(k + 1)-cores have k-factors

Siu On Chan∗  Michael Molloy∗

June 26, 2010

Abstract

We prove that almost surely the first non-empty (k+1)-core to arise during the random graph process will have a k-factor or will be k-factor-critical. Thus the threshold for the appearance of a k-regular subgraph is at most the threshold for the (k+1)-core. This improves a result of Pralat, Verstraete and Wormald [5] and proves a conjecture of Bollobas, 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 $G_{n,p} = c/n$, the random graph with n vertices where each of the $\binom{n}{2}$ possible edges appears independently with probability p. So for $c < c_k$, a.s.$^1$ $G_{n,p} = c/n$ has no non-empty k-core and hence a.s. has no k-regular subgraph. In [3], Bollobás, Kim and Verstraete studied the threshold for the appearance of a 3-regular subgraph, and determined that it is strictly larger than $c_3$. They also conjectured that the threshold for a k-regular subgraph is strictly larger than $c_k$ for all $k \geq 4$. Pretti and Weigt[6] used some statistical physics techniques to predict the opposite: for every $k \geq 4$, the threshold for the appearance of a k-regular subgraph is $c_k$. In other words, for every $c > c_k$, a.s. the k-core contains a k-regular subgraph. Those conflicting conjectures remain unresolved.

Bollobás, Kim and Verstraete also conjectured that if $c > c_{k+1}$ then a.s. the (k+1)-core of $G_{n,p} = c/n$ has a k-regular subgraph (see Conjecture 1.3 from [3]). We prove that conjecture here for k sufficiently large. They proved that $G_{n,p} = c/n$ a.s. contains a k-regular subgraph if $c > \rho_k n$ for a specific function $\rho_k = 4k + o(k)$; note that $\rho_k \approx 4c_k$.

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

$^1$A property holds almost surely (a.s.) if it holds with probability tending to 1 as $\lim_{n \to \infty}$. 
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 \( c_{k+2} < c < c_{k+2} + 10k \log k \) and let \( C \) denote the \((k+2)\)-core of \( G_{n,p=c/n} \). Pralat, Verstraete and Wormald[5] proved that if \( k \) is sufficiently large then a.s. (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 \( c_{k+1} < c < c_{k+1} + 10k \log k \) a.s. the \((k+1)\)-core, \( K \), of \( G_{n,p=c/n} \) 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 \( c > c_k \) a.s. the \( k \)-core of \( G_{n,p=c/n} \) neither contains a \( k \)-factor nor is \( k \)-factor-critical, because it a.s. contains many vertices of degree greater than \( k \) whose neighbours all have degree exactly \( k \).

By monotonicity, Theorem 1.1 implies that for any \( c > c_{k+1} \), a.s. the \((k+1)\)-core of \( G_{n,p=c/n} \) 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]. 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. [3] remarked that perhaps a.s. 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 \( k \).

Our proof makes use of Tutte’s \( f \)-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 \( X, Y \subset V(G) \), we use \( \lambda(X, Y) \) 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 \( \lambda(Q, Y) \) have different parities.

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

\[
\sum_{v \in W} \text{deg}_G(v) + k|R| \geq q(R, W) + k|W| + \sum_{v \in W} \text{deg}_{G-R}(v).
\]

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

\[
\sum_{v \in W} \text{deg}_G(v) + k|R| \geq q(R, W) + k|W| + \lambda(R, W).
\]

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

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

By Theorem 1.2, with $R := S, W := T$ this will suffice to prove part (a), since $\omega(K - R \cup W) \geq q(R, W)$.

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

$$\sum_{v \in T} \deg_K(v) + k|S| \geq \omega(K - S \cup T) + k|T| + \lambda(S, T) + k.$$  \hfill (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]). Thus, if (3) were to hold for all $S, T$ with $|S| \geq 1$ and $|S \cup T| \geq 2$ then this would establish part (b). This was indeed the case in [5]. Unfortunately there are some cases in our setting where (3) does not hold, so we need to instead focus directly on (1).

To see why our setting is a bit more delicate, 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 $G = K - x$.

Our proof follows the same outline as that of [5]. Their proof covered four separate cases for the sizes of $S, T$. In Case 1, 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. Their arguments for Cases 3 and 4 apply to the setting of this paper, so we didn’t need any new ideas there; we combine them into our Case 3. The reader who is already familiar with [5] may want to skip directly to Case 2 (in particular, Case 2b).

We close this introduction by noting that our main theorem extends to $G_{n,M}$, a model that permits a somewhat stronger statement. The random graph process begins with $n$ vertices and no edges, and then repeatedly adds an edge chosen uniformly at random from amongst those edges not yet present. $G_{n,M}$ is the graph obtained after $M$ steps.

**Theorem 1.3** There is an absolute constant $k_0$ such that for all $k \geq k_0$, a.s. $K$, the first non-empty $(k + 1)$-core to arise during the random graph process, 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.
1.1 Preliminaries

We will make use of the following lemmas from [5] concerning the structure of $K$. (Actually, their lemmas were stated a bit differently in that they were in terms of the $k$-core. But it is straightforward to adapt their proofs to obtain the statements below.)

**Lemma 1.4** (Lemma 2 of [5].) There is a constant $\gamma > 0$ (independent of $k$) such that a.s. for every set $X \subset V(K)$ of at most $\frac{1}{2}|V(K)|$ vertices, we have:

$$\lambda(X, K - X) \geq \gamma(k + 1)|X|.$$  

For the remainder of the paper, we use $\gamma$ to denote the constant from Lemma 1.4. We define:

$$s(n) = \log n/(2e c \log \log n).$$

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

**Lemma 1.5** For any constant $c > 0$, a.s. every subset $Y$ of the vertices of $G_{n, p = c/n}$ with $|Y| \leq 4s(n)$ has at most $|Y|$ edges.

Lemma 4 of [5] says:

**Lemma 1.6** If $k$ is sufficiently large then: a.s. for every subset $Y \subseteq V(K)$ with $|Y| \leq s(n)$, $K - Y$ contains a component with more than $|V(K)| - 2s(n)$ vertices.

**Proof:** Let $X$ be the union of the vertex sets of some components of $K - Y$, such that $|X| > s(n)$. We’ll show that if the a.s. properties from Lemmas 1.4 and 1.5 hold then $|X| > \frac{1}{2}|V(K)|$; this implies the lemma.

Consider any $Z \subset X$ where $|Z| = |Y|$. Thus $|Y \cup Z| \leq 2s(n)$ and so by Lemma 1.5 we can assume $\lambda(Y, Z) \leq |Y \cup Z| = 2|Z|$. Averaging over all such $Z \subseteq X$ yields $\lambda(Y, X) \leq 2|X| < \gamma k|X|$, for $k$ sufficiently large (since $\gamma$ does not depend on $k$). Since $\lambda(X, K - X) = \lambda(Y, X)$, the a.s. property of Lemma 1.4 implies $|X| > \frac{1}{2}|V(K)|$ as required.

We often use the following well-known bound which follows easily from Stirling’s Inequality:

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

And finally, recall that [4] established $c_k = k + \sqrt{k \log k} + o(\sqrt{k})$ and that the hypothesis of Theorem 1.1 requires $c < c_k + 10\sqrt{k \log k}$. Thus, for $k$ sufficiently large, we have:

$$c < 2k.$$
2 Proof of Theorem 1.1

We will consider three cases for the sizes of $S, T$ from (2) and (3). Recall that $s(n) = \log n/(2\varepsilon \log \log n)$.

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 Lemma 1.6, a.s. $K$ is such that the sizes of $C_1, \ldots, C_\ell$, the $\ell$ smallest components of $K - (S \cup T)$, must total less than $2s(n)$. So Lemma 1.5 implies that a.s. the subgraph $X$ induced by $S \cup T \cup C_1 \cup \ldots \cup C_\ell$ has no more edges than vertices. Let $X'$ be the graph obtained by contracting each $C_i$ into the single vertex $c_i$. Since $C_i$ has at most one cycle (by Lemma 1.5) and 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 $|E(X')| \geq (k + 1)\ell + \lambda(S, T)$. Since $X$ has no more edges than vertices and since each $C_i$ is connected, $|E(X')| \leq |V(X')| = |S| + |T| + \ell$. Therefore:

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

Since every vertex in $T$ has degree at least $k + 1$, this implies (2) for $|S| + \ell \geq 1$ and (3) for $|S| + \ell \geq 2$ (and $k \geq 3$).

If $|S| = \ell = 0$ and $S \cup T \neq \emptyset$ then we must have $|T| \geq 1, \omega(K - (S \cup T)) = 1$ and $\lambda(S, T) = 0$, and so (2) holds.

We aren’t required to prove (3) for $|S| = 0$. So we have only failed to prove (3) for the case $|S| = 1, \ell = 0$; in fact, (3) does not a.s. hold in this case. Proving (3) is required only to prove Theorem 1.1(b); i.e. to establish that if $k|K|$ is odd then $K - x$ has a $k$-factor for every $x \in V(K)$. We will establish that by showing directly that (1) holds for $G = K - x$. The fact that (3) holds for $K$ when $|S| \geq 2$ or $|S| = 1, \ell \geq 1$ implies that (1) holds for $G = K - x$ whenever $|R| \geq 1$ and whenever $|R| = 0$ and $(K - x) - W$ has more than one component (recall the discussion following the statement of (3)). So we can assume $R = \emptyset$ and $(K - x) - W$ has at most one component. Then (1) becomes:

$$\sum_{v \in W} \deg_{K-x}(v) \geq q(\emptyset, W) + k|W|.$$  

$K$ has minimum degree at least $k + 1$ and so $K - x$ has minimum degree at least $k$. Since $(K - x) - W$ has at most one component, $q(\emptyset, W) \leq 1$. So (1) holds if there at least one $v \in W$ with $\deg_{K-x}(v) \geq k + 1$. Let $Q$ be the only component of $K - x - W$. If every $v \in W$ has $\deg_{K-x}(v) = k$ then $\lambda(Q, W) = k|W| - 2E(W)$ which has the same parity as $k|Q|$ since $|Q| + |W| = |K| - 1$ and $k|K|$ is odd (as we are in Theorem 1.1(b)). Thus, $q(\emptyset, W) = 0$ and so (1) holds.

This proves that a.s. for every $S, T$ satisfying Case 1, (6) holds for $S, T$ and (1) holds for $R := S - x, W := T$ with $G := K - x$. 

5
To specify Case 2, we fix an absolute constant $\epsilon_0$, independent of $k$, chosen so that $\epsilon_0 < \frac{\gamma^2}{10^7}$ (recall $\gamma$ from Lemma 1.4).

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

We use the following two technical bounds, which are very much like bounds found in [5]. We defer the proofs until Section 3.

A.s. for every disjoint pair of sets $X, Y$ with $|X| \geq \frac{1}{200} |Y|$ and $|Y| \leq \epsilon_0 n$ we have:

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

A.s. for every disjoint pair of sets $S, T$ with $s(n) \leq |S| + |T| \leq \epsilon_0 n$ we have:

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

We use (4) 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 Lemma 1.4 to each component of $X$, we have $\lambda(X, S \cup T) \geq \gamma (k+1)|X|$. Therefore, letting $Y = S \cup T$ and recalling that, in Case 2, $|Y| \leq \epsilon_0 n$, (4) is violated unless $|X| < \frac{1}{200} |S \cup T|$. Since $\omega(K - S \cup T) \leq |X| + 1$, this implies that a.s. $K$ is such that for every $S, T$ in Case 2 we have:

$$\omega(K - S \cup T) < \frac{1}{200} (|S| + |T|) + 1 < \frac{1}{100} (|S| + |T|). \quad (6)$$

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

(5) and (6) imply that a.s. every pair $S, T$ with $s(n) \leq |S| + |T| \leq \epsilon_0 n$ and $|T| \leq 20k|S|$ satisfies:

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

where the last inequality uses $|T| \leq 20k|S|$.

This implies that a.s. (2) and (3) hold for every $S, T$ satisfying Case 2a.

**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 $\omega(K - S \cup T) + \lambda(S, T) \leq |T| + k|S| - k$. Above, we saw that (5) and (6) yield $\omega(K - S \cup T) + \lambda(S, T) \leq \frac{102}{100} |T| + \left(\frac{k}{2} + \frac{101}{100}\right)|S|$, which is less than $|T| + k|S| - k$ if $T$ is a lot smaller than $S$, e.g. in Case 2a. Throughout Case 2, that bound clearly yields $\omega(K - S \cup T) + \lambda(S, T) \leq 2|T| + k|S| - k$, which would suffice for (2) and (3) if $K$ were the $(k+2)$-core. So the analysis above sufficed to cover all of Case 2 in [5].
It is natural to try and tighten the proof of (5) to obtain: $$\lambda(S,T) < |T| + \frac{k}{4} |S|$$. Unfortunately, this approach fails - the proof of (5) uses a first moment calculation, and the $n\choose|T|$ term in that calculation is far too big. But instead of bounding $\lambda(S,T)$, we can bound $\lambda(S,N(S))$. The advantage of replacing $T$ by $N(S)$ is that the choice of the vertices in $S$ determines $N(S)$ and so the $n\choose|T|$ term is replaced by 1. We will obtain:

A.s. for every set $S \subset V(K)$ with $|S| \leq \frac{\alpha}{20k} n$ we have:

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

This yields that a.s. for every disjoint pair of sets $S, T$ as in Case 2b, we have:

$$\lambda(S,T) \leq \lambda(S,N(S)) - |N(S)\setminus T| \leq |N(S) \cap T| + \frac{k}{4} |S|. \quad (8)$$

We will also show a bound similar to (4):

A.s. for every pair of disjoint sets $S, X \subset V(K)$ with $|S| \leq \frac{\alpha}{20k} n$ and $|X| \geq |S|$ we have:

$$\lambda(X, (S \cup N(S)) \setminus X) \leq \frac{1}{2} \gamma k |X|. \quad (9)$$

The proofs of (7) and (9) appear in Section 3.

Next, we will bound $\omega(K - S \cup T)$. Consider any pair of sets $S, T$ with sizes as in Case 2b. First, we note that if $S = \emptyset$ then $|T| \geq s(n)$ and (6) implies that:

$$\omega(K - S \cup T) + k|T| + \lambda(S,T) \leq \frac{1}{100} (|S| + |T|) + kT + \lambda(S, T) = \frac{1}{100} |T| + k|T| < \sum_{v \in T} \deg_K(v),$$

and so (2) holds. (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 in Case 2b we have $|S| \leq \frac{\alpha}{20k} n$. If $|X| \geq \max(\frac{1}{200} |T\setminus N(S)|, |S|)$ then (4) with $Y = T\setminus N(S)$ and (9) imply:

$$\lambda(X, S \cup T) = \lambda(X, T\setminus N(S)) + \lambda(X, S \cup (T \cap N(S))) \leq \lambda(X, T\setminus N(S)) + \lambda(X, (S \cup N(S)) \setminus X) \leq \gamma k |X|,$$

which contradicts Lemma 1.4 unless $X = \emptyset$, since $\lambda(X, K - X) = \lambda(X, S \cup T)$. Since we can assume $|S| \geq 1$, this implies $|X| \leq \max(\frac{1}{200} |T\setminus N(S)|, |S|)$, which again since $|S| \geq 1$, implies $|X| \leq |S| + \frac{1}{200} |T\setminus N(S)| - 1$. Therefore

$$\omega(K - S \cup T) \leq |X| + 1 \leq |S| + \frac{1}{200} |T\setminus N(S)|.$$
This, along with (8) implies
\[ \omega(K - S \cup T) + \lambda(S, T) \leq |S| + \frac{1}{200}|T \setminus N(S)| + |T \cap N(S)| + \frac{k}{4}|S| \]
\[ = k|S| + |T| - \frac{199}{200}|T \setminus N(S)| - \left( \frac{3k}{4} - 1 \right)|S|. \]

This yields (2). It also implies (3) if |S| ≥ 2 and so (\(\frac{3k}{4} - 1\)) |S| > k. When |S| = 1, we can trivially strengthen (8) to \(\lambda(S, T) = |N(S) \cap T|\). That improves the above bound to
\[ \omega(K - S \cup T) + \lambda(S, T) \leq k|S| + |T| - \frac{199}{200}|T \setminus N(S)| - k + 1, \]
which implies (3) if at least one \(v \in T\) has degree \(k(v) \geq k + 2\) or if \(|T \setminus N(S)| \geq 1\).

So the only remaining case is where |S| = 1, \(T \subseteq N(S)\) and every vertex in \(T\) has degree \(k + 1\). Above, we proved that \(|X| < \max(\frac{1}{200}|T \setminus N(S)|, |S|)\) and so, in this case, \(|X| = 0\). We work directly with (1), proving that it holds for \(R := \emptyset, W := T, G = K - x\) with the same argument that was used in Case 1.

This proves that a.s., for every \(S, T\) satisfying Case 2b, (2) holds for \(S, T\) and (1) holds for \(R := S - x, W := T\) with \(G = K - x\).

**Case 3:** \(|S| + |T| \geq \epsilon_0 n\)

This is covered by Cases 3 and 4 from [5]. The proofs from that paper also apply to the setting of this paper (after a straightforward adjustment of some of the constants).

In particular, if \(|T| < \frac{1}{10} \epsilon_0 n\) then \(|S| > \frac{9}{10} \epsilon_0 n\). The same analysis as in Case 3 of [5] shows that a.s. every such \(S, T\) satisfies \(\lambda(S, T) \leq \frac{3}{4} k|S|\). Indeed, they use a straightforward bound on the tail of the degree sequence to show that a.s. \(G_{n,p=\epsilon_0/n}\) is such that \(\sum \deg(v)\) over all \(v \in T\) with \(\deg(v) > \frac{3}{2} c\) must be 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 \(\frac{3}{2} c|T| < \frac{3}{20} \epsilon_0 n\). So, using \(c < 2k\) and \(|S| > \frac{9}{10} \epsilon_0 n\), we obtain:
\[ \lambda(S, T) \leq \sum_{v \in T} \deg(v) < \epsilon_0 n + \frac{3}{20} \epsilon_0 n < \frac{1}{5} \epsilon_0 n < \frac{3}{4} k|S|. \]

Since \(\sum_{v \in T} d(v) \geq (k + 1)|T|\) and \(\omega(G - (S \cup T)) < n < \frac{1}{4} k|S| - 1\) for \(k > \frac{8}{\epsilon_0}\), (2) and (3) both hold.

If \(|T| \geq \frac{1}{10} \epsilon_0 n\) then the same argument that yielded (18) from [5] (the only difference is a trivial reworking of a few constants) yields that there exists \(\epsilon > 0\) such that a.s. \(\lambda(S, T) \leq k|S| + (1 - \epsilon) \sqrt{k \log k |T|}\) for every such \(S, T\). The degree sequence analysis preceding (18) in [5] (after replacing \(\epsilon\) by \(\frac{\epsilon}{2}\)) yields that for \(k\) sufficiently large, we a.s. have \(\sum_{v \in T} d(v) > (k + (1 - \frac{\epsilon}{2} \sqrt{k \log k}))|T|\) for every such \(T\). Since \(\omega(G - (S \cup T)) + 1 < n < \frac{3}{2} \sqrt{k \log k |T|} - k\) for \(k > 4/(\epsilon \epsilon_0)^2\), this yields (2) and (3).
Remark: It is in this final step that we require \( c \leq c_{k+1} + 10\sqrt{k \log k} < k + 12\sqrt{k \log k} \). Replacing 10 by any other constant would suffice.

Therefore, a.s. (2) and (3) hold for every \( S, T \) in Case 3.

**Proof of Theorem 1.1** We have proved that (2) holds for every \( S, T \), which implies that (1) holds for every \( R, W \) when \( G := K \). This establishes Theorem 1.1(a). We have proved that (2) holds for all but a few cases of \( S, T \); as described in the introduction, this implies that (1) holds when \( R := S - x, W := T \) and \( G := K - x \). For those few remaining cases, we showed directly that (1) holds. Thus (1) holds for all \( R, W \) when \( G := K - x \); this establishes Theorem 1.1(b).

\[ \square \]

We close this section by presenting the adaptation of our arguments to the \( G_{n,M} \) model.

**Proof of Theorem 1.3** It suffices to prove that all of the a.s. statements from our proof also hold when \( K \) is the first non-empty \((k+1)\)-core to arise during the random graph process. Specifically, these statements are: Lemmas 1.4, 1.5, 1.6, (4), (5), (7) and (9) and the bound on \( \lambda(S, T) \) corresponding to (18) from [5], as well as the degree sequence analysis from Case 3. All but Lemma 1.4 were proven to hold for the entire graph \( G_{n,p} = c/n \) when \( c < 2k \), rather than just for the \( k \)-core. Each of these properties are monotone (Lemma 1.5 is preserved under the addition of edges, the others are preserved under the deletion of edges), and so Theorem 2.2 of [2] implies that they all hold a.s. for \( G_{n,M} = \frac{1}{2}cn \) for any \( c < 2k \).

This implies that they will a.s. hold for the first \((k+1)\)-core to arise. Lemma 1.4 is Lemma 2 from [5] which, in turn, follows from Lemma 5.3 of [1]. That last lemma was proven for random graphs on a fixed degree sequence, whose degrees all lie between 3 and \( n^{0.02} \). It is well known that the first \((k+1)\)-core to arise is uniformly random on its degree sequence (see eg. [4]), and those degrees lie between \( k + 1 > 3 \) and the maximum degree of \( G_{n,M} \) which is a.s. less than \( \log n << n^{0.02} \). It follows that Lemma 1.4 also holds when \( K \) is the first non-empty \((k+1)\)-core to arise during the random graph process. The remainder of the proof is identical to that of Theorem 1.1. \[ \square \]

### 3 The remaining details

Here we provide the proofs of some of the technical statements from Case 2. Rather than working with the \((k+1)\)-core \( K \) directly, we will actually prove that the statements hold over the entire graph \( G_{n,p} = c/n \).

We begin with equations (4) and (5) from Case 2a.

A.s. for every disjoint pair of sets \( X, Y \) with \( |X| \geq \frac{1}{200} |Y| \) and \( |Y| \leq \epsilon_0 n \) we have:
\[ \lambda(X, Y) \leq \frac{1}{2} \gamma k |X|. \] \hspace{1cm} (4)

**Proof of (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 \( G_{n,p=c/n} \) that violate (4) is at most:

\[
\left( \frac{n}{yn} \right) \left( \frac{n}{xn} \right) \left( \frac{(yn)(xn)}{\frac{1}{2} \gamma k xn} \right) \left( \frac{c}{n} \right)^{\frac{1}{2} \gamma k xn} = \frac{1}{2} \gamma k xn
\]

\[
\left( \frac{e}{y} \right) \left( \frac{x}{e} \right)^{xn} \left( \frac{exyn^2c}{\frac{1}{2} \gamma k xn^2} \right) \left( \frac{c}{n} \right)^{\frac{1}{2} \gamma k xn}
\]

\[
\left( \frac{e}{y/200} \right)^{201xn} \left( \frac{4ey}{\gamma} \right) \left( \frac{e}{y/200} \right)^{\frac{1}{2} \gamma k xn} \left( \frac{e}{y/200} \right) > 1 \text{ and } c < 2k
\]

\[
\left( \frac{3200e^3y}{\gamma^2} \right)^{\frac{1}{2} \gamma k xn}
\]

\[
\left( \frac{1}{2} \right)^{xn} \text{ since } y \leq \epsilon_0 < \frac{1}{2} \left( \frac{\gamma^2}{3200c^3} \right) \text{ and } \frac{1}{4} \gamma k > 1.
\]

For each fixed \( x \), there are at most \( 200xn \) choices for \( y \), since \( s(n) < |Y| \leq 200|X| \). Therefore, summing over all \( x, y \) we find that the expected number of pairs \( X, Y \) violating (4) with \( |X| \geq \log n \) is less than:

\[
\sum_{|X| \geq \log n} 200|X| \left( \frac{1}{2} \right)^{|X|} = o(1).
\]

For \( |X| < \log n \) we have \( |Y| < 200 \log n \); i.e. \( y < \frac{200 \log n}{n} \). Thus \( \left( \frac{3200e^3y}{\gamma^2} \right)^{\frac{1}{2} \gamma k xn} \) \( < 1 \) \( \frac{c}{n} \) \( (101 \frac{1}{100} |T| + \frac{k}{2} |S|) \). 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 (6) is \( o(1) \).

\[ \square \]

A.s. for every disjoint pair of sets \( S, T \) with \( s(n) \leq |S| + |T| \leq \epsilon_0 n \) we have:

\[ \lambda(S, T) \leq \frac{101}{100} |T| + \frac{k}{2} |S|. \] \hspace{1cm} (5)

**Proof of (5):** Let \( \sigma n = |S| \) and \( \tau n = |T| \). For any choice of \( \sigma, \tau \), the expected number of such sets \( S, T \) in \( G_{n,p=c/n} \) violating (5) is at most:

\[
\left( \frac{n}{\sigma n} \right) \left( \frac{n}{\tau n} \right) \left( \frac{(\sigma n)(\tau n)}{\frac{101}{100} \tau n + \frac{k}{2} \sigma n} \right) \left( \frac{c}{n} \right)^{\frac{101}{100} \tau n + \frac{k}{2} \sigma n} < \left( \frac{e}{\sigma} \right)^{\sigma n} \left( \frac{e}{\tau} \right)^{\tau n} \left( \frac{e \sigma \tau n^2 c}{\frac{101}{100} \tau n + \frac{k}{2} \sigma n} \right)^{\frac{101}{100} \tau n + \frac{k}{2} \sigma n}
\]

\[ 10 \]
\[
\left( \frac{e}{\sigma} \right)^{\sigma n} \left( \frac{e}{\tau} \right)^{\tau n} \left( \frac{e^{\sigma \tau c}}{101 \frac{\tau}{100} + \frac{k}{2} \sigma} \right)^{101 \tau + \frac{k}{2} \sigma}.
\]

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

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

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

\[
\frac{e^{\sigma \tau c}}{101 \frac{\tau}{100} + \frac{k}{2} \sigma} < \left( \frac{\tau}{2e} \right)^{100} < e^{-1} < \left( \frac{\sigma}{2e} \right)^{\frac{2}{3}},
\]

while if \( \sigma \leq e^{-k/3} \) then for \( k \) sufficiently large we have:

\[
\frac{e^{\sigma \tau c}}{101 \frac{\tau}{100} + \frac{k}{2} \sigma} < e^{\sigma^{1/2} \tau^{1/2}} < e(2k)e^{-k/6} \sigma^{1/2} < \sigma^{1/2} < \left( \frac{\sigma}{2e} \right)^{\frac{2}{3}}.
\]

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

\[
\left( \frac{e}{\sigma} \right)^{\sigma n} \left( \frac{e}{\tau} \right)^{\tau n} \left( \frac{\tau}{2e} \right)^{100} \left( \frac{\sigma}{2e} \right)^{\frac{2}{3}} = \left( \frac{1}{2} \right)^{(\sigma + \tau)n}.
\]

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:

\[
\sum_{y=s(n)} y \left( \frac{1}{2} \right)^y = o(1).
\]

Now we turn to the results required for Case 2b. We begin with a technical lemma. Note that in Case 2b, we have \(|S| + 20k|S| \leq |S| + |T| \leq \epsilon_0 n\) and so \(|S| < \frac{\epsilon_0}{20k} n\).

**Lemma 3.1** A.s. every set \( S \) in \( G_{n,p=\epsilon/n} \) of size at most \( \frac{6n}{20k} \) satisfies:

1. \( (\frac{25|S \cup N(S)|}{\gamma n})^{\frac{1}{2}} \gamma k \frac{n^2}{|S|^2} < \frac{1}{2e^2} \). 
2. \( (\frac{25|N(S)|}{n})^{\frac{k}{4}} \frac{n}{|S|} < \frac{1}{2e} \).

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

\[
\left( \frac{25|N(S)|}{n} \right)^{\frac{k}{4}} \frac{n}{|S|} < \left( \frac{25|S \cup N(S)|}{\gamma n} \right)^{\frac{k}{4}} \frac{n}{|S|} < \left( \frac{25|S \cup N(S)|}{\gamma n} \right)^{\frac{1}{2}} \gamma k \frac{n^2}{|S|^2} \frac{|S|}{n} < \frac{1}{2e^2} \times \frac{\epsilon_0}{20k} < \frac{1}{2e}.
\]
i.e. (a) implies (b). So we will prove (a).

Let \( S^* \) be the \(|S|\) vertices of largest degree in \( G_{n,p} \) 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 \( G_{n,p=c/n} \) is \( \left( e^i/n e^{-c} + o(1) \right) n \). Standard methods (eg Lemma 3.10 of [2]) show that this number is concentrated enough that: A.s. for all \( i \) such that \( \frac{c^i}{n} \geq \sqrt{n} \) we have (a) at most \( \frac{c^i}{n} \) vertices have degree \( i \) and (b) at most \( \sum_{j \geq i} \frac{c^j}{n} \) vertices have degree at least \( i \). Also, it is well-known that the maximum degree in \( G_{n,p=c/n} \) is a.s. less than \( \log n \) (see eg. Exercise 3.5 of [2]). We will assume that these almost sure 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 clearly 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 \( \frac{c^i}{n} \geq |S| \), and \( i^* \geq I \) to be the largest integer for which \( \frac{c^{i^*}}{n} \geq \sqrt{n} \). It is easily verified that \( \sum_{i \geq 1} \frac{c^i}{n} < 2\frac{e^{i^*+1}}{(i^*+1)!} n < 2\sqrt{n} \), 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:

\[
D < \sum_{i=0}^{i^*} \frac{c^i}{i!} n + 2\sqrt{n} \log n.
\]

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

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

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

\[
\left( \frac{25 \times 48 c^2}{\gamma} \right)^2 \left( \frac{25 \times 12 c^{4c}}{(4c)!} \right)^{\frac{1}{2} \gamma k - 2} < \frac{1}{2e^2}.
\]
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 $|S| = \frac{e}{10} n$. Applying the analysis from Case 2, at that value of $|S|$ we have $|S| + D \leq 3I^c n = 3I|S| = 12c|S|$. Therefore, using the facts that $c < 2k$ and $|S| < \frac{e}{20k} n < \frac{e}{10c} n < \frac{e}{25 \times 25} n$ we have:

$$
\left( \frac{25|S \cup N(S)|}{\gamma n} \right)^{\frac{1}{2}} \frac{n_2}{|S|^2} < \left( \frac{25 \times 12c}{\gamma} \right)^2 \left( \frac{25 \times 12c|S|}{\gamma n} \right)^{\frac{1}{2}} \frac{n_2}{|S|^2} < \left( \frac{25 \times 24k}{\gamma} \right)^2 \left( \frac{1}{2} \right) \frac{1}{2} \frac{1}{2} < \frac{1}{2c^2},
$$

for $k$ sufficiently large.

Next we prove our bound on $\lambda(X, (S \cup N(S)) \backslash X)$:

A.s. for every pair of disjoint sets $S, X \subset V(K)$ with $|S| \leq \frac{e}{20k} n$ and $|X| \geq |S|$ we have:

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

**Proof of (9):** We show that there are a.s. no such sets violating (9) in $G_{n,p=\rho/n}$. We fix $|S| = \sigma n \leq \frac{e}{20k} n$, $|X| = x n \geq \sigma n$ and we let $\rho$ be the solution to $\left( \frac{25 \rho}{\gamma} \right)^{\frac{1}{2}} \frac{\rho}{\sigma} = \frac{1}{2c^2}$. By Lemma 3.1(a), a.s. $K$ 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 $\lambda(X, (S \cup N(S)) \backslash X) > \frac{1}{2} \gamma k |X|$. We first choose $S, X$, then expose $N(S)$. We assume that $|S \cup N(S)| \leq \rho n$ and bound the probability, under that assumption, that $\lambda(X, (S \cup N(S)) \backslash X) > \frac{1}{2} \gamma k |X|$. This yields a bound of at most:

$$
\left( \frac{n}{\sigma n} \right) \left( \frac{n}{x n} \right) \left( \frac{(\rho n)(x n)}{\frac{1}{2} \gamma k x n} \right) \left( \frac{c}{\frac{1}{n}} \right) \frac{1}{2} \gamma k x n
$$

$$
< \left( \frac{e}{\sigma} \right) \left( \frac{e}{x} \right) \left( \frac{e \rho}{\frac{1}{2} \gamma k} \right) \frac{1}{2} \gamma k x n
$$

$$
< \left( \frac{e}{\sigma} \right)^{2x n} \left( \frac{25 \rho}{\gamma} \right)^{\frac{1}{2} \gamma k x n} \frac{1}{2} \gamma k x n
$$

since $\sigma \leq x$ and $c < 2k$

$$
= \left( \frac{1}{2} \right)^{x n} \frac{1}{2} \gamma k x n
$$

$$
= \left( \frac{1}{2} \right)^{x n} \frac{1}{2} \gamma k x n
$$

The sum of $\left( \frac{1}{2} \right)^{x n}$ 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, a.s. there are no sets \( S, X \) violating (9) with \(|X| \geq \sqrt{\log n}\) and with \(|S \cup N(S)| \leq \rho n\). By Lemma 3.1(a), this implies that a.s. there are no \( S, X \) violating (9) 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 \( \ell \) 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 \(|V(H)| = |S| + |X| \leq \ell + 2 |X|\). But \( H \) has at most \(|S| + |X| + \frac{1}{2} \gamma k |X| = O(\sqrt{\log n})\) vertices. So by Lemma 1.5, a.s. no such \( H \) exists. \( \Box \)

Our final bound is that a.s. for every set \( S \subset V(K) \) with \(|S| \leq \frac{\epsilon n}{20k} n\) we have:

\[
\lambda(S, N(S)) \leq |N(S)| + \frac{k}{4} |S|. \tag{7}
\]

**Proof of (7):** Again, we prove that (7) a.s. holds for every such \( S \) in \( G_{n,p} = c/n \). Consider any set \( S \) of \( \sigma n \) vertices. Let \( \nu \) be the solution to \((25\nu)^{k/4}/\sigma = \frac{1}{2k}\). By Lemma 3.1(b), a.s. \(|N(S)| \leq \nu n\).

We expose the edges from \( S \) to \( N(S) \) as follows: 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 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 \( \frac{1}{4} |S| \) edges are exposed during the second step. So the expected number of sets \( S \) of size \( \sigma n \) that violate (7) and for which \(|N(S)| \leq \nu n\) is at most:

\[
\left( \frac{n}{\sigma n} \right) \times \left( \sigma n \times \nu n \right) \left( \frac{c}{n} \right)^{\frac{k}{4} |S|} \leq \left( \frac{e}{\sigma} \right)^{\sigma n} \left( \frac{ec \sigma n^2}{k} \right)^{\frac{k}{4} |S|} \leq \left( \frac{e}{\sigma} \right)^{\sigma n} (25\nu)^{\frac{k}{4} |S|} < \left( \frac{e}{\sigma} \right)^{\sigma n} \text{ since } c < 2k \]

\[
= \left( \frac{1}{2} \right)^{\sigma n} .
\]

The sum of \( \left( \frac{1}{2} \right)^{|S|} \) over all \(|S| \geq \log n\) is \( o(1) \). For the case where \(|S| < \log n\), we use the well known fact that a.s. the maximum degree in \( G_{n,p} = c/n \) is less than \( \log n \) (see eg. Exercise

\[\text{Exercise}\]

\[\text{We are grateful to an anonymous referee for suggesting this approach, which is simpler and more elegant than our original proof.}\]
3.5 in [2]). Thus, replacing υn above by |S| log n we obtain a smaller bound of

\[
\left( \frac{en}{|S|} \right)^{|S|} \left( \frac{25|S| \log n}{n} \right)^{\frac{|S|}{2}} < \frac{1}{n^2}.
\]

Multiplying this by the log \( n \) choices for \(|S| < \log n\) yields that the expected number of violating sets is \( o(1) \), thus establishing that a.s. (7) holds for all such \( S \).

\[\square\]

Acknowledgements

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

References


