Sharp threshold for $K_4$-percolation

Brett Kolesnik

Abstract

Graph bootstrap percolation is a variation of bootstrap percolation introduced by Bollobás. Let $H$ be a graph. Edges are added to an initial graph $G = (V, E)$ if they are in a copy of $H$ minus an edge, until no further edges can be added. If eventually the complete graph on $V$ is obtained, $G$ is said to $H$-percolate. We identify the sharp threshold for $K_4$-percolation on the Erdős-Rényi graph $G_{n,p}$. This refines a result of Balogh, Bollobás and Morris, which bounds the threshold up to multiplicative constants.

1 Introduction

Fix a graph $H$. Following Bollobás [4], $H$-bootstrap percolation is a cellular automaton that adds edges to a graph $G = (V, E)$ by iteratively completing all copies of $H$ missing a single edge. Formally, given a graph $G_0 = G$, let $G_{i+1}$ be $G_i$ together with every edge whose addition creates a subgraph that is isomorphic to $H$. For a finite graph $G$, this procedure terminates once $G_{\tau+1} = G_{\tau}$, for some $\tau = \tau(G)$. We denote the resulting graph $G_{\tau}$ by $\langle G \rangle_H$. If $\langle G \rangle_H$ is the complete graph on $V$, the graph $G$ is said to $H$-percolate, or equivalently, that $G$ is $H$-percolating.

Recall that the Erdős-Rényi [6] graph $G_{n,p}$ is the random subgraph of $K_n$ obtained by including each possible edge independently with probability $p$. In this work, we identify the sharp threshold for $K_4$-percolation on $G_{n,p}$.

**Theorem 1.1.** Let $p = \sqrt{\alpha/(n \log n)}$. If $\alpha > 1/3$ then $G_{n,p}$ is $K_4$-percolating with high probability. If $\alpha < 1/3$ then with high probability $G_{n,p}$ does not $K_4$-percolate.

Angel and Kolesnik [2] established the super-critical case $\alpha > 1/3$, via a connection with 2-neighbour bootstrap percolation (see [1, Section 1.1]). It thus remains to study the sub-critical case $\alpha < 1/3$. In this case, we also identify the size of the largest $K_4$-percolating subgraphs of $G_{n,p}$.
Theorem 1.2. Let $p = \sqrt{\alpha/(n \log n)}$, for some $\alpha \in (0, 1/3)$. With high probability the largest cliques in $\langle G_{n,p} \rangle K_4$ are of size $(\beta_* + o(1)) \log n$, where $\beta_*(\alpha) \in (0, 1/3)$ satisfies $\frac{3}{2} + \beta_* \log(\alpha\beta) - \alpha\beta^2/2 = 0$.

From the results in [2], it follows that with high probability $\langle G_{n,p} \rangle K_4$ has cliques of size at least $(\beta_* + o(1)) \log n$. Our contribution is to show that these are typically the largest cliques.

Balogh, Bollobás and Morris [3] study $H$-bootstrap percolation in the case that $G = G_{n,p}$ and $H = K_k$. The case $k = 4$ is the minimal case of interest. Indeed, all graphs $K_2$-percolate, and a graph $K_3$-percolates if and only if it is connected. Therefore the case $K_3$ follows by a classical result of Erdős and Rényi [6]. If $p = (\log n + \epsilon)/n$ then $G_{n,p}$ is $K_3$-percolating with probability $\exp(-e^{-\epsilon}(1 + o(1)))$, as $n \to \infty$.

Critical thresholds for $H$-bootstrap percolation are defined in [3] by

$$p_c(n, H) = \inf \{ p > 0 : \mathbb{P}(\langle G_{n,p} \rangle H = K_n) \geq 1/2 \}.$$  

In light of [Theorem 1.1] we find that $p_c(n, K_4) \sim 1/\sqrt{3n \log n}$, solving Problem 2 in [3]. Moreover, the same holds if the $1/2$ in the definition above is replaced by any probability in $(0, 1)$. It is expected that this property has a sharp threshold for $H = K_k$ for all $k$, in the sense that for some $p_c = p_c(k)$ we have that $G_{n,p}$ is $K_k$-percolating with high probability for $p > (1 + \delta)p_c$ and with probability tending to 0 for $p = (1 - \delta)p_c$. Some bounds for $p_c(n, K_k)$ are established in [3]. A main result of [3] is that $p_c(n, K_4) = \Theta(1/\sqrt{n \log n})$. For larger $k$ even the order of $p_c$ is open.

1.1 Seed edges

In [2, Theorem 1.2], a sharp upper bound for $p_c(n, K_4)$ is established by observing a connection with 2-neighbour bootstrap percolation (see Pollak and Riess [8] and Chalupa, Leath and Reich [5]). This process is defined as follows: Let $G = (V, E)$ be a graph. Given some initial set $V_0 \subset V$ of activated vertices, let $V_{t+1}$ be the union of $V_t$ and the set of all vertices with at least 2 neighbours in $V_t$. The sets $V_t$ are increasing, and so converge to some set of eventually active vertices, denoted by $\langle V_0, G \rangle_2$. A set $I$ is called contagious for $G$ if it activates all of $V$, that is, $\langle I, G \rangle_2 = V$. (Note that, despite the similar notation, $\langle \cdot \rangle_2$ has a different meaning than $\langle \cdot \rangle_H$ above for graphs $H$. In the present article, we only use $\langle \cdot \rangle_2$ and $\langle \cdot \rangle_{K_4}$.)

If $G = (V, E)$ has a contagious pair $\{u, v\}$, and moreover $(u, v) \in E$, then clearly $G$ is $K_4$-percolating (see [2, Lemma 1.3]). In this case we call $(u, v)$ a
seed edge and $G$ a seed graph. Hence $G$ is a seed graph if some contagious pair of $G$ is joined by an edge.

While it is possible for a graph to be $K_4$-percolating without containing a seed edge (see Section 2), we believe that the two properties are fairly close. In particular, they have the same asymptotic threshold. In [2] the sharp threshold for the existence of contagious pairs in $G_{n,p}$ is identified, and is shown to be $1/(2\sqrt{n\log n})$. It is also shown that if $p = \sqrt{\alpha/(n\log n)}$, then for $\alpha > 1/3$ with high probability $G_{n,p}$ has a seed edge, and so is $K_4$-percolating. If $\alpha < 1/3$ then the largest seed subgraphs of $G_{n,p}$ are of size $(\beta_* + o(1)) \log n$ with high probability, where $\beta_*$ is as defined in Theorem 1.2.

1.2 Outline

By the results in [2] discussed in the previous Section 1.1, to prove Theorems 1.1 and 1.2 it remains to establish the following result.

Proposition 1.3. Let $p = \sqrt{\alpha/(n\log n)}$, for some $\alpha \in (0, 1/3)$. For any $\delta > 0$, with high probability $\langle G_{n,p} \rangle_{K_4}$ contains no clique larger than $(\beta_* + \delta) \log n$, where $\beta_*$ is as defined in Theorem 1.2.

In other words, we need to rule out the possibility that some subgraph of $G_{n,p}$ is $K_4$-percolating and larger than $(\beta_* + \delta) \log n$.

For a graph $G = (V,E)$, let $V(G) = V$ and $E(G) = E$ denote its vertex and edge sets. For $H \subset G$, let $\langle H, G \rangle_2$ denote the subgraph of $G$ induced by $\langle V(H), G \rangle_2$ (see Section 1.1). It is easy to see that if $H \subset G$ is $K_4$-percolating, then so is $\langle H, G \rangle_2$. In particular, $G$ is a seed graph if $\langle e, G \rangle_2 = G$ for some seed edge $e \in E(G)$. On the other hand, if a $K_4$-percolating graph $G$ is not a seed graph, we show that there is some $K_4$-percolating subgraph $C \subset G$ of minimum degree 3 such that $\langle C, G \rangle_2 = G$. We call $C$ the 3-core of $G$. Hence, to establish Proposition 1.3 we require bounds for (i) the number of $K_4$-percolating graphs $C$ of size $q$ with minimum degree 3, and (ii) the probability that for a given set $I \subset [n]$ of size $q$ we have that $|\langle I, G_{n,p} \rangle_2| \geq k$.

We obtain an upper bound of $(2/e)^q q! q^q$ for the number of $K_4$-percolating 3-cores $C$ of size $q$. (This is much smaller than the number of seed subgraphs of size $q$, which in [2] is shown to be equal to $q! q^q e^{o(q)}$.) Further arguments imply that, for $p$ as in Proposition 1.3, with high probability $G_{n,p}$ has no such subgraphs $C$ larger than $(2\alpha)^{-1} \log n$. This already gives a strong indication that $1/3$ is indeed the critical constant, since as shown by Janson, Łuczak, Turova and Vallier [7, Theorem 3.1], $(2\alpha)^{-1} \log n$ is the critical size above which a random set is likely to be contagious.
Recently, Angel and Kolesnik [1] developed large deviation estimates for the probability that small sets of vertices eventually activate a relatively large set of vertices via the $r$-neighbour bootstrap percolation dynamics. These bounds complement the central limit theorems of [7]. This result, in the case of $r = 2$, plays an important role in the current work. For $2 \leq q \leq k$, let $P(q, k)$ denote the probability that for a given set $I \subset [n]$, with $|I| = q$, we have that $|\langle I, \mathcal{G}_{n,p}\rangle I| \geq k$.

**Lemma 1.4** ([1] Lemma 3.2). Let $p = \sqrt{\alpha / (n \log n)}$, for some $\alpha > 0$. Let $\varepsilon \in [0, 1)$ and $\beta \in [\beta_\varepsilon, 1/\alpha]$, where $\beta_\varepsilon = (1 - \sqrt{1 - \varepsilon})/\alpha$. Put $k_\alpha = \alpha^{-1} \log n$ and $q_\alpha = (2\alpha)^{-1} \log n$. Suppose that $q/q_\alpha \to \varepsilon$ and $k/k_\alpha \to \alpha \beta$ as $n \to \infty$. Then $P(q, k) = n^{\xi_\varepsilon + o(1)}$, where $\xi_\varepsilon = \xi_\varepsilon(\alpha, \beta)$ is equal to

$$\frac{\alpha \beta^2}{2} + \left\{ \begin{array}{ll}
(2\alpha \beta - \varepsilon)(2\alpha)^{-1} \log(e(\alpha \beta)^2/(2\alpha \beta - \varepsilon)), & \beta \in [\beta_\varepsilon, \varepsilon/\alpha); \\
\beta \log(\alpha \beta) - \varepsilon(2\alpha)^{-1} \log(\varepsilon/e), & \beta \in [\varepsilon/\alpha, 1/\alpha].
\end{array} \right.$$

(This estimate follows by [1] Lemma 3.2, setting $r = 2$, $\vartheta = (4\alpha)^{-1} \log n$ and $\delta = \alpha \beta$, in which case, in the notation of [1], we have $k_2 = k_\alpha$, $\ell_2 = q_\alpha$ and $\delta_\varepsilon = \alpha \beta_\varepsilon.$) Applying the lemma and the bound $(2/e)^q q^q$ for the number of $K_4$-percolating $3$-cores of size $q$, we deduce that the expected number of $K_4$-percolating subgraphs of $\mathcal{G}_{n,p}$ of size $k = \beta \log n$, for some $\beta \in [\beta_\varepsilon, 1/\alpha]$, is bounded by $n^{\mu + o(1)}$, where

$$\mu(\alpha, \beta) = 3/2 + \beta \log(\alpha \beta) - \alpha \beta^2/2,$$

leading to **Proposition 1.3**.

In closing, we remark that the proof of [2, Proposition 2.1] shows that the expected number of edges in $\mathcal{G}_{n,p}$ that are a seed edge for a subgraph of size at least $k = \beta \log n$, for $\beta \in (0, 1/\alpha]$, is bounded by $n^{\mu + o(1)}$. (Alternatively, we recover this bound from the case $\varepsilon = 0$ in [Lemma 1.4].) This suggests that perhaps $\mathcal{G}_{n,p}$ is as likely to $K_4$-percolate due to a seed edge as in any other way. That being said, the precise behaviour in the **scaling window** (the range of $p$ where $\mathcal{G}_{n,p}$ is $K_4$-percolating with probability in $[\varepsilon, 1 - \varepsilon]$) remains an interesting open problem. As mentioned above, the case of $K_3$-percolation follows by fundamental work of Erdős and Rényi [6]: With high probability $\mathcal{G}_{n,p}$ is $K_3$-percolating (equivalently, connected) if and only if it has no isolated vertices. It seems possible that $K_4$-percolation is more complicated. Perhaps, for $p$ in the scaling window, the probability that $\mathcal{G}_{n,p}$ has a seed edge converges to a constant in $(0, 1)$, and with non-vanishing probability $\mathcal{G}_{n,p}$ is $K_4$-percolating due to a small 3-core $C$ of size $O(1)$ such that $|\langle C, \mathcal{G}_{n,p}\rangle C| = n$. We hope to investigate this further in future work.
2 Clique processes

If a graph $G$ is $K_4$-percolating, we will often simply say that $G$ percolates, or that it is percolating. Following [3], we define the clique process, as a way to analyze $K_4$-percolation on graphs.

**Definition 2.1.** We say that three graphs $G_i = (V_i, E_i)$ form a triangle if there are distinct vertices $x, y, z$ such that $x \in V_1 \cap V_2$, $y \in V_1 \cap V_3$ and $z \in V_2 \cap V_3$. If $|V_i \cap V_j| = 1$ for all $i \neq j$, we say that the $G_i$ form exactly one triangle.

In [3] the following observation is made.

**Lemma 2.2.** Suppose that $G_i = (V_i, E_i)$ percolate.

(i) If $|V_1 \cap V_2| > 1$ then $G_1 \cup G_2$ percolates.

(ii) If the $G_i$ form a triangle then $G_1 \cup G_2 \cup G_3$ percolates.

Moreover, if the $G_i$ form multiple triangles (that is, if there are multiple triplets $x, y, z$ as above), then the percolation of $G_1 \cup G_2 \cup G_3$ follows by applying Lemma 2.2(ii) twice. Indeed, some $G_i, G_j$ have two vertices in common, and so $G' = G_i \cup G_j$ percolates, and $G'$ has two common vertices with the remaining graph $G_k$.

By these observations, the $K_4$-percolation dynamics are classified in [3] as follows (which we modify slightly here in light of the previous observation).

**Definition 2.3.** A clique process for a graph $G$ is a sequence $(S_t)_{t=1}^\tau$ of sets of subgraphs of $G$ with the following properties:

(i) $S_0 = E(G)$ is the edge set of $G$.

(ii) For each $t < \tau$, $S_{t+1}$ is constructed from $S_t$ by either (a) merging two subgraphs $G_1, G_2 \in S_t$ with at least two common vertices, or (b) merging three subgraphs $G_1, G_2, G_3 \in S_t$ that form exactly one triangle.

(iii) $S_\tau$ is such that no further operations as in (ii) are possible.

**Lemma 2.4.** Let $G$ be a finite graph and $(S_t)_{t=1}^\tau$ a clique process for $G$. For each $t \leq \tau$, $S_t$ is a set of edge-disjoint, percolating subgraphs of $G$. Furthermore, $\langle G \rangle_{K_4}$ is the edge-disjoint, triangle-free union of the cliques $\langle H \rangle_{K_4}, H \in S_\tau$. Hence $G$ percolates if and only if $S_\tau = \{G\}$. In particular, if two clique processes for $G$ terminate at $S_\tau$ and $S_\tau'$, then necessarily $S_\tau = S_\tau'$.
2.1 Consequences

The following corollaries of Lemma 2.4 are proved in [3].

**Lemma 2.5.** If $G = (V, E)$ percolates then $|E| \geq 2|V| - 3$.

In light of this, we define the excess of a percolating graph $G = (V, E)$ to be $|E| - (2|V| - 3)$. We call a percolating graph *edge-minimal* if its excess is 0. To prove Lemma 2.5, the following observations are made in [3].

**Lemma 2.6.** Suppose that $G_i = (V_i, E_i)$ percolate.

(i) If the $G_i$ form exactly one triangle, then the excess of $G_1 \cup G_2 \cup G_3$ is the sum of the excesses of the $G_i$.

(ii) If $|V_1 \cap V_2| = m \geq 2$, then the excess of $G_1 \cup G_2$ is the sum of the excesses of the $G_i$ plus $2m - 3$.

Hence, if $G$ is edge-minimal and percolating, then every step of any clique process for $G$ involves merging three subgraphs that form exactly one triangle. A special class of percolating graphs are *seed graphs*, as discussed in Section 1.1. In an edge-minimal seed graph $G$, every step of some clique process for $G$ involves merging three subgraphs, two of which are a single edge.

Finally, since in each step of any clique process for a graph $G$ either 2 or 3 subgraphs are merged, we have the following useful criterion for percolation.

**Lemma 2.7.** Let $G = (V, E)$ be a graph of size $n$, and $1 \leq k \leq n$. If there is no percolating subgraph $G' \subset G$ of size $k'$, for any $k' \in [k, 3k]$, then $G$ has no percolating subgraph larger than $k$. In particular, $G$ does not percolate.

3 Percolating graphs

In this section, we analyze the general structure of percolating graphs.

**Definition 3.1.** We say that a graph $G$ is irreducible if removing any edge from $G$ results in a non-percolating graph.

Clearly, a graph $G$ is percolating if and only if it has an irreducible percolating subgraph $G' \subset G$ such that $V(G) = V(G')$.

For a graph $G$ and vertex $v \in V(G)$, we let $G_v$ denote the subgraph of $G$ induced by $V - \{v\}$, that is, the subgraph obtained by removing $v$.

**Lemma 3.2.** Let $G$ be an irreducible percolating graph. If $v \in V(G)$ is of degree 2, then $G_v$ is percolating.
**Proof.** The proof is by induction on the size of $G$. The case $|V(G)| = 3$, in which case $G$ is a triangle, is immediate. Hence suppose that $G$, with $|V(G)| > 3$, percolates and some $v \in V(G)$ is of degree 2, and assume that the statement of the lemma holds for all graphs $H$ with $|V(H)| < |V(G)|$.

Let $(S_t)_{t=1}^{\tau}$ be a clique process for $G$. Let $e_1, e_2$ denote the edges incident to $v$ in $G$. Let $t < \tau$ be the first time in the clique process $(S_t)_{t=1}^{\tau}$ that a subgraph containing either $e_1$ or $e_2$ is merged with other (edge-disjoint, percolating) subgraphs. We claim that $S_{t+1}$ is obtained from $S_t$ by merging $e_1, e_2$ with a subgraph in $S_t$. To see this, we first observe that if a graph $H$ percolates and $|V(H)| > 2$ (that is, $H$ is not simply an edge), then all vertices in $H$ have degree at least 2. Next, by the choice of $t$, we note that none of the graphs being merged contain both $e_1, e_2$. Therefore, since $v$ is of degree 2, if one the graphs contains exactly one $e_i$, then it is necessarily equal to $e_i$, being a percolating graph of minimum degree 1. It follows that $v$ is contained in two of the graphs being merged, and hence that $S_{t+1}$ is the result of merging the edges $e_1, e_2$ with a subgraph in $S_t$, as claimed.

To conclude, note that if $t = \tau - 1$ then since $G$ percolates (and so $S_\tau = \{G\}$) we have that $S_{\tau-1} = \{e_1, e_2, G_v\}$, and so $G_v$ percolates. On the other hand, if $t < \tau - 1$, then $S_t$ contains 2 or 3 subgraphs, one of which contains $e_1$ and $e_2$. If $S_{\tau-1} = \{G_1, G_2\}$, where $e_1, e_2 \in E(G_1)$, say, then by the inductive hypothesis we have that $(G_1)_v$ percolates. Since $G_1, G_2$ are edge-disjoint, we have that $v \notin V(G_2)$, as otherwise $G_2$ would be a percolating graph with an isolated vertex. Hence, by Lemma 2.2(i), we find that $(G_1)_v \cup G_2 = G_v$ percolates. Similarly, if $S_{\tau-1} = \{G_1, G_2, G_3\}$, where $e_1, e_2 \in E(G_1)$, say, then by the inductive hypothesis and Lemma 2.2(ii), we find that $(G_1)_v \cup G_2 \cup G_3 = G_v$ percolates.

The induction is complete. $\blacksquare$

Recall (see Sections 1.1 and 1.2) that for graphs $H \subset G$, we let $\langle H, G \rangle_2$ denote the subgraph of $G$ induced by $\langle V(H), G \rangle_2$, that is, the subgraph of $G$ induced by the closure of $V(H)$ under the 2-neighbour bootstrap percolation dynamics on $G$. By Lemma 2.2(i), if $H \subset G$ is percolating then so is $\langle H, G \rangle_2$.

The following is an immediate consequence of Lemma 3.2.

**Lemma 3.3.** Let $G$ be an irreducible percolating graph. Then either

(i) $G = \langle e, G \rangle_2$ for some edge $e \in E(G)$, or else,

(ii) $G = \langle C, G \rangle_2$ for some percolating subgraph $C \subset G$ of minimum degree at least 3.

Furthermore,

(iii) the excess of $G$ is equal to the excess of $C$. 

7
We note that in case (i), $G$ is a seed graph and $e$ is a seed edge for $G$. In case (ii), we call $C$ the 3-core of $G$. If $G = C$ we say that $G$ is a 3-core.

It is straightforward to verify that all irreducible percolating graphs on $2 < k \leq 6$ vertices have a vertex of degree 2. There is however an edge-minimal percolating graph of size $k = 7$ with no vertex of degree 2, see Figure 1.

![Figure 1: The smallest irreducible percolating 3-core.](image)

### 3.1 Basic estimates

In this section, we use Lemma 3.3 to obtain upper bounds for irreducible percolating graphs. For such a graph $G$, the relevant quantities are the number of vertices in $G$ of degree 2, the size of its 3-core $C \subset G$, and its number of excess edges.

**Definition 3.4.** Let $I^\ell_q(k,i)$ be the number of labelled, irreducible graphs $G$ of size $k$ with an excess of $\ell$ edges, $i$ vertices of degree 2, and a 3-core $C \subset G$ of size $2 < q \leq k$. If $i = 0$, and hence $k = q$, we let $C^\ell_q(k) = I^\ell_k(k,0)$.

In the case $\ell = 0$, we will often simply write $I(k,i)$ and $C(k)$.

By Lemma 3.3(iii), if a graph $G$ contributes to $I^\ell_q(k,i)$ then its 3-core $C \subset G$ has an excess of $\ell$ edges. Also, as noted above, there are no irreducible 3-cores on $k \leq 6$ vertices. Hence, $I^\ell_q(k,i) = 0$ if $2 < q \leq 6$.

**Definition 3.5.** We define $I_2^\ell(k,i)$ to be the number of labelled, edge-minimal seed graphs of size $k$ with $i$ vertices of degree 2.

For convenience, we let $C(2) = 1$ and set $I^\ell_2(k,i) = 0$ and $C^\ell(2) = 0$ for $\ell > 0$ (in light of Lemma 3.3(iii)). Moreover, to simplify several statements in this work, if we say that a graph $G$ has a 3-core of size less than $q > 2$, we mean to include also the possibility that $q = 2$.

**Definition 3.6.** We let $I^\ell_q(k,i) = \sum_q I^\ell_i(k,i)$ denote the number of labelled, irreducible graphs $G$ of size $k$ with an excess of $\ell$ edges and $i$ vertices of degree 2.
We obtain the following estimate for $I^\ell(k, i)$ in the case that $\ell \leq 3$, that is, for graphs with at most 3 excess edges.

**Lemma 3.7.** For all $k \geq 2$, $\ell \leq 3$ and relevant $i$, we have that

$$I^\ell(k, i) \leq \left(\frac{2}{e}\right)^kk!k^{k+2\ell+i}.$$ 

In particular, $C^\ell(k) \leq \left(\frac{2}{e}\right)^kk!k^{2\ell}$. 

The method of proof gives bounds for larger $\ell$, however, as it turns out, percolating graphs with a larger excess can be dealt with using less accurate estimates (see Lemma 4.3).

The proof is somewhat involved, as there are several cases to consider, depending on the nature of the last step of a clique process for $G$. We proceed by induction: First, we note that the cases $i > 0$ follow easily, since if $G$ has $i$ vertices of degree 2, then removing such a vertex from $G$ results in a graph with $j \in \{i, i \pm 1\}$ vertices of degree 2. Analyzing this case leads to the constant $2/e$. The case $i = 0$ (corresponding to 3-cores) is the heart of the proof. The following observation allows the induction to go through in this case: If $G$ is a percolating 3-core, then in the last step of a clique process for $G$ either (i) three graphs $G_1, G_2, G_3$ are merged that form exactly one triangle on $T = \{v_1, v_2, v_3\}$, or else (ii) two graphs $G_1, G_2$ are merged that share exactly $m \geq 2$ vertices $S = \{v_1, v_2, \ldots, v_m\}$. We note that if some $G_j$ has a vertex $v$ of degree 2, then necessarily $v \in T$ in case (i), and $v \in S$ in case (ii) (as else, $G$ would have a vertex of degree 2). In other words, if a percolating 3-core is formed by merging graphs with vertices of degree 2, then all such vertices belong to the triangle that they form or the set of their common vertices.

**Proof.** It is easily verified that the statement of the lemma holds for $k \leq 4$. We prove the remaining cases by induction. For $k > 4$, we claim moreover that for all $\ell \leq 3$ and relevant $i$,

$$I^\ell(k, i) \leq A\zeta^kk!k^{k+2\ell} \quad (3.1)$$

where $\zeta = 2/e$ and $A = 6/(\zeta^55!5^5)$. The lemma follows, noting that $A < 1$ and $\binom{k}{i} \leq k^i$.

We introduce the constant $A < 1$ in order to push through the induction in the case $i = 0$, corresponding to 3-cores. The last step of a clique process for such a graph $G$ involves merging 2 or 3 subgraphs $G_i$. Informally, we
use the constant $A$ to penalize graphs $G$ such that at least two of the $G_i$ contain more than 4 vertices (that is, graphs $G$ formed by merging at least two “macroscopic” subgraphs).

By the choice of $A$, we have that (3.1) holds for $k = 5$. Indeed, note that $I(5, i) \leq \binom{5}{i} \binom{4}{2}$ for all $i \in \{1, 2, 3\}$ and $I^0(5, i) = 0$ otherwise. Assume that for some $k > 5$, (3.1) holds for all $4 < k' < k$, and all $\ell \leq 3$ and relevant $i$.

We begin with the case of graphs $G$ of size $k$ with at least one vertex of degree 2. This case follows easily by a recursive upper bound (and explains the choice of $\zeta = 2/e$).

Case 1 ($i > 0$). Suppose that $G$ is a graph contributing to $I^\ell(k, i)$, where $i > 0$. Let $v \in V(G)$ be the vertex of degree 2 in $G$ with the smallest index. By considering which two of the $k - i$ vertices of $G$ are neighbours of $v$, we find that $I^\ell(k, i)$ is bounded from above by

$$\binom{k}{i} \binom{k - i}{2} \sum_{j=0}^{2} \binom{2}{j} \frac{I^\ell(k - 1, i - 1 + j)}{\binom{k-1}{i-1+j}}.$$ 

In this sum, $j \in \{0, 1, 2\}$ is the number of neighbours of $v$ that are of degree 2 in the subgraph of $G$ induced by $V(G) - \{v\}$. Applying the inductive hypothesis, we obtain

$$I^\ell(k, i) \leq A \zeta^k \binom{k}{i} k! k^{k+2\ell} \cdot \frac{2}{\zeta} \left(\frac{k - 1}{k}\right)^k \leq A \zeta^k \binom{k}{i} k! k^{k+2\ell},$$

as required.

The remaining cases deal with 3-cores $G$ of size $k$, where $i = 0$. First, we establish the case $i = \ell = 0$, corresponding to edge-minimal percolating 3-cores. The cases $i = 0$ and $\ell \in \{1, 2, 3\}$ are proved by adapting the argument for $i = \ell = 0$.

Case 2 ($i = \ell = 0$). Let $G$ be a graph contributing to $C(k) = I(k, 0)$. Then, by Lemma 2.6, in the last step of a clique process for $G$, three edge-minimal percolating subgraphs $G_j$, $j \in \{1, 2, 3\}$, are merged which form exactly one triangle on some $T = \{v_1, v_2, v_3\} \subset V(G)$. Moreover, each $G_j$ has at most 2 vertices of degree 2, and if some $G_j$ has such a vertex $v$ then necessarily $v \in T$. Also if $k_j = |V(G_j)|$, with $k_1 \geq k_2 \geq k_3$, then $\sum k_j = k + 3$, $k_3 \geq 2$ or $k_3 \geq 4$, and $k_1, k_2 \geq 4$ (since if some $k_j = 3$ or some $k_j = k'_j = 2$, with $j \neq j'$, then $G$ would have a vertex of degree 2).

Since the inductive hypothesis only holds for graphs with more than 4 vertices, it is convenient to deal with the case $k_1 = 4$ separately: It is easily verified that the only edge-minimal percolating 3-cores of size $k$ with all
$k_1 \leq 4$ are of size $k \in \{7, 9\}$. These graphs are the graph in Figure 1 and the graph obtained from this graph by replacing the bottom edge with a copy of $K_4$ minus an edge. Hence it is straightforward to verify that the claim holds if $k \in \{7, 9\}$, and so in the arguments below we may assume that $k_1 > 4$. Moreover, since the graph in Figure 1 is the only irreducible percolating 3-core on $k = 7$ vertices, we may further assume that $k \geq 8$.

We take three cases, with respect to whether (i) $k_2 = 4$, (ii) $k_2 > 4$ and $k_3 \in \{2, 4\}$, or (iii) $k_3 > 4$.

**Case 2(i)** ($i = \ell = 0$ and $k_2 = 4$). Note that if $k_2 = 4$ then $k_3 \in \{2, 4\}$.

The number of graphs $G$ as above with $k_3 = 2$ and $k_2 = 4$ is bounded from above by

$$\binom{k}{k-3} \binom{k-3}{2} \binom{3}{1} 2^2 \sum_{j=0}^{2} \binom{2}{j} \frac{I(k-3, j)}{\binom{k-3}{j}}.$$  

Here the first binomial selects the vertices for the subgraph of size $k_1 = k - 3$, the next two binomials select the vertices for the triangle $T$, and the rightmost factor bounds the number of possibilities for the subgraph of size $k_1 = k - 3$ (recalling that it can have at most 2 vertices of degree 2, and if it contains any such vertex $v$, then $v \in T$). Applying the inductive hypothesis (recall that we may assume that $k_1 > 4$), the above expression is bounded by

$$A\zeta^k k! k^k \cdot \frac{(k-3)^{k-1}}{k^k} \frac{4}{4} \leq A\zeta^k k! k^k \cdot \frac{1}{k} \frac{4}{k^3 e^3}.$$  

Similarly, the number of graphs $G$ as above such that $k_1 = k_2 = 4$ is bounded by

$$\binom{k}{k-5} \binom{k-5}{2} \binom{3}{1} 2^3 \sum_{j=0}^{2} \binom{2}{j} \frac{I(k-5, j)}{\binom{k-5}{j}}.$$  

By the inductive hypothesis, this is bounded by

$$A\zeta^k k! k^k \cdot \frac{(k-5)^{k-3}}{k^k} \frac{4}{4} \zeta^5 \leq A\zeta^k k! k^k \cdot \frac{1}{k^3} \frac{4}{\zeta^5 e^5}.$$  

Altogether, we find that the number of graphs $G$ contributing to $C(k)$ with $k_2 = 4$, divided by $A\zeta^k k! k^k$, is bounded by

$$\gamma_1(k) = \frac{1}{k} \frac{4}{k^3 e^3} + \frac{1}{k^3} \frac{4}{\zeta^5 e^5}. \tag{3.2}$$
Case 2(ii) \((i = \ell = 0, k_2 > 4 \text{ and } k_3 \in \{2, 4\})\). For a given \(k_1, k_2 > 4\), the number of graphs \(G\) as above with \(k_3 = 2\) (in which case \(k_1 + k_2 = k + 1\)) is bounded by

\[
\left( \frac{k}{k_1, k_2 - 1} \right) \left( \frac{k_1}{2} \right) \left( \frac{k_2 - 1}{1} \right) 2^{k_3} \prod_{j=1}^{2} \sum_{i=0}^{2} \left( \begin{array}{c} 2 \\ i \end{array} \right) I(k_j, i) \left( \begin{array}{c} k_j \\ i \end{array} \right).
\]

Applying the inductive hypothesis, this is bounded by

\[
A \zeta^k k! k^k \cdot 2 \cdot 4^2 A \zeta^k_{k_1 + 2, k_2 + 2} \frac{k_1}{k_2}.
\]

Since \(k_2 = k + 1 - k_1\), we have that

\[
\frac{\partial}{\partial k_1} k_{k_1 + 2, k_2 + 2} = -k_{k_1 + 1, k_2 + 1} (k_1 k_2 \log(k_2/k_1) - 2(k_1 - k_2)).
\]

By the bound \(\log x \leq x - 1\), we see that

\[
k_1k_2 \log(k_2/k_1) - 2(k_1 - k_2) \leq -(k_2 + 2)(k_1 - k_2) \leq 0.
\]

Hence, setting \(k_1\) to be the maximum relevant value \(k_1 = k - 4\) (when \(k_2 = 5\)), we find

\[
\frac{k_{k_1 + 2, k_2 + 2}^k}{k^k} \leq \frac{5^7(k - 4)^{k-2}}{k^k} \leq \frac{1}{k^2} \frac{5^7}{e^4}
\]

for all relevant \(k_1, k_2\). Therefore, summing over the at most \(k/2\) possibilities for \(k_1, k_2\), we find that at most

\[
A \zeta^k k! k^k \cdot \frac{1}{k} A \zeta^2 4^2 5^7
\]

graphs \(G\) with \(k_3 = 2\) and \(k_2 > 4\) contribute to \(C(k)\).

The case of \(k_3 = 4\) is very similar. In this case, for a given \(k_1, k_2 > 4\) such that \(k_1 + k_2 = k - 1\), the number of graphs \(G\) as above is bounded by

\[
\left( \frac{k}{k_1, k_2 - 1, 2} \right) \left( \frac{k_1}{2} \right) \left( \frac{k_2 - 1}{1} \right) 2^{k_3} \prod_{j=1}^{2} \sum_{i=0}^{2} \left( \begin{array}{c} 2 \\ i \end{array} \right) I(k_j, i) \left( \begin{array}{c} k_j \\ i \end{array} \right),
\]

which, by the inductive hypothesis, is bounded by

\[
A \zeta^k k! k^k \cdot 2 \cdot 4^2 A \zeta^k_{k_1 + 2, k_2 + 2} \frac{k_1}{k_2}.
\]

12
Arguing as in the previous case, we see that the above expression is maximized when \( k_2 = 5 \) and \( k_1 = k - 6 \). Hence, summing over the at most \( k/2 \) possibilities for \( k_1, k_2 \), there are at most
\[
A \zeta^k k! k^k \cdot \frac{1}{k^3 \zeta} \cdot \frac{A 4^{2} 5^{7}}{e^{6}}
\]
graphs \( G \) that contribute to \( C(k) \) with \( k_3 = 4 \) and \( k_2 > 4 \).

We conclude that the number of graphs \( G \) that contribute to \( C(k) \) with \( k_2 > 4 \) and \( k_3 \in \{2, 4\} \), divided by \( A \zeta^k k! k^k \), is bounded by
\[
\gamma_2(k) = \frac{1}{k} A 4^{2} 5^{7} + \frac{1}{k^3 \zeta} \cdot \frac{A 4^{2} 5^{7}}{e^{6}}.
\]

**Case 2(iii) \((i = \ell = 0 \text{ and } k_3 > 4)\)**. For a given \( k_1, k_2, k_3 > 4 \) such that \( k_1 + k_2 + k_3 = k + 3 \), the number of graphs \( G \) as above is bounded by
\[
\left( \begin{array}{c} k \\ k_1, k_2 - 1, k_3 - 2 \end{array} \right) \left( \begin{array}{c} k_1 \\ 2 \end{array} \right) \left( \begin{array}{c} k_2 - 1 \\ 1 \end{array} \right) 2^{t_j} \sum_{j=1}^{3} \sum_{i=0}^{2} \left( \begin{array}{c} 2 \\ i \end{array} \right) I(k_j, i) \left( \begin{array}{c} k_j \\ i \end{array} \right).
\]

By the inductive hypothesis, this is bounded by
\[
A \zeta^k k! k^k \cdot 2^{2} 4^{3} A 2^{4} 3^{k_1+2} k_2+2 \frac{k_3+2}{k^k}.
\]

As in the previous cases considered, the above expression is maximized when \( k_2 = k_3 = 5 \) and \( k_1 = k - 7 \). Hence, summing over the at most \((k/2)^2\) choices for the \( k_i \), we find that the number of graphs \( G \) that contribute to \( C(k) \) with \( k_3 > 4 \), divided by \( A \zeta^k k! k^k \), is bounded by
\[
\gamma_3(k) = \frac{1}{k^3} \frac{A 2^{2} 3^{5} 5^{14} 4^{3}}{e^{7}}.
\]

Finally, combining (3.2), (3.3) and (3.4), we find that, for all \( k \geq 8 \),
\[
\frac{C(k)}{A \zeta^k k! k^k} \leq \sum_{i=1}^{3} \gamma_i(k) \leq \sum_{i=1}^{3} \gamma_i(8) \approx 0.23 < 1,
\]
completing the proof of Case 2.

It remains to consider the cases \( i = 0 \) and \( \ell \in \{1, 2, 3\} \), corresponding to 3-cores \( G \) with a non-zero excess. In these cases, it is possible that only 2 subgraphs are merged in the last step of a clique process for \( G \). We prove
the cases $\ell = 1, 2, 3$ separately, however they all follow by adjusting the proof of the Case 2.

First, we note that if two graphs $G_1, G_2$ are merged to form an irreducible percolating 3-core $G$, then necessarily each $G_i$ contains more than 4 vertices. This allows us to apply the inductive hypothesis in these cases (recall that we claim that (3.1) holds only for graphs with more than 4 vertices), without taking additional sub-cases as in the proof of Case 2. Moreover, recall that as noted above Case 2(i), we may assume in all of the following cases that $k \geq 8$ (since there are no irreducible percolating 3-cores on $k < 7$ vertices, and the case $k = 7$, corresponding to the graph in Figure 1, is easily verified).

Case 3 ($i = 0$ and $\ell = 1$). If $G$ contributes to $C^1(k)$, then by Lemma 2.6 in the last step of a clique process for $G$, there are two cases to consider:

(i) Three percolating subgraphs $G_j, j \in \{1, 2, 3\}$, are merged which form exactly one triangle $T = \{v_1, v_2, v_3\}$, such that for some $i_j \leq 2$ and $k_j, \ell_j \geq 0$ with $\sum k_j = k + 3$ and $\sum \ell_j = 1$, we have that $G_j$ contributes to $I^{b_j}(k_j, i_j)$. Moreover, if any $i_j > 0$, the corresponding $i_j$ vertices of $G_j$ of degree 2 belong to $T$.

(ii) Two percolating subgraphs $G_j, j \in \{1, 2\}$, are merged that share exactly two vertices $S = \{v_1, v_2\}$, such that for some $i_j \leq 2$ and $k_j$ with $\sum k_j = k + 2$, we have that the $G_j$ contribute to $I(k_j, i_j)$. Moreover, if any $i_j > 0$, the corresponding $i_j$ vertices of $G_j$ of degree 2 belong to $S$.

We claim that, by the arguments in Case 2 leading to (3.5) and the inductive hypothesis, the number of graphs $G$ satisfying (i), divided by $A\zeta^k k! k^{k+2}$, is bounded by

$$\gamma_1(k) + 2\gamma_2(k) + 3\gamma_3(k).$$

(3.6)

Indeed, note that if one of the graphs $G_i$ has an excess edge, then necessarily $k_i > 4$. Moreover, recall that graphs $G$ that contribute to $C(k)$, as considered in Cases 2(i),(ii),(iii) above, have exactly 1, 2, 3 such subgraphs $G_i$ with $k_i > 4$. The total number of such graphs $G$ is bounded by $\gamma_1(k), \gamma_2(k), \gamma_3(k)$, respectively, in these cases. Hence the claim follows, noting that if $G_i$ has an excess edge, then it contributes an extra factor of $k_i^2 < k^2$.

On the other hand, arguing along the lines as in Case 2, the number of graphs $G$ satisfying (ii), for a given $k_1, k_2 > 4$ such that $k_1 + k_2 = k + 2$, is bounded by

$$\binom{k}{k_1, k_2 - 2} \binom{k_1}{2} 2^{\ell_2} \prod_{j=1}^2 \sum_{i=0}^2 \binom{2}{i} I(k_j, i) (k_j).$$

14
By the inductive hypothesis, this is bounded by

\[ A \zeta^k k! k^k \cdot 2 \cdot 4^2 \frac{A \zeta^2 k_1^{k_1} k_2^{k_2}}{k^k}. \]

Arguing as in Case 2, we find that this expression is maximized when \( k_2 = 5 \) and \( k_1 = k - 3 \). Hence, summing over the at most \( k/2 \) choices for \( k_1, k_2 \), the number of graphs \( G \) satisfying (ii), divided by \( A \zeta^k k! k^k + 2 \), is at most

\[ \gamma_4(k) = \frac{1}{k^2} \frac{A \zeta^2 4^2 5^7}{e^3}. \]

(3.7)

Altogether, by (3.6) and (3.7), we conclude that, for all \( k \geq 8 \),

\[ \frac{C^1(k)}{A \zeta^k k! k^k + 2} \leq \gamma_1(8) + 2\gamma_2(8) + 3\gamma_3(8) + \gamma_4(8) \approx 0.43 < 1, \]

(3.8)

completing the proof of Case 3.

**Case 4** (\( i = 0 \) and \( \ell = 2 \)). This case is nearly identical to Case 3. By [Lemma 2.6](#) in the last step of a clique process for a graph \( G \) that contributes to \( C^2(k) \), either (i) three graphs that form exactly one triangle are merged whose excesses sum to 2, or else (ii) two graphs that share exactly two vertices are merged whose excesses sum to 1. Hence, by the arguments in Case 3 leading to (3.8), we find that, for all \( k \geq 8 \),

\[ \frac{C^2(k)}{A \zeta^k k! k^k + 4} \leq \gamma_1(8) + 3\gamma_2(8) + 6\gamma_3(8) + 2\gamma_4(8) \approx 0.63 < 1, \]

(3.9)

as required.

**Case 5** (\( i = 0 \) and \( \ell = 3 \)). Since \( \ell = 3 \), it is now possible that in the last step of a clique process for a graph \( G \) contributing to \( C^\ell(k) \), two graphs are merged that share 3 vertices. Apart from this difference, the argument is completely analogous to the previous cases.

If \( G \) contributes to \( C^3(k) \), then by [Lemma 2.6](#) in the last step of a clique process for \( G \), there are three cases to consider:

(i) Three percolating subgraphs \( G_{j}, j \in \{1, 2, 3\} \), are merged which form exactly one triangle \( T = \{v_1, v_2, v_3\} \), such that for some \( i_j \leq 2 \) and \( k_j, \ell_j \geq 0 \) with \( \sum k_j = k + 3 \) and \( \sum \ell_j = 3 \), we have that \( G_{i} \) contributes to \( I^k(k_j, i_j) \). If any \( i_j > 0 \), the corresponding \( i_j \) vertices of \( G_{j} \) of degree 2 belong to \( T \).

(ii) Two percolating subgraphs \( G_{j}, j \in \{1, 2\} \), are merged that share exactly two vertices \( S = \{v_1, v_2\} \), such that for some \( i_j \leq 2 \) and \( k_j, \ell_j \geq 0 \).
with $\sum k_j = k + 2$ and $\sum \ell_j = 2$, we have that the $G_j$ contribute to $I^{\ell_j}(k_j, i_j)$. If any $i_j > 0$, the corresponding $i_j$ vertices of $G_j$ of degree 2 belong to $S$.

(iii) Two percolating subgraphs $G_j$, $j \in \{1, 2\}$, are merged that share exactly three vertices $R = \{v_1, v_2, v_3\}$, such that for some $i_j \leq 3$ with $\sum k_j = k + 3$, we have that the $G_j$ contribute to $I(k_j, i_j)$. If any $i_j > 0$, the corresponding $i_j$ vertices of $G_j$ of degree 2 belong to $R$.

As in Case 4, we find by the arguments in Case 3 leading to (3.8) that the number of graphs $G$ satisfying (i) or (ii), divided by $A\zeta^k k! k^{k+6}$, is bounded by

$$\gamma_1(k) + 4\gamma_2(k) + 10\gamma_3(k) + 3\gamma_4(k).$$

(3.10)

On the other hand, by the arguments in Case 3 leading to (3.7), the number of graphs $G$ satisfying (iii), for a given $k_1, k_2 > 4$ such that $k_1 + k_2 = k + 3$, is bounded by

$$\left(\begin{array}{c} k \\ k_1, k_2 - 3 \end{array}\right) \left(\begin{array}{c} k_1 \\ 3 \end{array}\right) 3!^2 \prod_{j=1}^2 \sum_{i=0}^{3} \left(\begin{array}{c} 3 \\ i \end{array}\right) I(k_j, i) \frac{I(k_j, i)}{\binom{k}{i}}.$$

By the inductive hypothesis, this is bounded by

$$A\zeta^k k! k^6 \cdot 3!^2 A\zeta^3 \frac{k_1^{k_1+3} k_2^{k_2+3}}{k^k}.$$

This expression is maximized when $k_2 = 5$ and $k_1 = k - 2$. Hence, summing over the at most $k/2$ choices for $k_1, k_2$, the number of graphs $G$ satisfying (iii), divided by $A\zeta^k k! k^{k+6}$, is at most

$$\gamma_5(k) = \frac{1}{k^4} A\zeta^3 3!^5 8^2 2e^2.$$

(3.11)

Therefore, by (3.10) and (3.11), for all $k \geq 8$ we have that

$$\frac{C_3(k)}{A\zeta^k k! k^{k+6}} \leq \gamma_1(8) + 4\gamma_2(8) + 10\gamma_3(8) + 3\gamma_4(8) + \gamma_5(8) \approx 0.90 < 1,$$

completing the proof of Case 5.

This last case completes the induction. We conclude that (3.1) holds for all $k > 4$, $\ell \leq 3$ and relevant $i$. As discussed, Lemma 3.7 follows. ■
3.2 Sharper estimates

In this section, using Lemma 3.7, we obtain upper bounds for $I^\ell_q(k,i)$, which improve on those for $I^\ell(k,i)$ given by Lemma 3.7, especially when $q$ is significantly smaller than $k$. These are used in Section 5 to rule out the existence of large percolating subgraphs of $\mathcal{G}_{n,p}$ with few vertices of degree 2 and small 3-cores.

Lemma 3.8. Let $\varepsilon > 0$. For some constant $\vartheta(\varepsilon) \geq 1$, the following holds.

For all $k \geq 2$, $\ell \leq 3$, and relevant $q,i$, we have that

$$I^\ell_q(k,i) \leq \vartheta \psi_\varepsilon(q/k) k!k^{k+2\ell+i}$$

where

$$\psi_\varepsilon(y) = \max\{3/(2e) + \varepsilon, (e/2)^{1-2y^2}\}.$$

This lemma is only useful for $\varepsilon < 1/(2e)$, as otherwise $\psi_\varepsilon(y) \geq 2/e$ for all $y$, and so Lemma 3.7 gives a better bound. Note that, for any $\varepsilon < 1/(2e)$, we have that $\psi_\varepsilon(y)$ is non-decreasing and $\psi_\varepsilon(y) \to 2/e$ as $y \uparrow 1$, in agreement with Lemma 3.7. Moreover, $\psi_\varepsilon(y) = 3/(2e) + \varepsilon$ for $y \leq y_\ast$ and $\psi_\varepsilon(y) = (e/2)^{1-2y^2}$ for $y > y_\ast$, where $y_\ast = y_\ast(\varepsilon)$ satisfies

$$(e/2)^{1-2y_\ast^2} = 3/(2e) + \varepsilon.$$ 

We define $\hat{y} = y_\ast(0) \approx 0.819$, and note that $y_\ast(\varepsilon) \downarrow \hat{y}$, as $\varepsilon \downarrow 0$.

The general scheme of the proof is as follows: First, we note that the case $i = k - q$ follows easily by Lemma 3.7, since $I^\ell_q(k,k-q)$ is equal to $(k^{-q}/(q!))^{k-q}C^\ell(q)$. We establish the remaining cases by induction, noting that if a graph $G$ contributes to $I^\ell_q(k,i)$ and $i < k - q$, then there is a vertex $v \in V(G)$ of degree 2 with a neighbour not in the 3-core $C \subset G$. Therefore, either (i) some neighbour of $v$ is of degree 2 in the subgraph of $G$ induced by $V(G) - \{v\}$, or else (ii) there are vertices $u \neq w \in V(G)$ of degree 2 with a common neighbour not in $C$. This observation leads to an improved bound (when $q < k$) for $I^\ell_q(k,i)$ compared with that for $I^\ell_q(k,i)$ given by Lemma 3.7.

Proof. Let $\varepsilon > 0$ be given. We may assume that $\varepsilon < 1/(2e)$, as otherwise the statement of lemma follows by Lemma 3.7, noting that for any $q$, $I^\ell_q(k,i) \leq I^\ell(k,i)$. We claim that, for some $\vartheta(\varepsilon) \geq 1$ (to be determined below), for all $k \geq 2$, $\ell \leq 3$ and relevant $q,i$, we have that

$$I^\ell_q(k,i) \leq \vartheta {k\choose i} \psi_\varepsilon(q/k) k!k^{k+2\ell}.$$ 

(3.13)
Case 1 \((i = k - q)\). We first observe that [Lemma 3.7] implies the case \(i = k - q\). Indeed, if \(q = k\), in which case \(i = 0\), then (3.13) follows immediately by [Lemma 3.7] noting that \(I^\ell_k(k,0) = C^\ell(k)\) and \(\psi(1) = 2/e\).

On the other hand, if \(i = k - q > 0\) then

\[
I^\ell_q(k,k - q) = \left(\frac{k}{k - q}\right) \left(\frac{q}{2}\right)^{k - q} C^\ell(q),
\]

since all \(k - q\) vertices of degree 2 in a graph that contributes to \(I^\ell_q(k,k - q)\) are connected to 2 vertices in its 3-core of size \(q\). We claim that the right hand side is bounded by

\[
\left(\frac{k}{k - q}\right)(e/2)^{k - 2q}(q/k)^{2k}k!k^{2\ell}.
\]

Since \((e/2)^{k - 2q}(q/k)^{2k} \leq \psi(q/k)^k\), (3.13) follows. To see this, note that by [Lemma 3.7], we have that

\[
1 + \frac{2}{k - 1} \left(\frac{k - 2}{k - 1}\right)^k \psi_c(q/(k - 2))^{k - 2} \psi_c(q/(k - 1))^{k - 1} = 1 + O(1/k) \leq 1 + \delta,
\]

where

\[
\delta = \delta(\varepsilon) = \min \left\{1 - \frac{3(1 - y^*_s)^2}{3(2e)\varepsilon} + \varepsilon, 1 - \frac{3(1 - y^*_s)}{y^*_s}\right\}.
\]

Note that, since \(3(1 - y)/y^2 < 1\) for all \(y > (\sqrt{21} - 3)/2 \approx 0.791\), and recalling (see (3.12)) that \(y^*_s > \hat{y} \approx 0.819\), we have that \(\delta > 0\).

Select \(\delta(\varepsilon) \geq 1\) so that (3.13) holds for all \(k \geq k_\varepsilon\) and relevant \(q, \ell, i\). By Case 1, we have that \[3.13\] holds for all \(k, q\) in the case that \(i = k - q\). We establish the remaining cases \(i < k - q\) by induction: Assume that for some \(k > k_\varepsilon\), \[3.13\] holds for all \(k' < k\) and relevant \(q, \ell, i\).

In any graph \(G\) contributing to \(I^\ell_q(k,i)\), where \(i < k - q\), there is some vertex of degree 2 with at least one of its two neighbours not in the 3-core of \(G\). There are two cases to consider: either
(i) there is a vertex $v$ of degree 2 such that one of its two neighbours is of degree 2 in the subgraph of $G$ induced by $V(G) - \{v\}$, or else,

(ii) there is no such vertex $v$, however there are vertices $u \neq w$ of degree 2 with a common neighbour that is not in the 3-core of $G$.

Note that, in case (i), removing $v$ results in a graph with $j \in \{i, i+1\}$ vertices of degree 2. On the other hand, in case (ii), removing $u, w$ results in a graph with $j \in \{i-2, i-1, i\}$ vertices of degree 2. By considering the vertices $v$ or $u, w$ as above with minimal labels, we see that, for $i < k - q$, $I_q^i(k, i) / \binom{k}{i}$ is bounded by

$$I_q^i(k - 1, i + 1) \binom{k - i - q}{2} + I_q^i(k - 1, i) \binom{k - i - q}{k - i} (k - i-q)(k-i) + (k-i-q)(k-i)^2 \sum_{j=0}^{2} I_q^i(k-2, i-2+j) \binom{k-2}{i-2+j}.$$

Applying the inductive hypothesis, it follows that

$$I_q^i(k, i) \vartheta(k, i) \psi_q^i(q/k)^{k! k^{k+2 \ell}} \leq \Psi^i(q,k) \left[ 1 + \frac{2}{k-1} \left( \frac{k-2}{k} \right)^k \left( \frac{k}{k-1} \right) \psi_q^i(q/(k-2)^{k-2}) \right]$$

where

$$\Psi^i(q,k) = \frac{3k-q}{2k} \left( \frac{k-1}{k} \right)^k \frac{\psi_q^i(q/(k-1))^{k-1}}{\psi_q^i(q/k)^k}.$$

By the choice of $k$, and since $k \geq k^\varepsilon$, it follows that

$$I_q^i(k, i) \vartheta(k, i) \psi_q^i(q/k)^{k! k^{k+2 \ell}} \leq \Psi^i(q,k)(1 + \delta). \quad (3.14)$$

Next, we show that $\Psi^i(q,k) < 1 - \delta$, completing the induction. To this end, we take cases with respect to whether (i) $q/(k-1) \leq y_*$, (ii) $y_* \leq q/k$, or (iii) $q/k < y_* < q/(k-1)$.

**Case 2(i) $q/(k-1) \leq y_*$.** In this case $\psi_q^i(q/m) = \psi_q^i(q/m) = 3/(2e) + \varepsilon$, for each $m \in \{k-1, k\}$. It follows, by the choice of $\delta$, that

$$\Psi^i(q,k) \leq \left( \frac{k-1}{k} \right)^k \frac{3/2}{3/(2e) + \varepsilon} \leq \frac{3/(2e)}{3/(2e) + \varepsilon} < 1 - \delta,$$

as required.
Case 2(ii) \((y_\ast \leq q/k)\). In this case, for each \(m \in \{k - 1, k\}\), we have that \(\psi(q/m)^m = (e/2)^{m-2q}(q/m)^{2m}\). Hence

\[
\Psi_\varepsilon(q, k) = \frac{3}{e} \left( \frac{k}{k-1} \right)^{k-1} \frac{(k-q)(k-1)}{q^2} \leq \frac{3(1 - y)}{y^2},
\]

where \(y = q/k\). Since the right hand side is decreasing in \(y\), we find, by the choice of \(\delta\), that

\[
\Psi_\varepsilon(q, k) \leq \frac{3(1 - y_\ast)}{y_\ast^2} < 1 - \delta.
\]

Case 2(iii) \((q/k < y_\ast < q/(k - 1))\). In this case, \(\psi_\varepsilon(q/k) = 3/(2e) + \varepsilon\) and \(\psi_\varepsilon(q/(k - 1))^{k-1} = (e/2)^{k-1-2q}(q/(k - 1))^{2(k-1)}\). Hence

\[
\Psi_\varepsilon(q, k) = \frac{3}{e} \left( \frac{k}{k-1} \right)^{k-1} \frac{(k-q)(k-1)(e/2)^{k-2q(q/k)^2k}}{q^2} \frac{(3/(2e) + \varepsilon)^k}{3/(2e) + \varepsilon}.
\]

As in the previous case, we consider the quantity \(y = q/k\). The above expression is bounded by

\[
\frac{3(1 - y)}{y^2} \left( \frac{e/2}{3/(2e) + \varepsilon} \right)^k y^2.
\]

We claim that this expression is increasing in \(y \leq y_\ast\). By (3.12) and the choice of \(\delta\), it follows that

\[
\Psi_\varepsilon(q, k) \leq \frac{3(1 - y_\ast)}{y_\ast^2} < 1 - \delta,
\]

as required. To establish the claim, simply note that

\[
\frac{\partial}{\partial y} \left( \frac{2y}{e} \right)^{2k} = \frac{1}{y^3} \left( \frac{2y}{e} \right)^{2k} (2(1 - y)k + y - 2) > 0
\]

for all \(y \leq y_\ast\), since \(k \geq k_\varepsilon \geq 1/(1 - y_\ast)\).

Altogether, we conclude that \(\Psi_\varepsilon(q, k) \leq 1 - \delta\), for all relevant \(q\). By (3.14), it follows that

\[
I_\ell^k(q, i) \leq (1 - \delta^2) \vartheta \left( \frac{k}{i} \right) \psi_\varepsilon(q/k)^k k! k^{k+2\ell} \vartheta \left( \frac{k}{i} \right) \psi_\varepsilon(q/k)^k k! k^{k+2\ell},
\]

completing the induction. We conclude that (3.13) holds for \(k \geq 2, \ell \leq 3\) and relevant \(q, i\). Since \(\binom{k}{i} \leq k^i\), the lemma follows. \[\blacksquare\]
4 Percolating subgraphs with small cores

With Lemmas 1.4, 3.7, and 3.8 at hand, we begin to analyze the structure of percolating subgraphs of $G_{n,p}$. In this section, we show that for sub-critical $p$, with high probability $G_{n,p}$ has no percolating subgraphs larger than $(\beta^{*} + o(1)) \log n$ with a small 3-core. The non-existence of large percolating 3-cores is verified in the next Section 5, completing the proof of Proposition 1.3.

More specifically, we prove the following result.

Lemma 4.1. Let $p = \sqrt{\alpha/(n \log n)}$, for some $\alpha \in (0, 1/3)$. Then, for any $\delta > 0$, with high probability $G_{n,p}$ has no irreducible percolating subgraph $G$ of size $k = \beta \log n$ with a 3-core $C \subset G$ of size $q \leq (3/2) \log n$, for any $\beta \geq (\beta^{*} + \delta) \log n$.

Recall that (as discussed in Section 3.1), in statements such as this lemma, we mean also to include the possibility that $q = 2$ (corresponding to a seed graph $G$) when we say that the 3-core of a graph $G$ is less than $q > 2$.

First, we justify the definition of $\beta^{*}$ in Theorem 1.2.

Lemma 4.2. Fix $\alpha \in (0, 1/3)$. For $\beta > 0$, let

$$\mu(\alpha, \beta) = 3/2 + \beta \log(\alpha \beta) - \alpha \beta^{2}/2.$$  

The function $\mu(\alpha, \beta)$ is decreasing in $\beta$, with a unique zero $\beta^{*}(\alpha) \in (0, 3)$.

In particular, for $\alpha \in (0, 1/3)$, we have that $\beta^{*} \leq 1/\alpha$.

Proof. Differentiating $\mu(\alpha, \beta)$ with respect to $\beta$, we obtain $1 + \log(\alpha \beta) - \alpha \beta$. Since $\log x < x - 1$ for all positive $x \neq 1$, we find that $\mu(\alpha, \beta)$ is decreasing in $\beta$. Moreover, since $\alpha < 1/3$, we have that $\mu(\alpha, 3) < (3/2)(3\alpha - 1) < 0$. The result follows, noting that $\mu(\alpha, \beta) \to 3/2 > 0$ as $\beta \downarrow 0$. $lacksquare$

Recall that the bounds in Section 3.1 apply only to graphs with an excess of $\ell \leq 3$ edges. The following observation is useful for dealing with graphs with a larger excess.

Lemma 4.3. Let $p = \sqrt{\alpha/(n \log n)}$, for some $\alpha \in (0, 1/3)$. Then with high probability $G_{n,p}$ contains no subgraph of size $k = \beta \log n$ with an excess of $\ell$ edges, for any $\beta \in (0, 2]$ and $\ell > 3$, or any $\beta \in (0, 9]$ and $\ell > 27$.

Proof. The expected number of subgraphs of size $k = \beta \log n$ in $G_{n,p}$ with an excess of $\ell$ edges is bounded by

$$\binom{n}{k} \binom{\binom{k}{2}}{2k - 3 + \ell} p^{2k - 3 + \ell} \leq \left(\frac{e^{3}}{16 knp}\right)^{k} \left(\frac{e}{4 kp}\right)^{\ell - 3} \leq n^{\nu} \log^{\ell} n$$

21
where
\[ \nu(\beta, \ell) = -(\ell - 3)/2 + \beta \log(\alpha \beta e^3/16). \]
Note that \( \nu \) is convex in \( \beta \) and \( \nu(\beta, \ell) \to -(\ell - 3)/2 \) as \( \beta \downarrow 0 \). Note also that
\[ 2 \log(2/3 \cdot e^3/16) \approx -0.356 < 0 \]
and
\[ 9 \log(9/3 \cdot e^3/16) \approx 11.934 < 12. \]
Therefore \( \nu(2, \ell) < -(\ell - 3)/2 \) and \( \nu(9, \ell) < -(\ell - 27)/2 \). Hence, the first claim follows by summing over all \( k \leq 2 \log n \) and \( \ell > 3 \). The second claim follows, summing over all \( k \leq 9 \log n \) and \( \ell > 27 \).

**Definition 4.4.** Let \( E(q, k) \) denote the expected number of irreducible percolating 3-cores \( C \subset G_{n,p} \) of size \( q \) (or edges, if \( q = 2 \)), such that \( |\langle C, G_{n,p}\rangle_2| \geq k \).

Combining Lemmas 1.4, 3.7 and 4.3, we obtain the following estimate. Recall \( \mu \) as defined in Lemma 4.2.

**Lemma 4.5.** Let \( p = \sqrt{\alpha/(n \log n)} \), for some \( \alpha \in (0, 1/3) \). Let \( \varepsilon \in [0, 3\alpha] \) and \( \beta_{\varepsilon}, k_{\alpha}, q_{\alpha} \) be as defined in Lemma 1.4. Let \( \beta \in [\beta_{\varepsilon}, 1/\alpha] \). Suppose that \( q/q_{\alpha} \to \varepsilon \) and \( k/k_{\alpha} \to \alpha \beta \) as \( n \to \infty \). Then \( E(q, k) \leq n^{\mu_{\varepsilon} + o(1)} \), where \( \mu_{\varepsilon}(\alpha, \beta) = \mu(\alpha, \beta) \) for \( \beta \in [\varepsilon/\alpha, 1/\alpha] \),
\[
\mu_{\varepsilon}(\alpha, \beta) = \mu(\alpha, \beta) - \beta \log(\alpha \beta) + \frac{\varepsilon}{2\alpha} \log(\varepsilon/e) + \frac{2\alpha \beta - \varepsilon}{2\alpha} \log \left( \frac{e(\alpha \beta)^2}{2 \alpha \beta - \varepsilon} \right)
\]
for \( \beta \in [\beta_{\varepsilon}, \varepsilon/\alpha] \), and \( o(1) \) depends only on \( n \).

We note that \( \mu_{\varepsilon}(\alpha, \varepsilon/\alpha) = \mu(\alpha, \varepsilon/\alpha) \), as is easily verified.

**Proof.** By the proof of Lemma 4.3, the expected number of irreducible percolating 3-cores in \( G_{n,p} \) of size \( q \leq (3/2) \log n \) with an excess of \( \ell > 3 \) edges tends to 0 as \( n \to \infty \). Therefore, it suffices to show that, for all \( \ell \leq 3 \), we have that \( E(\ell, q) \leq n^{\mu_{\varepsilon} + o(1)} \), where \( E(\ell, q) \) is the expected number of irreducible percolating 3-cores \( C \subset G_{n,p} \) of size \( q = \varepsilon (2\alpha)^{-1} \log n \) with an excess of \( \ell \) edges, such that \( |\langle C, G_{n,p}\rangle_2| \geq k = \beta \log n \). For such \( \ell \), by Lemmas 1.4 and 3.7 we find that
\[
E(\ell, q) \leq \left( \frac{n}{q} \right) C(\ell)(q)p^{2q-3+\ell} P(q, k)
\]
\[
\leq q^{2\ell} p^{\ell-3} \left( \frac{2}{\varepsilon qnp^2} \right)^q P(q, k) \leq n^{\nu + o(1)}
\]
where
\[ \nu = 3/2 + \varepsilon (2\alpha)^{-1} \log(\varepsilon/e) + \xi_\varepsilon(\alpha, \beta) = \mu_\varepsilon(\alpha, \beta) \]
(and \( \xi_\varepsilon \) is as defined in Lemma 1.4), as required.

We aim to prove Lemma 4.1 by the first moment method. To this end, we first show that for some \( \varepsilon_* \in (0, 3\alpha) \), with high probability there are no irreducible percolating 3-cores in \( G_{n,p} \) of size \( \varepsilon(2\alpha)^{-1} \log n \), for all \( \varepsilon \in (\varepsilon, 3\alpha] \).

Moreover, we establish a slightly more general result that allows for graphs with \( i = O(1) \) vertices of degree 2, which is also used in the next Section 5.

**Lemma 4.6.** Let \( p = \sqrt{\alpha/(n \log n)} \) for some \( \alpha \in (0, 1/3) \). Fix some \( i_* \geq 0 \). Let \( \varepsilon_*(\alpha) \in (0, 3\alpha) \) satisfy \( 3/2 + \varepsilon(2\alpha)^{-1} \log(\varepsilon/e) = 0 \). Then, for any \( \eta > 0 \), with high probability \( G_{n,p} \) has no irreducible percolating subgraph \( G \) of size \( q = \varepsilon(2\alpha)^{-1} \log n \) with \( i \) vertices of degree 2, for any \( i \leq i_* \) and \( \varepsilon \in [\varepsilon_* + \eta, 3\alpha] \).

**Proof.** By Lemma 4.3, it suffices to consider subgraphs \( G \) with an excess of \( \ell \leq 3 \) edges. By Lemma 3.7, the expected number of such subgraphs is bounded by
\[
\left( \frac{n}{q} \right) p^{2q - 3 + \ell} I(q, i) \leq k^{2\ell + 4} (\frac{2}{e} knp^2)^q \leq n^{\nu + o(1)}
\]
where \( \nu(\varepsilon) = 3/2 + \varepsilon (2\alpha)^{-1} \log(\varepsilon/e) \). Noting that \( \nu \) is decreasing in \( \varepsilon < 1 \), and \( \nu(3\alpha) = (3/2) \log(3\alpha) < 0 \), the lemma follows.

Next, we plan to use Lemma 4.5 to rule out the remaining cases \( \varepsilon \leq \varepsilon_* + \eta \) (where \( \eta > 0 \) is a small constant, to be determined below). In order to apply Lemma 4.5, we first verify that for such \( \varepsilon \), we have that \( \beta_* \) is within the range of \( \beta \) specified by Lemma 4.5, that is, \( \beta_* \geq \beta_\varepsilon \).

**Lemma 4.7.** Fix \( \alpha \in (0, 1/3) \). Let \( \beta_\varepsilon, \beta_*, \varepsilon_* \) be as defined in Lemmas 1.4, 4.2 and 4.6. Then, for some sufficiently small \( \eta(\alpha) > 0 \), we have that \( \beta_* \geq \beta_\varepsilon \) for all \( \varepsilon \in [0, \varepsilon_* + \eta] \).

**Proof.** By Lemma 4.2 and the continuity of \( \mu(\alpha, \beta_\varepsilon) \) in \( \varepsilon \), it suffices to show that \( \mu(\alpha, \beta_\varepsilon) > 0 \), for all \( \varepsilon \in [0, \varepsilon_*] \). Let \( \delta_\varepsilon = 1 - \sqrt{1 - \varepsilon} \), so that \( \beta_\varepsilon = \delta_\varepsilon / \alpha \). Note that
\[
\mu(\alpha, \beta_\varepsilon) = 3/2 + (2\alpha)^{-1} (2\delta_\varepsilon \log \delta_\varepsilon + \delta_\varepsilon^2).
\]
Therefore, by the bound \( \log x \leq x - 1 \),
\[
\frac{\partial}{\partial \varepsilon} \mu(\alpha, \beta_\varepsilon) = (2\alpha)^{-1} (1 + \log(\delta_\varepsilon)/(1 - \delta_\varepsilon)) \leq 0.
\]
It thus suffices to verify that \( \mu(\alpha, \beta, \varepsilon) > 0 \). To this end note that, by the definition of \( \varepsilon \) (see Lemma 4.6),

\[
\mu(\alpha, \beta, \varepsilon) = (2\alpha)^{-1}(2\delta, \log \delta - \delta^2 - \varepsilon \log(\varepsilon/e)).
\]

By Lemma 4.6 we have that \( \delta_* = \delta(2 - \delta) \in (0, 1) \), and so \( \delta_* \in (0, 1) \). Hence the lemma follows if we show that \( \nu(\delta) > 0 \) for all \( \delta \in (0, 1) \), where

\[
\nu(\delta) = 2\delta \log \delta - (2 - \delta) \log(\delta(2 - \delta)/e).
\]

Note that

\[
\nu(\delta)/\delta = \delta \log \delta - (2 - \delta) \log(2 - \delta) + 2(1 - \delta).
\]

Differentiating this expression with respect to \( \delta \), we obtain

\[
\epsilon/\delta - 1 - \log(\delta/(2\delta - \epsilon)) \leq 0,
\]

by the inequality \( \log x \geq (x-1)/x \). Noting that \( \nu(1) = 0 \), the lemma follows.

It can be shown that for all sufficiently large \( \epsilon < \varepsilon_* \), we have that \( \beta_* < \epsilon/\alpha \). Therefore, we require the following bound.

**Lemma 4.8.** Fix \( \alpha \in (0, 1/3) \). Let \( \varepsilon \in [0, 1) \) and \( \beta, \mu \) be as defined in Lemmas 1.4 and 4.7. Then \( \mu(\alpha, \beta) \leq \mu(\alpha, \beta) \), for all \( \beta \in [\beta, 1/\alpha] \).

**Proof.** Since \( \mu(\alpha, \beta) = \mu(\alpha, \beta) \) for \( \beta \in [\epsilon/\alpha, 1/\alpha] \), we may assume that \( \beta < \epsilon/\alpha \). Let \( \delta = \alpha \beta \). Then

\[
\alpha(\mu(\alpha, \beta) - \mu(\alpha, \beta)) = \delta \log \delta - \frac{\epsilon}{2} \log(\epsilon/e) - \frac{2\delta - \epsilon}{2} \log\left(\frac{e\delta^2}{2\delta - \epsilon}\right).
\]

Differentiating this expression with respect to \( \delta \), we obtain

\[
\epsilon/\delta - 1 - \log(\delta/(2\delta - \epsilon)) \leq 0,
\]

by the inequality \( \log x \geq (x-1)/x \). Since \( \mu(\alpha, \epsilon/\alpha) = \mu(\alpha, \epsilon/\alpha) \), the lemma follows.

Finally, we prove the main result of this section.

**Proof of Lemma 4.4.** Select \( \eta > 0 \) as in Lemma 4.7. By Lemma 4.6 with high probability \( G_{n,p} \) has no percolating 3-core of size \( q = \epsilon(2\alpha)^{-1} \log n \), for any \( \epsilon \in [\varepsilon_* + \eta, 3\alpha] \). On the other hand, by the choice of \( \eta \), Lemmas 4.5, 4.7 and 4.8 imply that for any \( \beta \in [\beta, 1/\alpha] \), the expected number of irreducible percolating subgraphs of size \( k = \beta \log n \) with a 3-core of size \( q \leq (\epsilon_* + \eta)(2\alpha)^{-1} \log n \) is bounded by \( n^{\mu + o(1)} \), where \( \mu = \mu(\alpha, \beta) \). Hence the result follows by Lemma 4.2 summing over the \( O(\log n) \) possibilities for \( q \).
5 No percolating subgraphs with large cores

In the previous Section 4, it is shown that for sub-critical $p$, with high probability $G_{n,p}$ has no percolating subgraphs larger than $(\beta_* + o(1)) \log n$ with a 3-core smaller than $(3/2) \log n$. In this section, we rule out the existence of larger percolating 3-cores.

**Lemma 5.1.** Let $p = \sqrt{\alpha/(n \log n)}$, for some $\alpha \in (0, 1/3)$. Then with high probability $G_{n,p}$ has no irreducible percolating 3-core $C$ of size $k = \beta \log n$, for any $\beta \in [3/2, 9]$.

Before proving the lemma we observe that it together with Lemma 4.1 implies Proposition 1.3 As discussed in Section 1.2, our main Theorems 1.1 and 1.2 follow.

**Proof of Proposition 1.3.** Let $\delta > 0$ be given. By Lemma 4.2, without loss of generality we may assume that $\beta_* + \delta < 3$. Hence, by Lemmas 2.7 and 3.3, if $G_{n,p}$ has a percolating subgraph that is larger than $(\beta_* + \delta) \log n$, then with high probability it has some irreducible percolating subgraph $G$ of size $k = \beta \log n$ with a 3-core $C \subset G$ of size $q \leq k$ (or a seed edge, if $q = 2$), for some $\beta \in (\beta_* + \delta, 9]$. By Lemma 5.1, with high probability $q \leq (3/2) \log n$. Hence, by Lemma 4.1 with high probability $G_{n,p}$ contains no such subgraph $G$. Therefore, with high probability, all percolating subgraphs of $G_{n,p}$ are of size $k \leq (\beta_* + \delta) \log n$. $\blacksquare$

Towards Lemma 5.1, we observe that $G_{n,p}$ has no percolating subgraph with a small 3-core and few vertices of degree 2.

**Lemma 5.2.** Let $p = \sqrt{\alpha/(n \log n)}$, for some $\alpha \in (0, 1/3)$. Fix some $i_* \geq 1$. With high probability $G_{n,p}$ has no irreducible percolating subgraph $G$ of size $k \geq (3/2) \log n$ with a 3-core $C \subset G$ of size $q \leq (3/2) \log n$ and $i \leq i_*$ vertices of degree 2.

This is a simple consequence of Lemma 3.8 proved by a direct application of the first moment method and elementary calculus.

**Proof.** By Lemmas 3.3 and 4.3, we may assume that if $G_{n,p}$ has an irreducible percolating subgraph $G$ of size $k = \beta \log n$ with a 3-core of size $q \leq (3/2) \log n$, then $G$ has an excess of $\ell \leq 3$ edges. By Lemmas 4.1, 4.2 and 4.6, we may further assume that $\beta \in [3/2, 3]$ and $q = yk$, where $y\beta \in [0, 3/2 - \varepsilon]$, for some $\varepsilon(\alpha) > 0$. Without loss of generality, we assume that $\varepsilon < 1/(2e)$ is such that $\log(3/(2e) + \varepsilon) < -1/2$ (which is possible, since $1 + 2 \log(3/(2e)) \approx -0.189$).
By [Lemma 3.8](#) and since $\alpha < 1/3$, for some constant $\vartheta(\varepsilon) \geq 1$, the expected number of such subgraphs $G$ is bounded by

$$\binom{n}{k} p^{2k-3+\ell} I_{q}^{\ell}(k, i) \leq \vartheta k^{2\ell+i} p^{\ell-3} (kp\psi(y/k))^{k} \ll n^{\nu} \quad (5.1)$$

where

$$\nu(\beta, \psi(y)) = 3/2 + \beta \log(\beta/3) + \beta \log \psi(y)$$

and $\psi(y)$ is as in [Lemma 3.8](#) that is,

$$\psi(y) = \max\{3/(2e) + \varepsilon, (e/2)^{1-2y}y^{2}\}. \quad \text{(3.12)}$$

Recall that $\psi(y) = 3/(2e) + \varepsilon$ for $y < y_{*}$ and $\psi(y) = (e/2)^{1-2y}y^{2}$ for $y > y_{*}$, where $y_{*} = y_{*}(\varepsilon)$ is as defined by (3.12). Moreover, $y_{*} \downarrow \hat{y}$ as $\varepsilon \downarrow 0$, where $\hat{y} \approx 0.819$.

Therefore, to verify that with high probability $G_{n,p}$ has no subgraphs $G$ as in the lemma, we show that $\nu(\beta, \psi(y)) < -\delta$ for some $\delta > 0$ and all $\beta, y$ as above. Moreover, since $\nu$ is convex in $\beta$, it suffices to consider the extreme points $\beta = 3/2$ and $\beta = \min\{3, 3/(2y)\}$ in the range $y \in [0, 1 - \varepsilon']$, where $\varepsilon' = 2\varepsilon/3$.

Since $\psi(y) = 2/\varepsilon$, we have that $\nu(3/2, \psi(y)) = 0$. Hence, for some $\delta_{1} > 0$, we have that $\nu(3/2, \psi_{\varepsilon}(y)) < -\delta_{1}$ for all $y \in [0, 1 - \varepsilon']$. Next, for $\beta = \min\{3, 3/(2y)\}$, we treat the cases (i) $y \in [0, 1/2]$ and $\beta = 3$ and (ii) $y \in [1/2, 1 - \varepsilon']$ and $\beta = 3/(2y)$ separately. If $y \leq 1/2$, then $\psi(y) = 3/(2e) + \varepsilon$, in which case, by the choice of $\varepsilon$,

$$\nu(3, \psi_{\varepsilon}(y)) = \frac{3}{2}(1 + \log(3/(2e) + \varepsilon)) < 0. \quad \text{(5.1)}$$

On the other hand, for $y \geq 1/2$, we need to show that

$$\nu(3/(2y), \psi_{\varepsilon}(y)) = \frac{3}{2} \left(1 + \frac{1}{y} \log \left(\frac{\psi(y)}{2y}\right)\right) < 0. \quad \text{(5.1)}$$

To this end, we first note that differentiating $\nu(3/(2y), 3/(2e) + \varepsilon)$ twice with respect to $y$, we obtain

$$\frac{3}{2y^{2}} \left(3 + 2 \log \left(\frac{3/(2e) + \varepsilon}{2y}\right)\right) \geq \frac{3}{2} \left(3 + 2 \log \left(\frac{3}{4e}\right)\right) \approx 0.637 > 0.$$ 

Therefore it suffices to consider the extreme points $y = 1/2$ and $y = 1$. Noting that, by the choice of $\varepsilon$, we have that

$$\nu(3, 3/(2e)) = \frac{3}{2}(1 + 2 \log(3/(2e) + \varepsilon)) < 0$$

26
and
\[ \nu(3/2, 3/(2e)) = \frac{3}{2} \left( 1 + \log \left( \frac{3/(2e) + \varepsilon}{2} \right) \right) \]
\[ < \frac{3}{2} \left( 1 + 2 \log(3/(2e) + \varepsilon) \right) < 0, \]
it follows that \( \nu(3/(2y), 3/(2e) + \varepsilon) < 0 \) for all \( y \in [1/2, 1] \).

Next, we observe that differentiating \( \nu(3/(2y), (e/2)^{1-2y}y^2) \) with respect to \( y \), we obtain
\[ \frac{3}{2y^2} \left( 1 - \log(ey/4) \right) \geq 3 \log 2 > 0. \]
Therefore, since \( \nu(3/(2y), (e/2)^{1-2y}y^2) \to \nu(3/2, \psi_\varepsilon(1)) = 0 \) as \( y \uparrow 1 \), it follows that \( \nu(3/(2y), (e/2)^{1-2y}y^2) < 0 \) for all \( y \in [1/2, 1 - \varepsilon'] \). Altogether, there is some \( \delta_2 > 0 \) so that \( \nu(\min\{3, 3/(2y)\}, \psi_\varepsilon(y)) < -\delta_2 \) for all \( y \in [0, 1 - \varepsilon'] \).

Put \( \delta = \min\{\delta_1, \delta_2\} \). We conclude that \( \nu(\beta, \psi_\varepsilon(y)) < -\delta \), for all relevant \( \beta, y \). The lemma follows by (5.1), summing over the \( O(\log^2 n) \) choices for \( k \) and \( q \) and \( O(1) \) relevant values \( \ell \leq 3 \) and \( i \leq i_* \). \( \blacksquare \)

With Lemma 5.2 at hand, we turn to Lemma 5.1. The general idea is as follows: Suppose that \( G_{n,p} \) has an irreducible percolating 3-core \( C \) of size \( k = \beta \log n \), for some \( \beta \in [3/2, 9] \). By Lemma 4.3, we can assume that the excess of \( C \) is \( \ell \leq 27 \) edges. Hence, in the last step of a clique process for \( C \), either 2 or 3 percolating subgraphs are merged that have few vertices of degree 2 (as observed above the proof of Lemma 3.7 in Section 3.1). Therefore, by Lemma 5.2, each of these subgraphs is smaller than \( (3/2) \log n \), or else has a 3-core larger than \( (3/2) \log n \).

In this way, we see that considering a minimal such graph \( C \) is the key to proving Lemma 5.1. By Lemma 4.6, there is some \( \beta_1 < 3/2 \) so that with high probability \( G_{n,p} \) has no percolating subgraph of size \( \beta \log n \) with few vertices of degree 2, for all \( \beta \in [\beta_1, 3/2] \). Hence such a graph \( C \), if it exists, is the result of the unlikely event that 2 or 3 percolating graphs, all of which are smaller than \( \beta_1 \log n \) and have few vertices of degree 2, are merged to form a percolating 3-core that is larger than \( (3/2) \log n \). In other words, “macroscopic” subgraphs are merged to form \( C \).

Proof of Lemma 5.1. By Lemma 4.6, there is some \( \beta_1 < 3/2 \) so that with high probability \( G_{n,p} \) has no percolating subgraph of size \( \beta \log n \) with \( i \) vertices of degree 2, for any \( i \leq 15 \) and \( \beta \in [\beta_1, 3/2] \).
Suppose that $G_{n,p}$ has an irreducible $3$-core $C$ of size $k = \beta \log n$ with an excess of $\ell$ edges, for some $\beta \in [3/2, 9]$. By Lemma 4.3, we may assume that $\ell \leq 27$. Moreover, assume that $C$ is of the minimal size among such subgraphs of $G_{n,p}$. By Lemma 2.6, there are two possibilities for the last step of a clique process for $C$:

(i) Three irreducible percolating subgraphs $G_j$, $j \in \{1, 2, 3\}$, are merged which form exactly one triangle $T = \{v_1, v_2, v_3\}$, such that for some $i_j \leq 2$ and $k_j, \ell_j \geq 0$ with $\sum k_j = k + 3$ and $\sum \ell_j = \ell$, we have that the $G_j$ contribute to $I^{\ell_j}(k_j, i_j)$. If any $i_j > 0$, the corresponding $i_j$ vertices of $G_j$ of degree $2$ belong to $T$.

(ii) For some $m \leq (\ell + 3)/2 \leq 15$, two percolating subgraphs $G_j$, $j \in \{1, 2\}$, are merged that share exactly $m$ vertices $S = \{v_1, v_2, \ldots, v_m\}$, such that for some $i_j \leq m$ and $k_j, \ell_j \geq 0$ with $\sum k_j = k + m$ and $\sum \ell_j = \ell - (2m - 3)$, we have that the $G_j$ contribute to $I^{\ell_j}(k_j, i_j)$. If any $i_j > 0$, the corresponding $i_j$ vertices of $G_j$ of degree $2$ belong to $S$.

Moreover, in either case, by the choice of $C$, each $G_j$ is either a seed graph or else has a 3-core smaller than $(3/2) \log n$. Hence, by Lemmas 3.3 and 4.3, we may assume that each $\ell_j \leq 3$. Also, by Lemma 5.2 and the choice of $\beta_1$, we may further assume that all $G_j$ are smaller than $\beta_1 \log n$.

**Case (i).** Let $k, k_j, \ell_j$ be as in (i). Let $k_j - (j - 1) = \varepsilon_j k$, so that $\sum \varepsilon_j = 1$. Without loss of generality we assume that $k_1 \geq k_2 \geq k_3$. Hence $\varepsilon_1, \varepsilon_2$ satisfy $1/3 \leq \varepsilon_1 \leq \beta_1/\beta < 1$ and $(1 - \varepsilon_1)/2 \leq \varepsilon_2 \leq \min\{\varepsilon_1, 1 - \varepsilon_1\}$. The number of 3-cores $C$ as in (i) for these values $k, k_j, \ell_j$ is bounded by

$$
\binom{k}{k-k_1} \binom{k-k_1}{k_3-2} \binom{k_1}{2} \binom{k_2-1}{1} 2^{\varepsilon_2} \prod_{j=1}^{3} \sum_{i=0}^{2} \binom{2}{i} \frac{I^{\ell_j}(k_j, i)}{\binom{k_j}{i}}.
$$

Applying Lemma 3.7 and the inequality $k! < e(k/e)^k$, this is bounded by

$$
\binom{k}{k-k_1} \binom{k-k_1}{k_3-2} \frac{k^3}{2} (8ek^7)^3 \left(\frac{2}{e^2}\right)^{k+3} \prod_{j=1}^{3} k_j^{2k_j}.
$$

By the inequality $\binom{n}{k} < (ne/k)^k$, and noting that

$$
k_j^{2k_j} \leq (ek)^2(j-1)(k_j - (j - 1))^{2(k_j - (j - 1))},
$$

we see that the above expression is bounded, up to a factor of $n^{o(1)}$, by
\[(2e^{-2\eta(\varepsilon_1, \varepsilon_2)})^k k^{2k},\]

where

\[
\eta(\varepsilon_1, \varepsilon_2) = \left( \frac{e}{1 - \varepsilon_1} \right)^{1 - \varepsilon_1} \left( \frac{(1 - \varepsilon_1)e}{\varepsilon_3} \right)^\varepsilon_1 \varepsilon_1^{2\varepsilon_2} \varepsilon_2^{2\varepsilon_3} \\
= \frac{e^{1 - \varepsilon_1 + \varepsilon_3}}{(1 - \varepsilon_1)^{\varepsilon_2}} \varepsilon_1^{2\varepsilon_2} \varepsilon_2^{2\varepsilon_3}.
\]

Since there are \(O(1)\) choices for \(\ell\) and the \(\ell_i\), and since \(\alpha < 1/3\), the expected number of 3-cores \(C\) in \(\mathcal{G}_{n,p}\) of size \(k = \beta \log n\) with \(G_j\) of size \(k_j\) as in (i) is at most

\[
\binom{n}{k} p^{2k-3} \left( \frac{2}{e^2 \eta(\varepsilon_1, \varepsilon_2)k^2} \right)^k n^{o(1)} = p^{-3} \left( \frac{2}{e} \alpha \beta \eta(\varepsilon_1, \varepsilon_2) \right)^k n^{o(1)} \ll n^{\nu} \quad (5.2)
\]

where

\[
\nu(\beta, \varepsilon_1, \varepsilon_2) = \frac{3}{2} + \beta \log \left( \frac{2}{3\varepsilon} \beta \eta(\varepsilon_1, \varepsilon_2) \right).
\]

Since there are \(O(\log^3 n)\) possibilities for \(k\) and the \(k_j\), to show that with high probability \(\mathcal{G}_{n,p}\) has no subgraphs \(C\) as in (i), it suffices to show that for some \(\delta > 0\), we have that \(\nu(\beta, \varepsilon_1, \varepsilon_2) < -\delta\) for all relevant \(\beta\) and \(\varepsilon_j\). Moreover, since \(\nu\) is convex in \(\beta\), we can restrict to the extreme points \(\beta = 3/2\) and \(\beta = 3/(2\varepsilon_1) > \beta_1/\varepsilon_1\). To this end, observe that when \(\beta = 3/2\), we have that \(\nu < 0\) if and only if \(\eta < 1\). Similarly, when \(\beta = 3/(2\varepsilon_1)\), \(\nu < 0\) if and only if \(\eta < \varepsilon_1 e^{1 - \varepsilon_1}\). Since \(\varepsilon_1 e^{1 - \varepsilon_1} \leq 1\) for all relevant \(\varepsilon_1\), it suffices to establish the latter claim. To this end, we observe that

\[
\frac{\partial}{\partial \varepsilon_2} \eta(\varepsilon_1, \varepsilon_2) = \eta(\varepsilon_1, \varepsilon_2) \log \left( \frac{\varepsilon_2 \varepsilon_2}{(1 - \varepsilon_1)(1 - \varepsilon_1 - \varepsilon_2)} \right) \\
\geq \eta(\varepsilon_1, \varepsilon_2) \log(e/2) > 0
\]

for all relevant \(\varepsilon_2 \geq (1 - \varepsilon_1)/2\). Therefore, we need only show that

\[
\zeta(\varepsilon_1) = \frac{\eta(\varepsilon_1, \min\{\varepsilon_1, 1 - \varepsilon_1\})}{\varepsilon_1 e^{1 - \varepsilon_1}} < 1 - \delta
\]

for some \(\delta > 0\) and all relevant \(\varepsilon_1\). We treat the cases \(\varepsilon_1 \in [1/3, 1/2]\) and \(\varepsilon_1 \in [1/2, 1]\) separately.

For \(\varepsilon_1 \in [1/3, 1/2]\), we have

\[
\zeta(\varepsilon_1) = \frac{\eta(\varepsilon_1, \varepsilon_1)}{\varepsilon_1 e^{1 - \varepsilon_1}} = \frac{(e(1 - 2\varepsilon_1))^{1 - 2\varepsilon_1} \varepsilon_1^{4\varepsilon_1 - 1}}{(1 - \varepsilon_1)^{\varepsilon_1}}.
\]

29
Hence
\[
\frac{\partial}{\partial \varepsilon_1} \zeta(\varepsilon_1) = \zeta(\varepsilon_1) \left( \log \left( \frac{\varepsilon_1^4}{(1 - \varepsilon_1)^2} \right) + \frac{\varepsilon_1^2 + \varepsilon_1 - 1}{\varepsilon_1 (1 - \varepsilon_1)} \right).
\]

The terms \(\varepsilon_1^4/(1 - \varepsilon_1)^2\) and \((\varepsilon_1^2 + \varepsilon_1 - 1)/(\varepsilon_1 (1 - \varepsilon_1))\) are increasing for \(\varepsilon_1 \in [1/3, 1/2]\), as is easily verified. Hence \(\zeta(\varepsilon_1)\) is decreasing in \(\varepsilon_1\) for \(1/3 \leq \varepsilon_1 \leq x_1 \approx 0.439\) and increasing for \(x_1 \leq \varepsilon_1 \leq 1/2\). Therefore, since \(\zeta(1/3) = (e/6)^{1/3} < 1\) and \(\zeta(1/2) = 1/\sqrt{2} < 1\), we have that, for some \(\delta_1 > 0\), \(\zeta(\varepsilon_1) < 1 - \delta_1\) for all \(\varepsilon_1 \in [1/3, 1/2]\).

Similarly, for \(\varepsilon \in [1/2, 1)\), we have
\[
\zeta(\varepsilon_1) = \frac{\eta(\varepsilon_1, 1 - \varepsilon_1)}{\varepsilon_1 e^{1 - \varepsilon_1}} = (1 - \varepsilon_1)^{1 - \varepsilon_1} \varepsilon_1^{2 \varepsilon_1 - 1}.
\]

Hence
\[
\frac{\partial}{\partial \varepsilon_1} \zeta(\varepsilon_1) = \zeta(\varepsilon_1) \left( \log \left( \frac{\varepsilon_1^2}{1 - \varepsilon_1} \right) + \frac{\varepsilon_1 - 1}{\varepsilon_1} \right).
\]

Since \(\varepsilon_1^2/(1 - \varepsilon_1)\) and \((\varepsilon_1 - 1)/\varepsilon_1\) are increasing in \(\varepsilon_1 \in [1/2, 1)\), we find that \(\zeta(\varepsilon_1)\) is decreasing in \(\varepsilon_1\) for \(1/2 \leq \varepsilon_1 \leq x_2 \approx 0.692\) and increasing for \(x_2 \leq \varepsilon_1 < 1\). Note that \(\zeta(1/2) = 1/\sqrt{2} < 1\) and \(\zeta(1) = 1\). Hence, for some \(\delta_2 > 0\), \(\zeta(\varepsilon_1) < 1 - \delta_2\) for all \(\varepsilon_1 \in [1/2, \beta_1/\beta] \subset [1/2, 1)\).

Setting \(\delta' = \min\{\delta_1, \delta_2\}\), we find that \(\zeta(\varepsilon_1) < 1 - \delta'\) for all relevant \(\varepsilon_1\). It follows that, for some \(\delta > 0\), we have that \(\nu(\beta, \varepsilon_1, \varepsilon_2) < -\delta\), for all relevant \(\beta, \varepsilon_1, \varepsilon_2\). Summing over the \(O(\log^3 n)\) possibilities for \(k\) and the \(k_i\) and the \(O(1)\) possibilities for \(\ell\) and the \(\ell_i\), we conclude by (5.2) that with high probability \(G_{n,p}\) has no 3-cores \(C\) as in (i).

**Case (ii).** Let \(k, k_j, \ell, m\) be as in (ii). Let \(k_1 = \varepsilon_1 k\) and \(k_2 - m = \varepsilon_2 k\), so that \(\sum \varepsilon_j = 1\). Without loss of generality we assume that \(k_1 \geq k_2\). Hence \(\varepsilon_1, \varepsilon_2\) satisfy \(1/2 \leq \varepsilon_1 \leq \beta_1/\beta < 1\) and \(\varepsilon_2 = 1 - \varepsilon_1\). The number of 3-cores \(C\) as in (ii) for these values \(k, k_j, \ell, m\) is bounded by
\[
\binom{k}{k_2 - m} \binom{k_1}{m} m!^2 \prod_{j=1}^m \sum_{i=0}^m \binom{m}{i} \ell^{2k_1} \frac{\ell^{j}}{k_1^{i}}.
\]

By [Lemma 3.7](#) and the inequality \(k! < ek(k/e)^k\), this is bounded by
\[
\binom{k}{k_2 - m} \frac{k!^m}{m!} \left(2m! k^2 k^2\right)^2 \left(2/\varepsilon^2\right)^{2m} \prod_{j=1}^m k_{2j}^{2k_j}.
\]
By the inequality \binom{n}{k} < (ne/k)^k, and since
\[ k^{2k_2} < (ek)^{2m(k_2 - m)^2}, \]
the above expression is bounded by \((2e^{-2}\eta(\varepsilon_1, 1 - \varepsilon_1))^k k^{2k_2} n^\alpha(1)\), where \(\eta\) is as defined in Case (i). Therefore, by the arguments in Case (i), when \(\varepsilon_1 \geq 1/2\) and \(\varepsilon_2 = 1 - \varepsilon_1\), we find that with high probability \(G_{n,p}\) has no 3-cores \(C\) as in (ii).

The proof is complete. □

References

[1] O. Angel and B. Kolesnik, Minimal contagious sets in random graphs, preprint (2017), available at arXiv:1705.06815.
[2] ______, Thresholds for contagious sets in random graphs, preprint (2016), available at arXiv:1611.10167.
[3] J. Balogh, B. Bollobás, and R. Morris, Graph bootstrap percolation, Random Structures Algorithms 41 (2012), no. 4, 413–440. MR 2993128
[4] B. Bollobás, Weakly \(k\)-saturated graphs, Beiträge zur Graphentheorie (Kolloquium, Manebach, 1967), Teubner, Leipzig, 1968, pp. 25–31. MR 0244077.
[5] J. Chalupa, P. L. Leath, and G. R. Reich, Bootstrap percolation on a bethe lattice, J. Phys. C 21 (1979), L31–L35.
[6] P. Erdős and A. Rényi, On random graphs. I, Publ. Math. Debrecen 6 (1959), 290–297. MR 0120167.
[7] S. Janson, T. Łuczak, T. Turowska, and T. Vallier, Bootstrap percolation on the random graph \(G_{n,p}\), Ann. Appl. Probab. 22 (2012), no. 5, 1989–2047. MR 3025687.
[8] M. Pollak and I. Riess, Application of percolation theory to 2d-3d Heisenberg ferromagnets, Physica Status Solidi (b) 69 (1975), no. 1, K15–K18, (and R. Kopelman, private communication).

Brett Kolesnik
Department of Mathematics, University of British Columbia
Email: bkolesnik@math.ubc.ca