Weakly saturated random graphs

Zsolt Bartha\(^1\) | Brett Kolesnik\(^2\)

\(^1\)HUN-REN Alfréd Rényi Institute of Mathematics, Budapest, Hungary
\(^2\)Department of Statistics, University of Oxford, Oxford, UK

Correspondence
Brett Kolesnik, Department of Statistics, University of Oxford, Oxford, UK.
Email: brett.kolesnik@stats.ox.ac.uk

Abstract
As introduced by Bollobás, a graph \(G\) is weakly \(H\)-saturated if the complete graph \(K_n\) is obtained by iteratively completing copies of \(H\) minus an edge. For all graphs \(H\), we obtain an asymptotic lower bound for the critical threshold \(p_c\), at which point the Erdős–Rényi graph \(G_{n,p}\) is likely to be weakly \(H\)-saturated. We also prove an upper bound for \(p_c\), for all \(H\) which are, in a sense, strictly balanced. In particular, we improve the upper bound by Balogh, Bollobás, and Morris for \(H = K_r\), and we conjecture that this is sharp up to constants.

Keywords
bootstrap percolation, random graph, weak saturation

1 | INTRODUCTION

The concept of weak saturation in graphs was introduced by Bollobás [12]. Given graphs \(G\) and \(H\), the graph \(\langle G \rangle_H\) is obtained by iteratively completing copies of \(H\) minus an edge, starting with \(G\). Formally, set \(G_0 = G\), and for \(t \geq 1\), construct \(G_t\) by adding every edge not in \(G_{t-1}\) which if added to \(G_{t-1}\) creates a new copy of \(H\). We let \(\langle G \rangle_H = \bigcup_t G_t\) denote the result of this procedure. If \(\langle G \rangle_H\) is the complete graph on the vertex set of \(G\), that is, if all missing edges are eventually added, we say that \(G\) is weakly \(H\)-saturated, or that it \(H\)-percolates.

This process can be viewed as a type of cellular automaton [20, 26], of which bootstrap percolation (see, e.g., [1, 5, 13, 15, 17, 22, 25]) is a well-studied example. Balogh, Bollobás and Morris [6] introduced a random process called graph bootstrap percolation, taking \(G\) above to be the Erdős–Rényi [16] graph \(G_{n,p}\). The critical point \(p_c\), at which \(G_{n,p}\) is likely to \(H\)-percolate, is defined formally as

\[ p_c(n, H) = \inf \{ p > 0 : \mathbb{P}(\langle G_{n,p} \rangle_H = K_n) \geq 1/2 \}. \]
The purpose of this work is to obtain general bounds for $p_c$. Our first main result (see Theorem 4) establishes a nontrivial lower bound, which holds for all $H$. This follows by a general extremal result (see Proposition 9) that lower bounds the number of edges in so-called witness graphs, which add a given edge. This extends the bound in [6], proved in the case that $H = K_r$, to all $H$. As an application (see Theorem 2), we locate $p_c$ up to poly-logarithmic factors for all balanced (see Definition 1) graphs $H$, partially answering Prob. 1 in [6]. Our second main result (see Theorem 6) proves a sharper upper bound for all strictly balanced $H$. In particular, this improves the bound in [6] for $K_r$, when $r \geq 5$.

The primary focus in [6] is the case that $H = K_r$ is a complete graph (although some other graphs are also analyzed, see [6, Sec. 5]). Note that all graphs $K_2$-percolate (any missing edge is added at time $t = 1$) so trivially $p_c(n, K_2) = 0$. A graph $K_3$-percolates if and only if it is connected, so it follows

$$P(G_{n,p}=K_3) \rightarrow \exp(-e^{-c})$$

if $p = (\log n + c)/n$ by the fundamental work [16]. The next threshold of interest $p_c(n, K_4)$ is estimated in [6] up to constant factors, and the recent works [3, 4, 19] show that $p_c(n, K_4) \sim 1/\sqrt{3n \log n}$.

For $r \geq 5$, $p_c(n, K_r)$ is estimated in [6] up to poly-logarithmic factors. The upper bound for $p_c(n, K_r)$ proved in [6] holds for a more general class of graphs, which we now recall.

### 1.1 Balanced graphs

For a graph $H$, we let $v_H$ and $e_H$ denote its number of its vertices and edges, respectively, and we put

$$\lambda = (e_H - 2)/(v_H - 2).$$

Note that $H$-percolation is trivial if $v_H \leq 3$ or if the minimum degree $\delta_H = 1$. In the latter case, $p_c$ essentially coincides with the threshold for a copy of $H$ minus an edge in $G_{n,p}$ (see [6, Prop. 26]). That is, $p_c = \Theta(n^{-1/\lambda'})$, where

$$\lambda' = \min_{e \in E[H]} \max_{F \subset H \setminus e} e_F/v_F,$$

where $E[H]$ is the edge set of $H$, and $F$ is a subgraph of $H$ not containing the edge $e$. We therefore assume throughout this work that $\delta_H \geq 2$ and $v_H \geq 4$. In particular, this implies $\lambda \geq 1$. In fact, these assumptions hold for every graph satisfying the next definition.

**Definition 1.** We say that a graph $H$ is balanced if $v_H \geq 4$, and $(e_F - 1)/(v_F - 2) \leq \lambda$ for all subgraphs $F \subset H$ with $3 \leq v_F < v_H$.

This is related to the notion of a 2-balanced graph $G$, such that $(e_F - 1)/(v_F - 2)$ is maximized (over $F \subset G$ with $v_F \geq 3$) when $F = G$. This concept plays a role in, for example, [7, 14, 23, 24], where the maximal number of edges in an $H$-free subgraph (Turán’s problem) of $G_{n,p}$ is studied. Indeed, a graph $H$ is balanced as above if and only if $H \setminus e$ is 2-balanced, for all edges $e \in E[H]$. It also follows that $H$ is connected. See Appendix A.1 for a proof of these basic facts.

In [6], it is shown that $p_c(n, K_r) = n^{-1/\lambda+o(1)}$, as $n \to \infty$. The upper bound holds for balanced graphs $H$ (see [6, Prop. 3]). The lower bound, on the other hand, relies on the so-called witness set algorithm, which assigns to each $e \in E[H]$ a witness graph $W_e \subset G$ such that $e \in E[W_e]$. This algorithm yields an Aizenman–Lebowitz [1] type property (see Lemma 8), a standard tool from the theory of bootstrap percolation. A lower bound for $p_c$ is obtained by the first moment method in [6] using this, together with the fact (Lemma 9 in [6]) that if $H = K_r$ then a witness graph on $k$ vertices has at least $\lambda(k-2)+1$ edges. The proof, however, is somewhat abstract and lengthy. The authors state that “the proof is delicate, and does not seem to extend easily to other graphs.”
In this work, we present a short and simple proof (see Proposition 9 below) that works directly with the $H$-percolation dynamics, and naturally for all graphs $H$. Using this, we obtain the following result, which answers Prob. 1 in [6] in the case that $H$ is balanced.

**Theorem 2.** If $H$ is balanced (see Definition 1) then $p_c(n, H) = n^{-1/\lambda + o(1)}$.

We note that Bayraktar and Chakraborty [9] studied the case that $H = K_{r,s}$ is a complete bipartite graph. They find $p_c$ up to poly-logarithmic factors in the range of $r, s$ where $K_{r,s}$ is balanced, partially answering Prob. 5 in [6]. These results follow by Theorem 2, as a special case.

We also note that Theorem 2 is used in the recent work [8] to locate $p_c$ when $H = G_{\kappa,1/2}$ is a random graph, answering Prob. 6 in [6].

### 1.2 General lower bound

Theorem 2 follows by the upper bound in [6] and a general lower bound for $p_c(n,H)$, that holds for all $H$ (satisfying our baseline assumptions $v_H \geq 4$ and $\delta_H \geq 2$), which we now describe.

**Definition 3.** For a graph $H$ with $v_H \geq 4$, we put

$$\lambda_* = \min_{\gamma \geq 2} \frac{e_H - e_F - 1}{v_H - v_F},$$

minimizing over all subgraphs $F \subset H$ with $2 \leq v_F < v_H$.

Note that $\lambda_* > 0$ if and only if $\delta_H \geq 2$, and we continue to restrict our attention to graphs $H$ with this property.

It is easy to see that $\lambda_* \leq \lambda$, with equality if and only if $H$ is balanced (see Lemma 10). In Section 2.3 below, we show that a witness graph $W_{\gamma}$ for some $\gamma \in E[(G)_H]$ with $k \geq v_H$ vertices has at least $\lambda_* (k - \gamma H) + e_H - 1$ edges. Note that $K_r$ is balanced, and so this reduces to $\lambda (k - 2) + 1$ in that case, recovering Lemma 9 in [6]. A general lower bound follows.

**Theorem 4.** For any graph $H$ with $v_H \geq 4$ and $\delta_H \geq 2$,

$$p_c(n, H) \geq \Omega(n^{-1/\lambda_*} (\log n)^{1/\lambda_* - 1}).$$

Note that, in the case that $H = K_4$, this lower bound includes the correct poly-logarithmic factor (recall that $p_c(n, K_4) \sim 1/\sqrt{3n \log n}$, as discussed above), since $K_4$ is balanced and so $\lambda_* = \lambda = 2$. In [6], the “double-dumbbell” $H = DD_r$ (two copies of $K_r$, $r \geq 4$, joined by a pair of disjoint edges) is given as an example of an (unbalanced) graph for which $p_c = n^{-1/\gamma + o(1)}$, with $\gamma \in (\lambda', \lambda)$ (recall $\lambda'$ defined above Definition 1). We note that, in this instance, $\gamma = \left[ \left( \frac{r}{2} \right) + 1 \right] / r = \lambda_*$. On the other hand, Bidgoli, Mohammadian and Tayfeh-Rezaie [11] have studied the cases $H = K_{2,t}$. For these unbalanced graphs, even finding the correct power $\gamma$ in $p_c = n^{-1/\gamma + o(1)}$ remains open, except in the specific case $H = K_{2,4}$ where $p_c = \Theta(n^{-10/13}).$ However, note that $\lambda_* = 1$ for this graph.

Towards a full solution to Prob. 1 in [6], it would be interesting to determine the class of graphs for which $p_c = n^{-1/\lambda_* + o(1)}$. We note that Theorem 2 shows that this class includes all balanced graphs.

### 1.3 Strictly balanced graphs

Finally, we turn our attention to the specific class of balanced graphs, which includes the cases $H = K_r$, $r \geq 5$, of natural interest.
Definition 5. We call $H$ strictly balanced if the inequality $(e_F - 1)/(v_F - 2) < \lambda$ is strict in Definition 1.

Note that $\lambda > 1$ for strictly balanced graphs. Indeed, recall that we assume that $\delta_H \geq 2$, and consider any $F \subset H$ with three vertices and at least two edges.

For this class of graphs, we prove a sharper upper bound.

Theorem 6. If $H$ is strictly balanced then $p_\gamma(n, H) \leq O(n^{-1/\lambda})$.

Note that $K_3$ is balanced, but not strictly. For all $r \geq 5$, however, the graphs $K_r$ are strictly balanced. It is somewhat tempting to suspect that $p_\gamma(n, H) = \Theta(n^{-1/\lambda})$ for all strictly balanced graphs, but we would not go so far as to make that conjecture.

1.4 | Outline

The general lower bound in Theorem 4 is proved in Section 2, using Proposition 9 below, which lower bounds the number of edges in a witness graph. This result generalizes a result about cliques $K_r$, proved by a different strategy in [6, Lemma 9], to all graphs $H$.

Finally, in Section 3, we prove the upper bound in Theorem 6 for strictly balanced graphs. This is the most technical part of the article, where two rounds of the second moment method are required. First we bound the probability that a given edge is added by a specific type of witness graph, called an “$H$-ladder,” with an appropriately chosen height. Then we show that these events are roughly independent enough to ensure that a large proportion of all edges are added in this way, and then full percolation follows by sprinkling. We note that $H$-ladder graphs were introduced in [6, Sec. 2]. A finer analysis of the ways in which pairs of such graphs can overlap is key to the sharper upper bound in Theorem 6.

1.5 | Notation

For a graph $G = (V, E)$, we denote its vertex set by $V(G) = V$ and its edge set by $E(G) = E$, with sizes $v_G = |V(G)|$ and $e_G = |E(G)|$. We write $F \subset G$ for a (not necessarily induced) subgraph $F$ of $G$. We denote by $G \setminus e$ the graph obtained from $G$ by keeping its vertex set, and deleting the edge $e$ from its edge set. We similarly use $G \setminus \{e_1, \ldots, e_k\}$ when the edges $e_1, \ldots, e_k$ are removed. For graphs $G_1, G_2, \ldots, G_k$, we denote by $\bigcup_{i=1}^k G_i$ the graph with vertex set $\bigcup_{i=1}^k V[G_i]$ and edge set $\bigcup_{i=1}^k E[G_i]$, whereas $\bigcap_{i=1}^k G_i$ is the graph with vertex set $\bigcap_{i=1}^k V[G_i]$ and edge set $\bigcap_{i=1}^k E[G_i]$. For an edge $e = \{x, y\}$, for ease of notation, we often simply write $e$ to denote the graph $G$ with $V(G) = \{x, y\}$ and $E(G) = \{e\}$.

Throughout the paper we use the notation $f = O(g)$ or $f \gtrless O(g)$ for functions $f$ and $g$ if $f \lesssim cg$ for some universal constant $c$ not depending on any of the arguments of $f$ and $g$. Similarly, we use $f = \Omega(g)$ or $f \gtrsim \Omega(g)$ for the opposite inequality, and $f = \Theta(g)$ when both of these statements hold.

2 | A GENERAL LOWER BOUND

In this section, we obtain a lower bound for $p_\gamma$ that holds for all graphs $H$ (with $\delta_H \geq 2$ and $v_H \geq 4$). We first recall, in Sections 2.1 and 2.2, the results from [6] which we use. Then, in Section 2.3, we prove Theorem 4.

2.1 | Witness set algorithm

In [6, Sec. 3.1], the witness set algorithm (WSA) is introduced, which assigns a witness graph $W_e \subset G$ to each $e \in E(G)_H$ such that $e \in E(W_e)_H$. These graphs are defined in time with the percolation
dynamics in the following way. Let \( E_t \) denote the set of edges \( E[G_t] \setminus E[G_{t-1}] \) added at time \( t \). For any edge \( e \in E[G] \), \( W_e \) consists of the single edge \( e \) with its two endpoints as vertices. Then, at step \( t \geq 1 \), the WSA defines simultaneously for each \( e \in E_t \) the witness graph \( W_e = \bigcup_{f \in E[H \setminus e]} W_f \), where \( H_e \) is a copy (chosen arbitrarily if not unique) of \( H \) completed by the addition of \( e \) at time \( t \). Since \( H_e \setminus e \subset G_{t-1} \), this procedure is well-defined.

**Definition 7.** For a witness graph \( W_e \), we define \( v_{W_e} - 2 \), the number of vertices in \( W_e \) besides the endpoints of \( e \), to be its size.

We note that the “size” of a graph sometimes refers to its number of edges. We will not follow this convention.

A key property of this construction is the following Aizenman–Lebowitz [1] type property (cf. Lemma 13 in [6]), as is easily observed.

**Lemma 8.** Suppose that \( W_e \), for some \( e \in E[(G)_{H}] \), is of size at least \( k \) for some \( k \geq v_H - 2 \). Then, for some \( k' \in [k, e_H k] \), there is an \( f \in E[(G)_{H}] \) so that \( W_f \) is of size \( k' \).

**Proof.** Let \( M_1 \) be the maximal size of a witness graph \( W_f \), for \( f \in E[G_t] \). Note that \( M_0 = 0 \) and \( M_1 = v_H - 2 \) (assuming \( E_t \neq \emptyset \)). Then for any \( e \in E_{t+1}, t \geq 1 \), \( W_e \) is of size at most \( v_H - 2 + (e_H - 1)M_1 \leq e_H M_1 \), since \( M_t \geq M_1 = v_H - 2 \). Therefore \( M_{t+1} \leq e_H M_t \). Hence, if \( M_t \geq k \) for some \( t \), then \( M_s \in [k, e_H k] \) for some \( s \leq t \).

### 2.2 Red edge algorithm

In [6] a lower bound for \( p_c \), in the special case of \( H = K_r \), is obtained using Lemma 8 together with a lower bound for the number of edges in a witness graph. Specifically, it is shown that if \( W_e \) is of size \( k \), then it has at least \( \alpha k + 1 \) edges. A key tool in this regard is the red edge algorithm (REA), which is based on WSA (see Section 2.1 above). Informally, for a given edge \( e \in E[(G)_{H}] \setminus E[G] \) (not in \( G \) but eventually added by the \( H \)-percolation dynamics), REA describes the construction of the witness graph \( W_e \) one step at a time by running WSA, but ignoring steps that do not contribute to the construction of \( W_e \). All involved edges which are not in \( G \) are colored red.

REA is discussed in Sec. 3.1 of [6]. For completeness, we briefly describe the construction here. First, we “slow down” the \( H \)-percolation dynamics, so that in each step a single new edge \( e' \) is added. Recall (see Section 2.1) that \( H_{e'} \) is the copy of \( H \) which \( e' \) completes (that is, \( W_{e'} = \bigcup_{f \in E[H \setminus e']} W_f \)). Let \( e_1, \ldots, e_m = e \) be the edges for which \( W_{e_j} \subset W_e \) for \( j \in \{1, \ldots, m\} \) that are added (in that order by the “slowed down” dynamics) until finally \( e \) is added. Color all \( e_1, \ldots, e_m \) red. Then

\[
W_e = (H_1 \cup \ldots \cup H_m) \setminus \{e_1, \ldots, e_m\},
\]

where \( H_j = H_{e_j} \). Hence, REA has \( m \) steps. In the \( j \)th step, a copy \( H_j \) of \( H \) is added and one of its edges \( e_j \) is colored red. Note that \( e_j \notin \bigcup_{i < j} E[H_i] \), however, some of the other edges in \( H_j \) may already be in \( \bigcup_{i < j} E[H_i] \).

### 2.3 Proof of the lower bound

In this section, we give a proof of Theorem 4 that is based on the Aizenman–Lebowitz type property of Lemma 8, and the following lower bound on the number of edges of witness graphs, which we will prove later (as a special case of Lemma 12 below). Recall the definition of \( \lambda_n \) given in Section 1.
**Proposition 9.** If $W_e$ is a witness graph for an edge $e \in E(G)[H]$ on $k \ge v_H$ vertices, then $W_e$ has at least $\lambda_s(k - v_H) + e_H - 1$ edges.

The next lemma shows, in particular, that

$$\lambda_s(k - v_H) + e_H - 1 \ge \lambda_s(k - 2) + 1,$$

with equality if $H$ is balanced.

**Lemma 10.** We have $\lambda_s \le \lambda$, with equality if and only if $H$ is balanced.

**Proof.** The case when $F$ consists of a single edge $e \in E[H]$, with its two endpoints as vertices, shows that $\lambda_s \le \lambda$. To see the second claim, note that, for $F \subset H$ with $3 \le v_F < v_H$,

$$\lambda = \frac{e_H - \lambda(v_F - 2) - 2}{v_H - v_F} \le \frac{e_H - e_F - 1}{v_H - v_F},$$

if and only if $\lambda(v_F - 2) \ge e_F - 1$.

Since $K_r$ is balanced, and so $\lambda_s = \lambda$ in this case, we obtain Lemma 9 in [6] as a special case of Proposition 9. We note here that, as per [21], this result, in the special case of $K_r$, can also be alternatively obtained using [18].

We note here that once Lemma 8 and Proposition 9 have been established, the proof of our general lower bound for $p_e$ follows straightforwardly, along the same lines as the proof of Proposition 8 in Sec. 3.2 of [6].

**Proof of Theorem 4.** Fix $e \in E(K_n)$. Let $p = \alpha n^{-1/\lambda_s} (\log n)^{1/\lambda_s - 1}$. We show that, for $\alpha > 0$ sufficiently small, $e \notin E(\langle G_{n,p} \rangle)[H]$ with high probability.

If $e \in E(\langle G_{n,p} \rangle)[H]$, then by Lemma 8 either (1) $e \in E(\langle G_{n,p} \rangle)$, (2) $W_e$ is of size $k \in [v_H - 2, \log n]$, or else, (3) some $W_f'$ is of size $k' \in (\log n, e_H \log n]$. By Proposition 9, and the remark after it, a witness graph of size $k$ has at least $\lambda_s k + 1$ edges. There are at most $k[O(k^{\lambda_s})]^k$ graphs with exactly $\lambda_s k + 1$ edges on a given set of $k + 2$ vertices (using $\binom{m}{c} \le (me/c)^c$). Therefore, taking a union bound,

$$\mathbb{P}(e \in E(\langle G_{n,p} \rangle)[H])$$

$$\le p + p \sum_{k=v_H-2}^{\log n} k[O(np^{\lambda_s} k^{\lambda_s - 1})]^k + n^2 p \sum_{k'=\log n}^{e_H \log n} k'[O(np^{\lambda_s} (k')^{\lambda_s - 1})]^{k'}$$

$$\le p[O(\log n)]^2 (1 + n^2 + O(\log(n))) \ll 1,$$

for $\alpha$ sufficiently small.

We turn now to the proof of Proposition 9. To this end, it is useful to define an increasing sequence $H_1 \subset \cdots \subset H_n$ of auxiliary graphs associated with REA. (We note that something similar appears in [6], however, using hyper-graphs. For our purposes, it suffices to use graphs.) Recall (see Section 2.2) that in the $j$th step of REA a copy $H_j$ of $H$ is added and one of its new (not in $\bigcup_{i<j} H_i$) edges $e_j$ is colored red. Let $H_0$ be the empty graph. Then the graph $H_j$ is obtained from $H_{j-1}$ by adding a new vertex $v_j$, which we associate with $H_j$. For each edge $e' \in E(H_j) \cap E[\bigcup_{i<j} H_i]$ we add an edge from $v_j$ to
where $k = \max \{ i < j : e' \in H_j \}$. Finally, delete any redundant edges, to ensure that $H_j$ is a simple graph. Note that $v_k$ is associated with the most recently (before step $j$) added copy of $H$ containing $e'$. Finally, we put $H_e = H_m$.

**Definition 11.** For each connected component $C$ of $H_j$, we refer to $C = \bigcup_{i \in V[C]} H_i$ as the corresponding component in $\bigcup_{j=1}^i H_i$.

Components of $\bigcup_{j=1}^i H_i$ can share vertices but not edges. Indeed, when $H_j$ is added in step $j$ of REA, all components in $\bigcup_{j=1}^i H_i$ with at least one edge in $H_j$ are merged with $H_j$ to obtain a component of $\bigcup_{j=1}^i H_i$. Also note that $\bigcup_{j=1}^i H_i$ has only one component (that is, $H_e$ is connected). To see this, simply recall that $W_e = \bigcup_{f \in \{ H_i \}} W_f$, and so note that each $W_f$ has an edge $f \in H_e$, by the construction of $W_e$. Therefore we obtain Proposition 9 by the next result (which also implies Lemma 10 in [6] as a special case, once the terminology there is unpacked).

**Lemma 12.** Let $W_e$ be a witness graph for an edge $e \in E[(G)_H] \setminus E[G]$ on $k \geq v_H$ vertices. Then, after any $j \geq 1$ number of steps of the corresponding instance of REA, any component $C$ of $\bigcup_{j=1}^i H_i$ has at least $\lambda_s(v_C - v_H) + e_H - 1$ non-red edges.

**Definition 13.** We let $\mathcal{V}_s \subset \{ 2, \ldots, v_H - 1 \}$ be the set for which $\lambda_s$ is attained by some subgraphs $F \subset H$ with $v_F \in \mathcal{V}_s$. We put

$$\xi = \min_{F \subset H: v_F \in \{ 2, \ldots, v_H - 1 \}} \frac{e_H - e_F - 1}{v_H - v_F} - \lambda_s,$$

minimizing over $F \subset H$ with $v_F \in \{ 2, \ldots, v_H - 1 \} \setminus \mathcal{V}_s$.

Roughly speaking, we shall see that the most efficient way to add a new copy $H_j$ of $H$, in any given step $j$ of REA, is to ensure that the number of vertices in $V(H_j) \cap V(C')$ is in $\mathcal{V}_s$, for each of the components $C'$ in $\bigcup_{j=1}^i H_i$ that have an edge in common with $H_j$ (except in the simple case $H_j \setminus e_j \subset C'$, when only the red edge $e_j$ is added to $C'$ to obtain $C$). Note that all such components are merged with $H_j$ in the $j$th step of REA to form a new component $C$. The quantity $\xi$ is related to the minimum “cost” of a nonoptimal merge in REA.

We will also use the following distinction of steps in REA.

**Definition 14.** We call the $j$th step in REA a type-$1$ step if $H_j \setminus e_j$ is contained in some component of $\bigcup_{j=1}^{i-1} H_{i-1}$ as a subgraph. Otherwise, we call it a type-$2$ step. In particular, we include steps in which a new component is formed as a trivial instance of a type-$2$ step.

**Proof of Lemma 12.** The proof is by induction on the number of steps $j$ taken. The base case $j = 1$ is trivial, since $H_1 \setminus e_1$ has $v_H$ vertices and $e_H - 1$ edges. Likewise, the same reasoning applies if in some step $j > 1$ a new component is created. Hence suppose that in step $j > 1$, the addition of $H_j$ causes exactly $h \geq 1$ (edge-disjoint) components $C_1, \ldots, C_h$ (each with at least one edge in $H_j$) to merge with $H_j$ into a single component $C$. By assumption, we assume that the $C_i$ have $k_i$ vertices and $\lambda_s(k_i - v_H) + e_H - 1 + \ell_i$ non-red edges, for some $\ell_i \geq 0$. Let $k$ denote the number of vertices in $C$. Note that

$$k = v_H + \sum_i (k_i - \varepsilon_i - \eta_i),$$

where $\varepsilon_i$ is the number of vertices in $V[C_i] \setminus V[H_j]$, and $\eta_i$ is the number of vertices in $(V[C_i] \setminus V[H_j]) \setminus \bigcup_{j \neq i} V[C_j]$ (i.e., other vertices in $C_i$ that are in a previously considered
\(C_r\). To complete the proof, we show that \(C\) has at least \(\lambda_s(k - v_H) + e_H - 1\) nonred edges. We distinguish two scenarios based on whether the \(j\)th step is a type-1 or type-2 step.

**Case 1.** If \(H_j \setminus e_j\) is contained in some component of \(\bigcup_{i=1}^{j-1} H_i\) as a subgraph, then necessarily \(h = 1\), since the components \(C_i\) are edge-disjoint. In this case, the result follows immediately, since then \(k = k_1\) and a single red edge (and no nonred edges) is added to form \(C\).

**Case 2.** On the other hand, suppose that no \(C_i\) contains \(H_j \setminus e_j\) as a subgraph. It is more convenient to start with \(H_j\) and color \(e_j\) red, and then merge the \(C_i\) with it one at a time. In these dynamics, any edge in \(C_i \cap H_j\) that is red in \(C_i\) remains red after merging. Initially, we have the \(e_H - 1\) nonred edges in \(H_j\). In the \(i\)th substep (when we merge \(C_i\)), \(k_i - \varepsilon_i - \eta_i\) vertices are added. Note that, by the choice of \(\lambda_s\) and \(\xi\), at least

\[
(v_H - \varepsilon_i)(\lambda_s + \xi 1_{e \not\in V_s}) + 1_{e = v_H},
\]

of the \(e_H - 1\) edges in \(H_j \setminus e_j\) are not in \(C_i\). Therefore, since the components \(\{C_j\}_{i \leq i}\) are edge-disjoint, the number of non-red edges increases by at least

\[
\lambda_s(k_i - \varepsilon_i) + \ell_i + (v_H - \varepsilon_i)\xi 1_{e \not\in V_s} + 1_{e = v_H},
\]

in the \(i\)th substep. Altogether, summing over all \(i\), we find that \(C\) has \(k\) vertices and at least

\[
(e_H - 1) + \lambda_s \sum_i (k_i - \varepsilon_i) + \sum_i [\ell_i + (v_H - \varepsilon_i)\xi 1_{e \not\in V_s} + 1_{e = v_H}],
\]

nonred edges. Finally, note that the above is equal to

\[
(e_H - 1) + \lambda_s(k - v_H) + \sum_i \eta_i + \sum_i [\ell_i + (v_H - \varepsilon_i)\xi 1_{e \not\in V_s} + 1_{e = v_H}]
\]

\[
= [\lambda_s(k - v_H) + e_H - 1] + \sum_i [\ell_i + \lambda_s \eta_i + (v_H - \varepsilon_i)\xi 1_{e \not\in V_s} + 1_{e = v_H}]
\]

\[
\geq \lambda_s(k - v_H) + e_H - 1,
\]

as required.

### 3 Upper Bound for Strictly Balanced \(H\)

Next, we prove the upper bound in Theorem 6. We show that, for a strictly balanced graph \(H\), with high probability \(\langle G_{n,p} \rangle_H = K_n\), if \(p = (\alpha / n)^{1/\lambda}\) and \(\alpha > 0\) is sufficiently large.

Two applications of the second moment method are involved. First we show that with probability bounded away from 0 (and tending to 1 as \(\alpha \to \infty\)) any given edge \(e \in E[K_n]\) is added in \(\langle G_{n,p} \rangle_H\) due to a simple type of witness graph, called an \(H\)-ladder. These graphs were considered in [6]. The main difference here is that we consider induced \(H\)-ladders, resulting in an easier analysis of correlations (overlapping ladders). Then we show that the events that two given edges are added in \(\langle G_{n,p} \rangle_H\) by induced \(H\)-ladders (of suitable heights) are roughly independent. Hence, a significant proportion (tending to 1 as \(\alpha \to \infty\)) of all \(\binom{n}{2}\) edges in \(K_n\) are included in \(\langle G_{n,p} \rangle_H\). Full percolation is then easily deduced (by Turán’s Theorem and sprinkling).
3.1 H-ladders

We consider (as in Sec. 2 of [6]) the following type of edge-minimal witness graph, where the associated graph (as in the discussion before Definition 11) is a path.

**Definition 15.** An H-ladder $L$ of height $h$ (see Figure 1) is a graph constructed using $h$ copies $S_i$ (called steps) of $H$ minus two nonincident edges (called rungs) $\{u_{i-1}, v_{i-1}\}$ and $\{u_i, v_i\}$ such that, for each $1 < i \leq h$, we have that $V(S_i) \cap \bigcup_{j \neq i} V(S_j) = \{u_{i-1}, v_{i-1}\}$. We then obtain $L$ as the union of the $S_i$ and the top rung $\{u_h, v_h\}$. We call $(v_H - 2)h$ the size, $h$ the height and $\{u_0, v_0\}$ the base of $L$. We often write $k = (v_H - 2)h$.

By induction on $h$, it is easy to see that $L$ is a witness graph for its base edge. Note that $L$ has $\lambda k + 1$ edges, the minimal possible number by Proposition 9. To see this note that each $S_i$ has $e_H - 2 = \lambda(v_H - 2)$ edges. Since only the top rung $\{u_h, v_h\}$ is included in $L$, it has only $\lambda(v_H - 2)h + 1 = \lambda k + 1$ edges in total.

It is convenient, although slightly informal, to speak of vertices and edges of a ladder $L$ that are “above” and “below” its various rungs, etc. In this sense, note that $\lambda$ is the average number of edges “sent down the ladder” by vertices “above” the base.

Let us note here that we will, beginning in the next section, restrict to a specific class of $H$-ladders (see Definition 18 below) that is simpler and suffices for our purposes. However, before doing so, we first establish the following result that holds for $H$-ladders in general.
In [6] (see Lemma 6) it is shown that any subgraph $X \subset L$ containing $x + 2 < k + 2$ vertices of $L$, including those in its base, has at most $\lambda x$ edges. Equality is obtained if $X = \bigcup_{i \leq h'} S_i$ for some $1 \leq h' < h$. We prove the following estimate, which bounds the inefficiency of edge sharing in the other cases.

Since $H$ is strictly balanced (see Definition 13),

$$\xi = \min \frac{e_H - e_F - 1}{v_H - v_F} - \lambda > 0,$$

minimizing over $F \subset H$ with $3 \leq v_F < v_H$. The case of $F$ with $v_F = v_H - 1$ gives the bound

$$\xi \leq \delta_H - 1 - \lambda \leq \delta_H - 2$$

(recall that $\delta_H \geq 2$, and so $\lambda \geq 1$).

**Lemma 16.** Let $L$ be an $H$-ladder of size $k = (v_H - 2)h$. Let $X$ be a proper induced subgraph of $L$ that contains $x$ vertices above the base of $L$. Then $X$ has at most $\lambda x - \xi \sigma$ edges, where $\sigma$ is the number of steps $S_i \notin X$ of $L$ such that $X$ contains at least one vertex in $V[S_i] \setminus \{u_{i-1}, v_{i-1}\}$.

Note that $\sigma = 0$ if and only if $x = 0$ or $X = \bigcup_{i \leq h'} S_i$ for some $1 \leq h' < h$. This result, in particular, implies Lemma 6 in [6] (without the condition $\lambda \geq 2$). Also note that we do not require that $X$ contains the base vertices of $L$. This allows for an easier inductive proof, and will be useful for analyzing overlapping ladders with different bases (Lemma 21 below).

**Proof.** The proof is by induction on the height $h$ of $L$. If $x = 0$ (and so also $\sigma = 0$) the statement is trivial. Thus we assume $x \geq 1$.

**Base case.** If $h = 1$ then $1 \leq x \leq v_H - 2$ and $\sigma = 1$ (since $X$ is proper). There are $\lambda(v_H - 2) + 1$ edges in $L$.

**Case 1a.** If $x = v_H - 2$, then at least one vertex in the base of $L$ is not in $X$. Hence there are at least $\delta_H - 1 \geq \xi + 1$ edges in $E[L] \setminus E[X]$, and so at most $\lambda x - \xi$ in $E[X]$.

**Case 1b.** If $1 \leq x < v_H - 2$, then there are at least $(\lambda + \xi)(v_H - 2 - x) + 1$ edges in $E[L] \setminus E[X]$, and so at most $\lambda x - \xi(v_H - 2 - x) \leq \lambda x - \xi$ in $E[X]$.

**Inductive step.** Suppose $h > 1$. Let $L' \subset L$ be the ladder of height $h - 1$ based at the first rung $\{u_1, v_1\}$ of $L$.

**Case 2.** If $S_1 \subset X$ or $V[X \cap S_1] \subset \{u_0, v_0\}$, then the result follows immediately by the inductive hypothesis applied to $L'$ (since in either case $X \cap L' \subset L'$ is proper).

**Case 3.** Suppose that $X$ contains $x_1 \geq 1$ vertices in $S_1 \setminus \{u_0, v_0\}$ and $S_1 \notin X$. Then, by the base case, $X$ contains at most $\lambda x_1 - \xi - 1_{u_1, v_1 \in V[X]}$ edges in $S_1$ (since $h > 1$, the edge $\{u_1, v_1\} \notin E[L]$).

**Case 3a.** If $X \cap L' = L'$ (in which case $\sigma = 1$, and $u_1, v_1 \in V[X]$) the claim follows, since then there are $\lambda x_1 - \xi - 1$ edges in $X$ below the first rung, $\lambda(k - x_1) + 1$ above, and so $\lambda x - \xi$ in total.

**Case 3b.** Otherwise, applying the inductive hypothesis to the remaining $x - x_1$ vertices of $X$ in $L'$, it follows that there are at most $\lambda(x - x_1) - \xi(\sigma - 1)$ edges in $X \cap L'$. Hence $L$ has at most $\lambda x - \xi \sigma$ edges.

### 3.2 H-Ladders in $\mathcal{G}_{n,p}$

Having established Lemma 16, we turn to the upper bound for $p_c$. We first obtain a lower bound on the probability that a given edge $e \in E[K_n]$ is the base of an $H$-ladder of height $h$ in $\mathcal{G}_{m,p}$. This gives a
lower bound on the probability that $e \in E[\langle G_{n,p} \rangle]$. We then verify the approximate independence for different bases. This strategy thus involves two applications of the second moment method. As already discussed, we restrict to the case of induced $H$-ladders, since this simplifies the analysis of correlations. Furthermore, we also restrict our attention to a specific type of $H$-ladder, defined as follows.

**Definition 17.** Fix two nonincident edges $e_t, e_b$ in $H$, and a copy $T$ of $H \setminus \{e_t, e_b\}$ labeled in some arbitrary way such that
- the vertices in $e_b$ are labelled by $\{1, 2\}$, and
- all other vertices (not in $e_b$) in $H$ are labeled by $\{3, 4 \ldots, v_H\}$.

We call $T$ the template.

We fix $T$ for the rest of this work. We use the template $T$ to define a simple class of $H$-ladders, where, informally speaking, copies of $T$ are stacked on top of each other.

**Definition 18.** A (labelled) $H$-ladder (see Definition 15) is uniform if, for each of its steps $S_i$, the function $\phi_i$ for which
- $\phi_i(u_i-1) = 1$,$\phi_i(v_{i-1}) = 2$, and
- $\phi_i(w_k) = k + 2$, for $w_k \in V[S_i] \setminus \{u_{i-1}, v_{i-1}\}$ of $k$th largest label,

is an isomorphism from $S_i$ to the template $T$.

Note that, since $T$ is fixed, there are exactly

$$\binom{k}{v_H - 2, \ldots, v_H - 2} = \frac{k!}{(v_H - 2)! h}$$

uniform $H$-ladders of size $k = (v_H - 2)h$ (and height $h$) on a given set of $k + 2$ vertices and with a given base $e$. Indeed, the conditions in Definition 18 imply that the only freedom in selecting such an $H$-ladder is in choosing which vertices are in each of the $h$ sets $V[S_i] \setminus \{u_{i-1}, v_{i-1}\}$.

**Definition 19.** For $\varepsilon \in (0, 1)$, put

$$\alpha_{\varepsilon} = \exp \left[\left(\frac{v_H + 1}{\varepsilon} + \frac{1 - \varepsilon}{4}\right) \log \left(\frac{v_H - 2}{\varepsilon}\right)\right], \quad (3.1)$$

and

$$\beta_{\varepsilon} = \frac{\varepsilon \xi}{\lambda(v_H + 1)(v_H - 2) \log(v_H - 2)}. \quad (3.2)$$

Note that $\alpha_{\varepsilon} \to \infty$ and $\beta_{\varepsilon} \to 0$, as $\varepsilon \to 0$. We also note that $\alpha_{\varepsilon}$, $\beta_{\varepsilon}$ are chosen so that (3.4), (3.2), and (3.8) below simplify nicely (but are, of course, not the only choice of $\alpha$, $\beta$ for which the following lemmas hold).

First we show that a given edge in $K_n$ is the base of an $H$-ladder with probability tending to 1 as $np^1 \to \infty$.

**Lemma 20.** Fix $\varepsilon \in (0, 1)$. Put $np^1 = \alpha_{\varepsilon}(v_H - 2)!^{1/(v_H - 2)}$. Then any given $e \in E[K_n]$ is the base of an induced uniform $H$-ladder of height $h = \beta_{\varepsilon} \log n$ in $G_{n,p}$ with probability at least $\gamma_{\varepsilon} - o(1)$, where

$$\gamma_{\varepsilon} = 1 - \frac{1}{\alpha_{\varepsilon}^{v_H - 2} - 1}.$$
For ease of exposition, we write quantities such as $h = \beta_x \log n$ as is, instead of replacing them with their integer parts.

**Proof.** Let $N_k$ denote the number of induced uniform $H$-ladders in $G_{n,p}$ of size $k = (v_H - 2)h$ with a given base $e$. Noting that $k \ll n$ and $k^2 p \ll 1$ (and using the standard bound $\binom{n}{k} \geq (n - k)^k/k!$) we find that

$$
\mathbb{E}N_k = \binom{n - 2}{k} \frac{k!}{(v_H - 2)!} p^{jk+1} (1 - p)^{\left(\frac{k + 2}{2}\right) - (jk + 1)} \geq p \alpha_x^k (1 - o(1)).
$$

(3.3)

Since, by (3.1) and (3.2),

$$
\lambda \beta_x (v_H - 2) \log \alpha_x = 1 + \frac{(1 - \varepsilon)k}{4(v_H + 1)} > 1,
$$

(3.4)

it follows that $\mathbb{E}N_k \gg 1$.

Let $L_1, L_2, \ldots$ enumerate all uniform $H$-ladders in $K_n$ of size $k$ that are based at $e$. Let $A_i$ be the event that $L_i$ is an induced subgraph of $G_{n,p}$. Following Sec. 4.3 of [2], and using symmetry,

$$
\mathbb{P}(N_k = 0) \leq \frac{\mathbb{E}(N_k^2)}{(\mathbb{E}N_k)^2} - 1 = \frac{1}{\mathbb{E}N_k} + \frac{1}{(\mathbb{E}N_k)^2} \sum_{i \neq j} \mathbb{P}(A_i \cap A_j) - 1
$$

$$
= \frac{1}{\mathbb{E}N_k} \left[ 1 + \sum_{i > 1} \mathbb{P}(A_i|A_1) \right] - 1.
$$

(3.5)

Since we are considering induced subgraphs of $G_{n,p}$, for any $i \neq 1$, we have $\mathbb{P}(A_i|A_1) = 0$ unless $V[L_i] \neq V[L_1]$. Let $S_i$, $1 \leq i \leq h$, denote the steps of $L_1$.

**Case 1.** First, we consider the case that $L_i$ “breaks cleanly” from $L_1$ at one of its rungs, that is, $L_i \cap L_1 = e$ or $L_i \cap L_1 = \bigcup_{1 \leq h'} S_i$, for some $1 \leq h' < h$.

If $L_i \cap L_1 = e$ then $L_i$ is an $H$-ladder of height $h$ that is edge-disjoint from $L_1$. Hence $\mathbb{P}(A_i|A_1) \leq p^{jk+1}$. Similarly, if $L_i \cap L_1 = \bigcup_{1 \leq h'} S_i$, for some $1 \leq h' < h$, then $L_i$ and $L_1$ agree up to height $h'$. The part of $L_i$ that is “above” the intersection $L_i \cap L_1$ is an $H$-ladder (based at the $h'$th rung $\{u_{h'}, v_{h'}\}$ of $L_1$) of height $h - h'$ that is edge-disjoint from $L_1$. Hence, in this case, $\mathbb{P}(A_i|A_1) \leq p^{jk'h'}$, where $k' = (v_H - 2)(h - h')$. Summing over all such $L_i$, using (3.3),

$$
\frac{1}{\mathbb{E}N_k} \sum \mathbb{P}(A_i|A_1) \leq \frac{1}{\mathbb{E}N_k} \sum_{h' = 0}^{h-1} n^{k'} p^{jk'h'} (v_H - 2)^{h-h'}
$$

$$
\leq (1 + o(1)) \sum_{h' = 0}^{h-1} \alpha_x^{-(v_H - 2)h'}
$$

$$
\leq \frac{1 + o(1)}{1 - 1/\alpha_x^{v_H - 2}}.
$$

**Case 2.** Next, we show that all other cases are of lower order. If $L_i$ does not “break cleanly” (as in Case 1) from $L_1$ then by Lemma 16, $\mathbb{P}(A_i|A_1) \leq p^{\beta(k-x)+1+\varepsilon \sigma}$, where $x$ is the
number of vertices in \(X = L_1 \cap L_1\) above the base of \(L_1\), and \(\sigma \geq 1\) is the number of \(S_i \not\subset X\) such that \(V[X] \cap (V[S_i] \setminus \{u_{i-1}, v_{i-1}\}) \neq \emptyset\).

For any such \(L_i\), let \(s \geq 0\) be the number of maximal subgraphs \(\bigcup_{i=h_i}^{h_2} S_i \subset X\), \(h_1 \leq h_2\). For convenience, we refer to these as the subladders that \(L_1\) and \(L_i\) have in common. Let \(y \geq 0\) denote the number of other vertices in \(X\) (not inside a common subladder). Note that \(s + y \geq 1\) since \(\sigma \geq 1\). We claim that

\[
\sigma \geq \max \{1, (s - 1 + y)/v_H\} \geq (s + y)/(v_H + 1). \tag{3.6}
\]

To see this, note that there are at least \(2(s - 1)\) vertices of \(X\) in steps \(S_i \not\subset X\), since if \(\bigcup_{i=h_i}^{h_2} S_i \subset X\) is maximal and \(h_1 > 1\), then \(S_{h_i-1} \not\subset X\) and \(u_{h_i-1}, v_{h_i-1} \in V[X]\).

Next, we claim that, for given \(x, s, y\), there are at most

\[
\binom{h + 1}{2s} s! \binom{h}{s} 2^s \binom{n}{k} (k - x + y)! \leq \left[O(k^3)\right]^{s+y} n^{k-x},
\]

ladders \(L_i\) to consider. To see this, observe that:

1. there are \(\binom{h + 1}{2s}\) ways to select the \(s\) common subladders (since this corresponds to selecting \(2s\) rungs in \(L_1\)) and \(s!\) ways to choose the order in which they can appear in \(L_i\),
2. there are at most \(\binom{h}{s}\) possibilities for where the sub-ladders are located in \(L_i\) (choose a height for the top rung of each), and \(2^s\) ways to decided whether the top rung (as it appears in \(L_1\)) of each sub-ladder is the top or bottom rung as it appears in \(L_i\), and
3. the final three factors bound the choices for the \(k - x + y\) other vertices in \(L_i\) and their locations in \(L_i\).

Note that in (2) we need not also consider the possibility that the labels in a top or bottom rung of a common subladder are reversed (with respect to \(e\)) in \(L_i\) as compared with how they appear in \(L_1\). Indeed, doing so would either produce a nonuniform \(H\)-ladder, or else one that is equivalent to \(L_1\).

Hence, summing over all such \(L_i\) with given \(x, s, y\), we find (using \(\alpha_\epsilon > 1, x \geq 1, s + y \geq 1, (3.3)\) and (3.6)) that

\[
\frac{1}{\mathbb{E}N_k} \sum \mathbb{P}(A_i|A_1) \leq \frac{1}{\mathbb{E}N_k} \left[O(k^3)\right]^{s+y} n^{k-x} p^{2(k-x)+1+\xi_\sigma} \\
\leq O\left[(\log n)^3 (np^{k-x})^{1+\xi/(v_H+1)}\right] \\
\leq O\left[(\log n)^3 (v_H - 2)! h_p^{\xi/(v_H+1)}\right] \ll n^{-\theta_\epsilon},
\]

where, by (3.2),

\[
\theta_\epsilon = \frac{\xi}{\lambda(v_H + 1)} - \beta_\epsilon (v_H - 2) \log(v_H - 2) = \frac{(1 - \epsilon)\xi}{\lambda(v_H + 1)} > 0. \tag{3.7}
\]

Since there are only \(O(k^3)\) relevant \(x, s, y\) the same holds summing over all \(L_i\) (not included in Case 1).
Therefore, combining the two cases, we find that
\[
\frac{1}{\mathbb{E}N_k} \sum_{i \geq 1} \mathbb{P}(A_i | A_1) \leq \frac{1 + o(1)}{1 - 1/\alpha_k^{v_H-2}},
\]
and so, by (3.5),
\[
\mathbb{P}(N_k > 0) \geq 1 - \frac{1 + o(1)}{\alpha_k^{v_H-2} - 1} = \gamma_\varepsilon - o(1).
\]

Next, by another application of the second moment method, we show that with high probability a significant proportion (tending to 1 as \( \alpha \to \infty \)) of edges in \( K_n \) are bases of \( H \)-ladders in \( \mathcal{G}_{n,p} \), and so included in \( \langle \mathcal{G}_{n,p} \rangle_H \).

**Lemma 21.** Fix \( \varepsilon \in (0,1) \). Put \( np^3 = \alpha_\varepsilon (v_H - 2)!/(v_H-2) \). Then, with high probability, there are at least \( (\gamma_\varepsilon - \varepsilon) \binom{n}{2} \) edges in \( K_n \) which are bases of induced uniform \( H \)-ladders of height \( h = \beta_\varepsilon \log n \) in \( \mathcal{G}_{n,p} \).

**Proof.** Let \( e_1, e_2, \ldots \) enumerate the edges of \( K_n \) and let \( E_i \) denote the event that \( e_i \) is the base of an induced uniform \( H \)-ladder of height \( h \) in \( \mathcal{G}_{n,p} \). Similarly, let \( E'_i \) denote the event that \( e_i \) is the base of a (not necessarily induced) uniform \( H \)-ladder of height \( h \) in \( \mathcal{G}_{n,p} \). We show that
\[
\frac{1}{\binom{n}{2}} \sum_{i \neq j} \mathbb{P}(E_i \cap E_j) \leq 1 + o(1),
\]
from which, together with Lemma 20, the result follows (as then the number of such edges \( \sum_i 1_{E_i} \), divided by its expectation \( \binom{n}{2} \mathbb{P}(E_1) \), converges to 1 in probability, see again Sec. 4.3 of [2] and the technique used in (3.5)).

To this end, we bound the event \( E_i \cap E_j \) by the union of events (1) \( E_i \circ E'_j \) that there are edge-disjoint (not necessarily induced) uniform \( H \)-ladders of heights \( h \) based at \( e_i \) and \( e_j \), and, (2) \( E_{ij} \) that there is an induced uniform \( H \)-ladder of height \( h \) based at \( e_j \) that includes an edge of such a ladder based at \( e_i \).

By the van den Berg–Kesten (BK) inequality [10] (the events \( E'_j \) are increasing) and symmetry,
\[
\mathbb{P}(E'_i \circ E'_j) \leq \mathbb{P}(E'_j)^2 = (1 + o(1))\mathbb{P}(E_j)^2,
\]
where the last equality follows since \( k^2 p \ll 1 \). It thus suffices to show that
\[
\sum_{i \neq j} \mathbb{P}(E_{ij}) \ll n^4,
\]
since by Lemma 20, the probability \( \mathbb{P}(E_i) \geq \gamma_\varepsilon - o(1) \) (and so, in particular, bounded away from 0 as \( n \to \infty \)).

Let \( L_1 \) be a fixed uniform \( H \)-ladder of height \( h \) in \( K_n \) based at \( e_1 \), and let \( A_1 \) denote the event that \( L_1 \) is an induced subgraph of \( \mathcal{G}_{n,p} \). For \( j > 1 \), let \( B_j \) be the event that there is an induced uniform \( H \)-ladder in \( \mathcal{G}_{n,p} \) of height \( h \) based at \( e_j \) that includes at least one edge in \( L_1 \). As in the previous proof, let \( N_k \) denote the number of induced uniform \( H \)-ladders in
$G_{n,p}$ of size $k = (v_H - 2)h$ with a given base $e$. Note that $\mathbb{E}(N_k) \leq p\alpha_k^e$. Hence, by symmetry, we have
\[
\sum_{i \neq j} \mathbb{P}(E_{ij}) \leq n^2 \sum_{j > 1} \mathbb{P}(E_{1j}) \\
\leq n^2 \mathbb{E}(N_k) \sum_{j > 1} \mathbb{P}(B_j|A_1) \\
\leq n^2 p\alpha_k^e \sum_{j > 1} \mathbb{P}(B_j|A_1).
\]

Hence, it suffices to show that
\[
p\alpha_k^e \sum_{j > 1} \mathbb{P}(B_j|A_1) \ll n^2.
\]

Finally, we estimate $\sum_{j > 1} \mathbb{P}(B_j|A_1)$ by a union bound, considering the expected (conditioned on $A_1$) number of induced $H$-ladders $L$ of height $h$ based at some $e \neq e_1$ that include at least one edge of $L_1$, and hence at least one vertex not in its base $e_1$.

At this point, the argument is similar to the proof of Lemma 20, and so we only sketch the details. As before, we take two cases with respect to whether $L$ and $L_1$ “intersect cleanly” (i.e., if $L \cap L_1 = \bigcup_{i \leq h'} S_i$ for some $1 \leq h' < h$, where $S_i$ are the steps of $L_1$) or not. Note that if $L$ and $L_1$ share at least one common edge, then we cannot have that $L \cap L_1 = e_1$, since if $A_1$ occurs then $e \notin E[G_{n,p}]$.

**Case 1.** If $L$ and $L_1$ “intersect cleanly” then there are $O(h)$ possibilities for where (i.e., the height at which) $L \cap L_1$ is located in $L$. Apart from this, by an argument similar to that in Case 1 in proof of Lemma 20, we see that the expected number of such $L$ is at most $\left(\frac{n}{2}\right) O(hp\alpha_k^e/n^2) \leq O(kp\alpha_k^e)$. The compensating factor $1/n^2$ here is due to the fact that there are $(v_H - 2)h' + 2$ vertices, but only $\lambda(v_H - 2)h'$ edges, in $L \cap L_1$.

**Case 2.** Otherwise, if $L$ and $L_1$ do not “intersect cleanly” then, arguing as in Case 2 in the proof of Lemma 20, the expected number of such $L$ in this case is $\ll \left(\frac{n}{2}\right)n^{-\theta_r}p\alpha_k^e$, where $\theta_r > 0$ is as defined in (3.7).

Altogether,
\[
\frac{p\alpha_k^e}{n^2} \sum_{j > 1} \mathbb{P}(B_j|A_1) \leq O((p\alpha_k^e)^2 (kn^{-2} + n^{-\theta_r})) = O((p\alpha_k^e)^2 n^{-\theta_r}) \ll 1,
\]

since, by (3.4) and (3.7),
\[
\frac{2}{\lambda}(-1 + \lambda\beta_e(v_H - 2) \log \alpha_e) - \theta_r = -\frac{\theta_r}{2} < 0. \tag{3.8}
\]

### 3.3 The upper bound

With Lemma 21 at hand, we obtain our upper bound for $p_c$ by an adaptation of the argument found at the end of Sec. 2 in [6].

**Proof of Theorem 6.** Let $p = (\alpha/n)^{1/\lambda}$. We show that for $\alpha > 0$ sufficiently large, $\langle G_{n,p}\rangle_H = K_n$ with high probability.
Let $G = (V, E)$ be a graph. If only $\epsilon n$ vertices $v \in V$ have degree $d_v \geq \delta(n - 1)$, for some $\delta$, then $|E| \leq (\epsilon + (1 - \epsilon)\delta)n^2/2$. Hence, if $|E| > \gamma(n^2/2)$, there is a set $S \subset V$ of size satisfying $|S|/n \geq (\gamma - \delta)/(1 - \delta)$ so that all $v \in S$ have $d_v \geq \delta(n - 1)$.

Therefore, by Lemma 21, for $\alpha > 0$ large (and so $\gamma$ close to 1) with high probability there is a set $S$ of size $\Omega(n)$ such that all neighborhoods $N_v$ in $(G_{n,p})_H$ of vertices $v \in S$ are larger than $(3/4)n$. As a result, all $u, v \in S$ have $|N_u \cap N_v| \geq n/2$. Also, for $\alpha$ large enough, all induced subgraphs of $(G_{n,p})_H$ of size $n/4$ contain a copy of $K_vH$ by Turán’s Theorem. Hence all edges between vertices in $S$ are in $(G_{n,p})_H$.

Once a percolating subgraph $S$ of size $\Omega(n)$ has been established, the result follows easily by sprinkling, as in [6]. For completeness, we sketch the argument. Consider a random graph $G_{n,p'}$ that is independent of $G_{n,p}$ with $(\log n)/n \ll p' \ll p$. Since $H$ is strictly balanced, such a $p'$ exists, as $p = \Omega(n^{-1/\lambda})$ for some $\lambda > 1$ (see below Definition 5).

Due to $(\log n)/n \ll p'$, with high probability, in the graph $G_{n,p'}$, all vertices outside of $S$ have at least $v_H - 2$ neighbors in $S$. Hence, $(G_{n,p} \cup G_{n,p'})_H = K_n$ with high probability. This implies the result, noting that $G_{n,p} \cup G_{n,p'}$ is a random graph with edge probability $1 - (1 - p)(1 - p') \sim p$. ■

ACKNOWLEDGMENTS
We are indebted to Yuval Peled for noticing an issue in a previous version of this work that had claimed a lower bound (matching the upper bound in Theorem 6) for all strictly balanced graphs.

DATA AVAILABILITY STATEMENT
Data sharing is not applicable, as no datasets were generated or analyzed.

ORCID
Zsolt Bartha https://orcid.org/0000-0002-1986-6571
Brett Kolesnik https://orcid.org/0000-0003-2604-6093

REFERENCES
1. M. Aizenman and J. L. Lebowitz, Metastability effects in bootstrap percolation, J. Phys. A 21 (1988), no. 19, 3801–3813.
2. N. Alon and J. H. Spencer, “The probabilistic method,” Wiley series in discrete mathematics and optimization, fourth ed., John Wiley & Sons, Inc., Hoboken, NJ, 2016.
3. O. Angel and B. Kolesnik, Sharp thresholds for contagious sets in random graphs, Ann. Appl. Probab. 28 (2018), no. 2, 1052–1098.
4. O. Angel and B. Kolesnik, Large deviations for subcritical bootstrap percolation on the Erdős–Rényi graph, J. Stat. Phys. 185 (2021), no. 2, 8–16.
5. J. Balogh, B. Bollobás, H. Duminil-Copin, and R. Morris, The sharp threshold for bootstrap percolation in all dimensions, Trans. Am. Math. Soc. 364 (2012), no. 5, 2667–2701.
6. J. Balogh, B. Bollobás, and R. Morris, Graph bootstrap percolation, Random Struct. Algorithm. 41 (2012), no. 4, 413–440.
7. J. Balogh, R. Morris, and W. Samotij, Independent sets in hypergraphs, J. Am. Math. Soc. 28 (2015), no. 3, 669–709.
8. Z. Bartha, B. Kolesnik, and G. Kronenberg. $H$-percolation with a random $H$, Preprint available at arXiv : 2312.03663.
9. E. Bayraktar and S. Chakraborty, $K_\alpha$, graph bootstrap percolation, Electron. J. Combin 29 (2022), no. 1, 1.46–1.17.
10. J. van den Berg and H. Kesten, Inequalities with applications to percolation and reliability, J. Appl. Probab. 22 (1985), no. 3, 556–569.
11. M. Bidgoli, A. Mohammadian, and B. Tayfeh-Rezaie, *On K2,t-bootstrap percolation*, Graphs and Combinatorics 37 (2021), no. 3, 731–741.

12. B. Bollobás, “Weakly k-saturated graphs,” Beiträge zur Graphentheorie (Kolloquium, Manebach, 1967), B. G. Teubner Verlagsgesellschaft, Teubner, Leipzig, 1968, pp. 25–31.

13. J. Chalupa, P. L. Leath, and G. R. Reich, *Bootstrap percolation on a Bethe lattice*, J. Phys. C 21 (1979), L31–L35.

14. D. Conlon and W. T. Gowers, *Combinatorial theorems in sparse random sets*, Ann. Math. 184 (2016), no. 2, 367–454.

15. A. C. D. van Enter, *Proof of Straley’s argument for bootstrap percolation*, J. Stat. Phys. 48 (1987), no. 3-4, 943–945.

16. P. Erdős and A. Rényi, *On random graphs. I*, Publ. Math. Debrecen 6 (1959), 290–297.

17. A. E. Holroyd, “Sharp metastability threshold for two-dimensional bootstrap percolation”, Probab. Theory Relat. Fields 125 (2003), no. 2, 195–224.

18. G. Kalai, “Weakly saturated graphs are rigid,” Convexity and graph theory (Jerusalem, 1981), Vol 87, North-Holland Math. Stud., North-Holland, Amsterdam, 1984, pp. 189–190.

19. B. Kolesnik, *The sharp K4-percolation threshold on the Erdős–Rényi random graph*, Electron. J. Probab. 27 (2022), 13–23.

20. J. von Neumann, Theory of self-reproducing automata, University of Illinois Press, Champaign, IL, 1966.

21. Y. Peled. Personal communication.

22. M. Pollak and I. Riess, *Application of percolation theory to 2d-3d Heisenberg ferromagnets*, Phys. Status Solidi B 69 (1975), no. 1, K15–K18.

23. D. Saxton and A. Thomason, *Hypergraph containers*, Invent. Math. 201 (2015), no. 3, 925–992.

24. M. Schacht, *Extremal results for random discrete structures*, Ann. Math. 184 (2016), no. 2, 333–365.

25. R. H. Schonmann, *On the behavior of some cellular automata related to bootstrap percolation*, Ann. Probab. 20 (1992), no. 1, 174–193.

26. S. Ulam, “Random processes and transformations.” Proceedings of the international congress of mathematicians, Vol 2, Amer. Math. Soc, Cambridge, MA, 1950, pp. 264–275.

How to cite this article: Z. Bartha and B. Kolesnik, *Weakly saturated random graphs*, Random Struct. Alg. 65 (2024), 131–148. https://doi.org/10.1002/rsa.21210

APPENDIX A: SUPPLEMENTARY FACTS

A.1 Balanced graphs

We note here some basic facts about balanced graphs $H$. Recall Definition 1 and the definition of 2-balanced graphs, which appears below Definition 1. Also recall that we assume throughout this work that $\delta_H \geq 2$ and $v_H \geq 4$. Hence $e_H \geq v_H$.

**Lemma 22.** For any graph $H$, we have that $H$ is balanced if and only if $H \setminus e$ is 2-balanced for all edges $e \in E[H]$.

**Proof.** Suppose that $H \setminus e$ is 2-balanced for all edges $e \in H$. Let $F$ be a proper subgraph of $H$ with $v_F \geq 3$. Let $e \in E[H] \setminus E[F]$. Since $H' = H \setminus e$ is 2-balanced, it follows that $(e_F - 1)/(v_F - 2) \leq (e_{H'} - 1)/(v_{H'} - 2) = \lambda$. Thus $H$ is balanced.

On the other hand, if $H$ is balanced, then for any proper subgraph $F$ of some $H' = H \setminus e$ with $v_F \geq 3$, $(e_F - 1)/(v_F - 2) \leq \lambda = (e_{H'} - 1)/(v_{H'} - 2)$. Thus $H'$ is 2-balanced.

**Lemma 23.** For any graph $H$, if $H$ is balanced then it is connected.

**Proof.** Assuming that $H$ is balanced, we show that there is at least one edge between any two non-empty sets $V_1, V_2$ that partition the vertex set of $H$. Let $v_i$ and $e_i$ be the number of
vertices and edges, respectively, in the subgraph of $H$ induced by $V_i$, and $e_{12}$ the number of edges in $H$ between $V_1$ and $V_2$, so that $e_H = e_1 + e_2 + e_{12}$. If either $v_i \leq 2$ or $e_1 + e_2 \leq 3$ then $e_{12} \geq 1$, since $\delta_H \geq 2$ and $e_H \geq v_H \geq 4$. Hence assume that both $v_i \geq 3$ and $e_1 + e_2 \geq 4$. Then both

$$\frac{e_i - 1}{v_i - 2} \leq \lambda = \frac{e_1 + e_2 + e_{12} - 2}{v_1 + v_2 - 2}.$$  

Taking a weighted average, with weights $v_i - 2$, it follows that

$$\frac{e_1 + e_2 - 2}{v_1 + v_2 - 4} \leq \lambda,$$

and so

$$e_{12} \geq \left(\frac{v_1 + v_2 - 2}{v_1 + v_2 - 4} - 1\right)(e_1 + e_2 - 2) > 0.$$

▪