset$ when $k|K|$ is even. So part (a) follows from Theorem 1.2.

Part (b): For any vertex $x \in V(K)$, (3) applied to some S, T with $x \in S$ implies (1) for $G = K - x, R = S - x, W = T$. So Lemma 3.1(b) implies that (1) holds for $G = K - x$, for all $R, W \subseteq V(K)$, except when $R = \emptyset$ and $G - W$ is connected. In that case, (1) becomes:

$$\sum_{v \in W} \deg_G(v) \geq q(\emptyset, W) + k|W|.$$

K has minimum degree at least $k+1$ and so $G = K - x$ has minimum degree at least k. Since $G - W$ is connected, $q(\emptyset, W) \leq 1$. So (1) holds if there is at least one $v \in W$ with $\deg_G(v) \geq k+1$. Thus we assume that $\deg_G(v) = k$ for every $v \in W$. So letting Q be the only component of $G - W$, we have $e(Q, W) = k|W| - 2|E(W)|$. This has the same parity as $k|Q|$ since $|Q| + |W| = |K| - 1$ and $k|K|$ is odd (as we are in part (b)). Thus, $q(\emptyset, W) = 0$ and so (1) holds. So part (b) follows from Theorem 1.2. $\square$

So it only remains to prove our main lemma.

Proof of Lemma 3.1: We note that (8), (9) below imply (2), (3) respectively, since $\deg_k(v) \geq k+1$ for all v.

$$|T| + k|S| \geq \omega(K - S \cup T) + e(S, T). \tag{8}$$

$$|T| + k|S| \geq \omega(K - S \cup T) + e(S, T) + k. \tag{9}$$

We will consider three cases for the sizes of S, T. In all but the last, we actually prove (8), (9). Recall that $s(n) = \log n/(2ec \log\log n)$.

Case 1: $|S| + |T| \leq s(n)$.

The proof of this case is similar to that from [5]. Let $\omega(K - (S \cup T)) = \ell + 1$. By (P3), K is such that the sizes of $C_1, ..., C_\ell$, the ℓ smallest components of $K - (S \cup T)$, must total less than $2s(n)$. So (P2) implies that the subgraph X induced by $S \cup T \cup C_1 \cup ... \cup C_\ell$ has no more edges than vertices. Let X' be the multigraph obtained from X by contracting each C_i into a single vertex which we denote c_i. Since, by (P2), C_i has at most one cycle, and since every vertex of C_i has degree at least $k+1$ in X, it follows that $\deg(c_i) \geq k+1$. Since each c_i is only adjacent to vertices in $S \cup T$ we have $|E(X')| \geq (k+1)\ell + e(S, T)$. Since each C_i is connected, the contractions reduce the number of edges by at least as much as the number of vertices. Since X has no more edges than vertices, $|E(X')| \leq |V(X')| = |T| + |S| + \ell$. Therefore $|T| + |S| + \ell \geq (k+1)\ell + e(S, T)$ and so:

$$|T| + k|S| \geq k\ell + e(S, T) + (k-1)|S| = \omega(K - (S \cup T)) + e(S, T) + (k-1)(|S| + \ell) - 1.$$

This implies (8), and hence (2), for $|S| + \ell \geq 1$ and implies (9), and hence (3) for $|S| + \ell \geq 2$ (and $k \geq 3$).

Note that we only need to establish (3) when $|S| + \ell \geq 2$. So it only remains to prove (2) when $|S| = \ell = 0$. In that case, since $S \cup T \neq \emptyset$, we must have $|T| \geq 1, \omega(K - (S \cup T)) = \ell + 1 = 1$ and $e(S, T) = 0$, and so (8) and hence (2) holds.

For the next case, recall the constant $\epsilon_0 > 0$ from Section 2.

Case 2: $s(n) \leq |S| + |T| \leq \epsilon_0 n$

We use (P4) to bound $\omega(K - S \cup T)$. Let X be the set of vertices in all components of $K - S \cup T$ that have size at most $\frac{1}{2}|V(K)|$. By applying (P1) to each component of X, we have $e(X, S \cup T) \geq \gamma(k+1)|X| > \frac{1}{2}\gamma k|X|$. Therefore, letting $Y = S \cup T$ and using that $|Y| \leq \epsilon_0 n$, we see that (P4) is violated unless $|X| < \frac{1}{200}|S \cup T|$. Since $\omega(K - S \cup T) \leq |X| + 1$, this implies that K is such that:

$$\omega(K - S \cup T) < \frac{1}{200}(|S| + |T|) + 1 < \frac{1}{100}(|S| + |T|), \tag{10}$$

as $|S| + |T| \geq s(n) > 200$.

Case 2a: $|T| \leq 20k|S|$.

Since $s(n) \leq |S| + |T| \leq \epsilon_0 n$, (P5) and (10) imply:

$$\omega(K - S \cup T) + e(S,T) < \frac{1}{100}(|T| + |S|) + \frac{101}{100}|T| + \frac{k}{2}|S| = \frac{102}{100}|T| + (\frac{k}{2} + \frac{1}{100})|S|. \quad (11)$$

Since $|T| \leq 20k|S|$ and $|S| \geq \frac{|S|}{2} + \frac{|T|}{40k} > \frac{s(n)}{40k} > 100$, we have for k sufficiently large:

$$\frac{102}{100}|T| + (\frac{k}{2} + \frac{1}{100})|S| < |T| + \frac{91k}{100}|S| < |T| + k|S| - k,$$

and so (11) implies (8), (9) and hence (2), (3).

Case 2b: $|T| > 20k|S|$.

Note that, since $|S| + |T| \leq \epsilon_0 n$, we have $|S| \leq \frac{\epsilon_0}{20k}n$.

This case contains most of the new ideas for this paper. To prove (2) and (3), it would suffice to show $\omega(K - S \cup T) + e(S,T) \leq |T| + k|S| - k$. In Case 2a, we saw that (P5) and (10) yield (11), which was less than $|T| + k|S| - k$ since, in that case, T was a lot smaller than S. Throughout Case 2, (11) clearly yields $\omega(K - S \cup T) + e(S,T) \leq 2|T| + k|S| - k$, which would suffice for (2) and (3) if K were the $(k+2)$-core. So in [5], the analysis for Case 2a sufficed to cover all of Case 2.

It is natural to try and tighten the proof of Lemma 2.5 (namely that (P5) holds w.h.p.) to obtain: $e(S,T) < |T| + \frac{k}{2}|S|$. Unfortunately, this approach fails - the proof of Lemma 2.5 uses a first moment calculation, and the $\binom{n}{|T|}$ term in that calculation is far too big. But instead of bounding $e(S,T)$, we can bound $e(S, N(S))$. The advantage of replacing T by $N(S)$ is that the choice of the vertices in S determines $N(S)$ and so the $\binom{n}{|T|}$ term is replaced by 1. This is what led us to consider (P6).

Since $20k|S| < |T| < \epsilon_0 n$, we have $|S| < \frac{\epsilon_0}{20k}n$ and so (P7) yields:

$$e(S,T) \leq e(S, N(S)) - |N(S) \backslash T| \leq |N(S)| + \frac{k}{4}|S| - |N(S) \backslash T| = |N(S) \cap T| + \frac{k}{4}|S|. \quad (12)$$

Next, we will bound $\omega(K - S \cup T)$. First, we note that if $S = \emptyset$ then $e(S,T) = |S| = 0$ and so (10) easily implies (8) and hence (2). (We can also show that (3) holds, but it is not required to hold when $S = \emptyset$.) Thus, we will assume $|S| \geq 1$.

Recall that we defined X to be the set of vertices in all components of $K - S \cup T$ of size at most $\frac{1}{2}|V(K)|$ and so $|X| \geq \omega(K - S \cup T) - 1$. Recall also that $|S| \leq \frac{\epsilon_0}{20k}n$. If $|X| \geq \max(\frac{1}{200}|T \backslash N(S)|, |S|)$ then (P4) with $Y = T \backslash N(S)$ and (P6) imply:

$$e(X, S \cup T) = e(X, T \backslash N(S)) + e(X, S \cup (T \cap N(S))) \leq e(X, T \backslash N(S)) + e(X, (S \cup N(S)) \backslash X) \leq \gamma k|X|.$$

Each component of the graph induced by X has size at most $\frac{1}{2}|V(K)|$. So we can apply (P1) to each such component and sum to obtain a contradiction unless $X = \emptyset$, since $e(X, K - X) =$

$e(X, S \cup T)$. Since we assume $|S| \geq 1$, this implies $|X| < \max(\frac{1}{200}|T \backslash N(S)|, |S|)$, which again since $|S| \geq 1$, implies $|X| \leq |S| + \frac{1}{200}|T \backslash N(S)| - 1$. Therefore

$$\omega(K - S \cup T) \leq |X| + 1 \leq |S| + \frac{1}{200}|T \backslash N(S)|.$$

This, along with (12) implies

$$\begin{aligned} \omega(K - S \cup T) + e(S, T) &\leq |S| + \frac{1}{200}|T \backslash N(S)| + |T \cap N(S)| + \frac{k}{4}|S| \\ &= k|S| + |T| - \frac{199}{200}|T \backslash N(S)| - \left(\frac{3k}{4} - 1\right)|S|. \end{aligned}$$

This yields (8). It also implies (9) if $|S| \geq 2$ and so $\left(\frac{3k}{4} - 1\right)|S| > k$.

The only remaining task is to prove (9) for the case $|S| = 1$ and $K - S \cup T$ is disconnected, which implies $|X| \geq 1$. Above, we proved that $|X| < \max(\frac{1}{200}|T \backslash N(S)|, |S|)$ and so we must have $|T \backslash N(S)| > 200$. Since $|S| = 1$, we can trivially strengthen (12) to $e(S, T) = |N(S) \cap T|$. That improves the above bound to

$$\omega(K - S \cup T) + e(S, T) \leq k|S| + |T| - \frac{199}{200}|T \backslash N(S)| - k + 1,$$

which implies (9) since $|T \backslash N(S)| \geq 200$. So we have proven (8), (9) and hence (2), (3).

Case 3: $|S| + |T| \geq \epsilon_0 n$

The analysis in Cases 3 and 4 of [5] also applies to the setting of this paper (after a straightforward adjustment of some of the constants), to cover this remaining case.

If $|T| < \frac{1}{10}\epsilon_0 n$ then $|S| > \frac{9}{10}\epsilon_0 n$. So by (P8), $e(S, T) < \frac{3}{4}k|S|$. Since $\omega(G - S \cup T) < n < \frac{1}{4}k|S| - k$ for $k > \frac{8}{\epsilon_0}$, this yields (8), (9) and hence (2), (3).

If $|T| \geq \frac{1}{10}\epsilon_0 n$ then we apply (P9), noting that $\omega(G - S \cup T) + 1 < n < \frac{\epsilon}{2}\sqrt{k \log k}|T| - k$ for $k > 4/(\epsilon\epsilon_0)^2$, to obtain (8), (9) and hence (2), (3). $\square$

Acknowledgements

We thank Pawel Pralat, Jacques Verstraete and Nick Wormald for their helpful discussions, and an anonymous referee for suggesting several improvements to the presentation and a better proof of (7).

References

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

[2] B. Bollobás, *Random Graphs.* 2nd edition. Cambridge University Press, 2001.

[3] B. Bollobás, J.H. Kim and J. Verstraete, *Regular subgraphs of random graphs.* Rand. Struc. & Alg. **29** (2006), 1 - 13.

[4] B. Pittel, J. Spencer and N. Wormald, *Sudden emergence of a giant k-core in a random graph.* J.Comb.Th.(B) **67** (1996), 111 - 151.

[5] P. Pralat, J. Verstraete and N. Wormald, *On the threshold for k-regular subgraphs of random graphs.* Submitted.

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

[7] W. Tutte, *The factors of graphs.* Canad. J. Math **4** (1952), 314 - 328.

[8] D. West, *Introduction to Graph Theory.* 2nd edition. Prentice Hall, 2001.