THE CRITICAL ACTIVATION DENSITY IN GRAPH BOOTSTRAP PERCOLATION

BRETT KOLESNIK, TAMÁS MAKAI, RAJKO NENADOV,
XAVIER PÉREZ-GIMÉNEZ, PAWEŁ PRAŁAT, AND MAKSIM ZHUKOVSKII

Dedicated to Brendan McKay and Nick Wormald

ABSTRACT. In graph bootstrap percolation, edges of an Erdős–Rényi random graph $\mathscr{G}_{n,p}$ are initially active. Activation spreads to other edges of the complete graph K_n by an iterative process governed by a fixed graph H, whereby an edge becomes active whenever it is the only inactive edge in a copy of H. If all edges of K_n are eventually activated, we say the process H-percolates. The case $H = K_3$ corresponds to the classical sharp threshold for connectivity in $\mathscr{G}_{n,p}$. When $H = K_4$, there are close connections with 2-neighbor bootstrap percolation from statistical physics. Varying H produces a wide range of behaviors.

In this work, for every graph H, we locate the critical H-percolation threshold $p_c(n,H)$, answering a question of Balogh, Bollobás, and Morris. Our general methods recover and improve several previous results. The location of $p_c(n,H)$ is related to a critical limiting density $\rho(H)$ of graphs that most efficiently activate a given edge. Introducing the parameter $\rho(H)$ raises several questions. For instance, it remains open whether $\rho(H)$ is computable in general, and its expression appears to indicate when the H-percolation threshold is sharp.

1. INTRODUCTION

In highly influential works, Ulam [29] and von Neumann [30] introduced cellular automata in an attempt to understand the mystery of biological self-replication. Although such processes evolve according to local rules, they can nonetheless lead to complex global behavior. Perhaps the most famous example is Conway's *Game of Life* [19], in which, at any given moment, each square in an infinite square grid takes value 0 or 1 (representing the absence and presence of life). The game consists of a simple set of rules by which squares change values, depending only on the values of those nearby. Despite its apparent simplicity, the game gives rise to a wide diversity of behavior, and it is generally undecidable whether a game will ever end.

Date: May 15, 2026.

2010 *Mathematics Subject Classification.* 05C35; 05C80; 82B43; 68Q80.

Key words and phrases. bootstrap percolation; cellular automaton; critical threshold; phase transition; random graph; weak saturation.

Shortly after the work of von Neumann, in the context of extremal combinatorics, Bollobás [13] introduced an intriguing class of dynamics that evolve in graphs. The update rule for this model is governed by a fixed graph H. To begin, we start with a graph G and let K_G denote the clique on $V(G)$, the vertex set of G. The edges in $G_0 = G$ are said to be *initially active*. In each step $t \geqslant 1$, a graph G_t induced by all edges activated by time t is obtained as follows: At time $t \geqslant 1$, we *activate* each edge e such that $H_e \setminus e \subseteq G_{t-1}$ for some copy $H_e \subseteq K_G$ of H containing e, and obtain G_t by adding all such edges e to G_{t-1}. In other words, an edge becomes active if it is the only inactive edge in an otherwise active copy of H in K_G. The resulting graph $\cup_{t\geqslant 0} G_t$ with all *eventually active* edges, obtained once the *H-dynamics* stabilize, is denoted by $\langle G \rangle_H$. The graph G is said to be *weakly H-saturated* if $\langle G \rangle_H = K_G$. The main result in [13] identifies, for all $3 \leqslant r \leqslant 6$, the minimum number $\mathrm{wsat}(n, K_r) = (r-2)n - \binom{r-1}{2}$ of edges in a weakly K_r-saturated graph on n vertices. This was later generalized, using a variety of methods, for all $r \geqslant 3$, in independent works by Alon [2], Frankl [17], Kalai [21, 22], and Lovász [24].

Remarkably, $\mathrm{wsat}(n, K_r)$ coincides with the minimal number $\mathrm{sat}(n, K_r)$ of edges in a *strongly K_r-saturated* graph, with the property that adding *any* edge creates a copy of K_r (i.e., one step of the H-dynamics turns the graph into K_n). In identifying $\mathrm{sat}(n, K_r)$, Erdős, Hajnal and Moon [16] showed there is a unique such graph, obtained by connecting each of $n-(r-2)$ many vertices to all vertices in a clique of size $r-2$. On the other hand, due to the more challenging, algorithmic nature of weak saturation, less is known about weakly K_r-saturated graphs. For instance, there is no explicit description of all weakly K_r-saturated graphs with $\mathrm{wsat}(n, K_r)$ many edges in the literature. One way to create such a graph, which we call an *r-Bollobás graph* (see [13, Fig. 4]), is to start with a clique on $r-2$ vertices and then iteratively add the remaining $n-(r-2)$ vertices one at time, in each step adding edges to exactly $r-2$ previously added vertices; but these are by no means the only K_r-percolating graphs with the minimal number of edges.

One of the most well-studied of all cellular automata is called *bootstrap percolation*. This model was introduced by Chalupa, Leath and Reich [15] to explain the sharp drop-off in the strength of a magnet observed as the concentration of a non-magnetic impurity reaches a critical value. A large literature has developed around the subject, which continues to remain active; see, e.g., the survey by Morris [25]. In bootstrap percolation, all vertices in a graph are initially *infected* independently with probability p. Further vertices become infected according to local rules. The most popular rule (although there are many variations) is the *r-neighbor rule*, under which a vertex becomes infected once it has at least r infected neighbors. Notice the connection with weak K_r-saturation: If the clique of size $r-2$ in an r-Bollobás graph is initially infected, then by induction all other vertices will become infected by the $(r-2)$-neighbor rule.

More recently, Balogh, Bollobás and Morris [8] introduced *graph bootstrap percolation*, starting the H-dynamics with an Erdős–Rényi random graph $G = \mathscr{G}_{n,p}$. In this context, if $\langle \mathscr{G}_{n,p} \rangle_H = K_n$ we say that the process *H-percolates*. Recalling that sites are initially infected independently with probability p in classical bootstrap percolation, it is thus quite natural to initialize the H-dynamics with $\mathscr{G}_{n,p}$, as then edges in K_n are initially active independently with probability p. The *critical H-percolation threshold* is defined as

$$p_c(n,H) = \inf\{p > 0 : \mathbb{P}(\langle \mathscr{G}_{n,p} \rangle_H = K_n) \geqslant 1/2\}. \tag{1}$$

As further motivation, we recall that weakly K_r-saturated graphs, and weakly H-saturated graphs more generally, are not well understood. As such, it is of interest to study the structure of weakly H-saturated *random* graphs. When p is close to p_c the graph $\mathscr{G}_{n,p}$ will, in some sense, favor the least-cost instances of such graphs. For example, when $\mathscr{G}_{n,p}$ is a barely super-critical K_4-percolating graph it is likely to contain a 4-Bollobás graph [23]. On the other hand, when $r \geqslant 5$, barely super-critical K_r-percolating random graphs $\mathscr{G}_{n,p}$ with high probability do not contain an r-Bollobás graph, but instead K_r-percolate in other ways that remain quite mysterious [11].

1.1. Main results.

One of the main problems in this area, stated as Problem 1 in [8], it to find the limit (assuming it exists)

$$\ell(H) = \lim_{n\to\infty} \frac{\log p_c(n,H)}{\log n},$$

for *every* graph H. Our first result answers this question, which the authors state [8, p. 438] "would represent a major step" towards the "ultimate aim of this line of research."

Definition 1. For a graph H, let $\mathscr{A}(H)$ be the set of all pairs (e,A) such that A is a graph and e is an edge in $\langle A \rangle_H \setminus A$. We define the *critical H-activation density* as

$$\rho(H) = \inf_{(e,A)\in\mathscr{A}(H)} \max_{e\subset F\subseteq A\cup e} \frac{e(F)-1}{v(F)-2}. \tag{2}$$

Above, and throughout this work, we let $v(F)$ and $e(F)$ denote the number of vertices and edges in a graph F.

Note that, if $e(H) = 0$ then $\rho(H) = \infty$ (by convention, since $\mathscr{A}(H) = \emptyset$); whereas, if $e(H) = 1$ then $\rho(H) = 0$. On the other hand, for $e(H) \geqslant 2$, one can easily verify that $0 < \rho(H) < \infty$. Throughout this work, we will for convenience assume that H has no isolated vertices, since adding isolated vertices to H has no effect on $p_c(n,H)$, for all large enough n.

We note that the "-1" and "-2" in (2) compensate for the edge e and its vertices. Indeed, the quantity inside the infimum can be viewed as the 2-density of $G = A \cup e$, "rooted" at e. We recall that the *2-density* $m_2(G)$ of a graph G is the maximum of $(e(F)-1)/(v(F)-2)$, over all subgraphs $F \subseteq G$ with $v(F) \geqslant 3$. When maximizing in (2), we require that $e \subset F$.

Our first main result locates $\ell(H)$.

Theorem 2. *For any graph H,*

$$\ell(H) = -1/\rho(H).$$

This result includes the trivial cases that $e(H) = 0$ (when $\rho(H) = \infty$ and $\ell(H) = 0$) and also $e(H) = 1$ (when $\rho(H) = 0$ and $\ell(H) = -\infty$), but we are mainly interested in the cases that $e(H) \geqslant 2$ so that $0 < \rho(H) < \infty$. The cases that $\rho(H) \leqslant 1$ are relatively simple. In Section 4 below, we prove hitting time results for all such H. These results, in particular, improve upon some of the results for specific H presented in [8, §5]. For the case that $\rho(H) > 1$ we prove the following.

Theorem 3. *Suppose that $\rho = \rho(H) > 1$. Then $p_c = p_c(n,H)$ satisfies*

$$\Omega(1/\log n) = n^{1/\rho} p_c = O(\log^2 n). \tag{3}$$

In fact, in Section 3 below, we prove a more technical upper bound of

$$n^{1/\rho} p_c = O(\log^{2/\rho + 2/a} n), \tag{4}$$

where if ρ is rational then a/b is its expression as an irreducible fraction, and if ρ is irrational then a can be taken to be arbitrarily large. For all $\rho > 1$, we have $a + \rho \leqslant a\rho$, so (4) implies the upper bound in (3).

We note that (4) is slightly weaker than the bound $n^{1/\rho} p_c = O(\log^{4/r} n)$ proved in [8, p. 416] for cliques $H = K_r$, with $r \geqslant 4$. However, (4) holds for *all* graphs H, with $\rho > 1$. In any case, we suspect (see Section 1.6 below) that $p_c = O(n^{-1/\rho})$ might hold for all such H.

1.2. **Outline.** A brief overview of our proof strategy is given in Section 1.4, after recalling the notion of a witness graph in Section 1.3. In Section 1.5, we recall the notion of a balanced graph H, and discuss how our results apply in this well-studied case. A number of open problems are presented in Section 1.6. In Section 2 we prove the lower bound and in Section 3 we prove the upper bound in our main result, Theorem 3. Finally, in Section 4, we prove hitting time results for all H with a leaf or with $\rho(H) \leqslant 1$.

1.3. **Witness graphs.** We note that $\rho(H)$ is related to the notion of a *witness graph*, as introduced in [8]. For any graph G, and each edge e in $\langle G\rangle_H$, the *witness graph algorithm (WGA)* [8] finds an inclusion-minimal subgraph $W_e \subseteq G$ such that e is in $\langle W_e\rangle_H$. The WGA is an inductive procedure:

- Put $W_e = e$ for each initially active edge e in $G_0 = G$.
- Then, for each edge e in G_t, activated at some time $t \geqslant 1$, we let H_e (chosen arbitrarily if not unique) be the copy of H that e completes. We put
$$W_e = \bigcup_{f\in E(H_e\setminus e)} W_f,$$
noting that $H_e \setminus e \subseteq G_{t-1}$, and so W_f has been defined by time $t-1$ for all edges f in $H_e \setminus e$.

By induction, for all edges e in $\langle G\rangle_H \setminus G$, the witness graph W_e is an inclusion-minimal subgraph $W_e \subset G$ for which $(e,W_e) \in \mathscr{A}(H)$. Therefore, if $(e,A) \in \mathscr{A}(H)$, then, by taking $G = A$, the WGA can be used to find an inclusion-minimal subgraph $W_e \subseteq A$ for which $(e,W_e) \in \mathscr{A}(H)$.

Since [8], most bounds on p_c have proceeded as follows:

- To obtain an upper bound on p_c, a reasonably efficient class of witness graphs (typically of logarithmic size) is identified. If the correlations of such structures in $\mathscr{G}_{n,p}$ can be controlled, they can be used to show that H-percolation occurs for all large enough p.
- To obtain a lower bound on p_c, a special property of the WGA is used, which is reminiscent of a property discovered by Aizenmann and Lebowitz [1] for 2-neighbor bootstrap percolation on $\mathbb{Z}^2$: Specifically, if we define the *size* of a witness graph W_e to be $v(W_e)-2$ (i.e., its number of vertices outside of e), then it can be seen (see [8, Lemma 13] and [9, Lemma 8]) that, started from any graph G, in each time step of the H-dynamics the maximum size of any witness graph defined so far increases by at most a factor of $e(H)$. Therefore, to show that some $e \in E(K_n)$ is, with high probability, *not* added by the H-dynamics to $\mathscr{G}_{n,p}$ it suffices to show that, with high probability, (1) e has no witness graph of size at most $\gamma \log n$ (for an appropriately chosen constant $\gamma > 0$) and (2) no other edge f has a witness graph of size between $\gamma \log n$ and $e(H)\gamma \log n$.

1.4. **Our approach.** A crucial barrier in applying the above proof strategy for arbitrary H is that in optimally efficient $(e,A) \in \mathscr{A}(H)$, the maximum in (2) is not necessarily achieved for $F = A\cup e$. In particular, this influences the correlations mentioned above.

When working with (2), we need to consider the *densest* subgraphs of witness graphs, rather than the witness graphs themselves. Recall that in the definition of $\rho(H)$, for each $A \in \mathscr{A}(H)$, we maximize the ratio $(e(F)-$

$1)/(v(F)-2)$ over all subgraphs $e \subset F \subseteq A \cup e$. If some F maximizes this ratio, we call $D = F \setminus e$ a *densest part* of A (with respect to e). For most graphs H that have been studied, such as cliques K_r (and, more generally, the balanced graphs H discussed in Section 1.5 below), in the most efficient witness graphs W (e.g., in K_r-ladders, as in Figure 1 below), the entire graph $D = W$ is *the* densest part of W.

In proving the lower bound in Theorem 3, we show that an Aizenmann–Lebowitz-type property continues to hold when looking at the dense parts D of a general witness graph W (for which possibly $D \neq W$). The key idea is to look at the *maximal* such $D_* \subseteq W$, and then argue that after each time step of WGA the maximum size over all such D_* defined so far increases by at most a factor of $e(H)$.

The upper bound in Theorem 3 is more complicated. We start by selecting some $(e, A_*) \in \mathscr{A}(H)$ such that

$$\max_{e \subset F \subseteq A_* \cup e} \frac{e(F)-1}{v(F)-2} < \rho(H) + \frac{1}{\log n}.$$

Note that A_* depends on n, and is nearly as efficient as possible at activating a given edge e. The difficultly, however, lies in the fact that we have no control whatsoever over the size of A_*. Indeed, A_* may even be *much* larger than n, the size of $\mathscr{G}_{n,p}$. As such, proof techniques used in previous works do not apply. To overcome this challenge, we introduce a novel technique of embedding witness graphs in $\mathscr{G}_{n,p}$: we will in some sense "unfold" the graph A_* (to essentially obtain a covering graph of A_*), and then recursively, one small extension at a time, "fold" a modified version of A_* into $\mathscr{G}_{n,p}$.

1.5. **Balanced graphs.** Prior to the current work, the most general results concerned the following class of graphs: Following [8], we say that a graph H is *balanced* if $v(H) \geqslant 4$ and (cf. (2) above)

$$\frac{e(F)-1}{v(F)-2} \leqslant \lambda(H) := \frac{e(H)-2}{v(H)-2}, \tag{5}$$

for all subgraphs $F \subset H$ with $3 \leqslant v(F) < v(H)$.[1]

For instance, all cliques K_r with $r \geqslant 4$ are balanced. Naturally, H is *strictly balanced* if (5) holds with $\leqslant$ replaced by $<$. Cliques K_r are strictly balanced for all $r \geqslant 5$; however, K_4 is balanced but not strictly. In some sense, "most" graphs H are strictly balanced since, as shown by Bartha, Kronenberg and the first author [10], almost surely, $\mathscr{G}_{\infty,1/2}[[r]]$ is strictly balanced for all large r, where $\mathscr{G}_{\infty,1/2}$ is the *Rado graph*, where every pair of vertices from the

[1] In [8] it is further assumed that $e(H)/2 \geqslant v(H)-1$. This implies $\lambda(H) \geqslant 2$, which is used in the proof of Lemma 6 in [8, p. 419]. However, in light of Lemma 16 in [9], only the minimal assumption that $v(H) \geqslant 4$ is required.

infinite set $\mathbb{Z}_{>0}$ is adjacent with probability $1/2$, and $\mathscr{G}_{\infty,1/2}[[r]] \sim \mathscr{G}_{r,1/2}$ is its induced subgraph on the first r vertices.

A main result in [8, Theorem 1] shows that, for all $r \geqslant 5$,

$$p_c(n,K_r) = n^{-1/\lambda(K_r)+o(1)}. \tag{6}$$

The proof of the lower bound introduces the WGA, as described above, and also shows that when $H = K_r$, any witness graph W with $v(W)$ vertices has at least

$$e(W) \geqslant \lambda(K_r)(v(W)-2)+1 \tag{7}$$

many edges.

On the other hand, the proof of the upper bound works for all balanced graphs H, using a special type of witness graph called an *H-ladder*. An H-ladder L of height h is obtained by stringing together a series of h many copies of H in such a way that adjacent pairs intersect in a single edge. To obtain L, all such shared edges are removed, together with another edge e from one of the copies of H at an end of the series. By induction, the H-dynamics activate all such removed edges, and L is a witness graph for e. An H-ladder L of height h has $v(L) = h(v(H)-2)+2$ many vertices and

$$e(L) = h(e(H)-2)+1 = \lambda(H)(v(L)-2)+1$$

many edges; see Figure 1 (cf. [8, Fig. 1]).

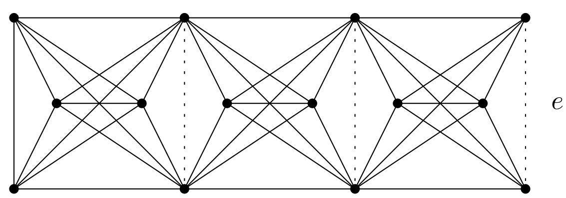

FIGURE 1. A K_6-ladder L of height $h=3$. All edges that join consecutive copies of K_6 and the edge e at the end of the ladder are added to L by the K_6-dynamics, so that $(e,L) \in \mathscr{A}(K_6)$.

In subsequent work by Bartha and the first author [9], it is shown that (6) extends to all balanced graphs H. A key step in this proof is to show (see [9, Proposition 9]) that, for any H (with at least 4 vertices and minimum degree at least 2), a witness graph W with $v(W)$ many vertices has at least

$$e(W) \geqslant \lambda_*(H)(v(W)-v(H))+e(H)-1 \tag{8}$$

many edges, where

$$\lambda_*(H) = \min_F \frac{e(H)-e(F)-1}{v(H)-v(F)},$$

minimizing over all subgraphs $F \subset H$ with $2 \leqslant v(F) < v(H)$ vertices. In [9, Theorem 4] it is shown that

$$p_c(n,H) \geqslant n^{-1/\lambda_*(H)+o(1)}.$$

Prior to the current work, this was the best general lower bound on p_c.

It can be shown (see [9, Lemma 10]) that $\lambda_*(H) \leqslant \lambda(H)$, with equality if and only if H is balanced. Hence, in particular, (7) above extends from K_r to all balanced graphs H, noting that when $\lambda_*(H) = \lambda(H)$ the right hand side in (8) simplifies to

$$\lambda(H)(v(W) - v(H)) + e(H) - 1 = \lambda(H)(v(W) - 2) + 1.$$

As such, as shown in [9, Theorem 2],

$$p_c(n,H) = n^{-1/\lambda(H)+o(1)}, \tag{9}$$

for all balanced graphs H.

We note that, by Theorem 2 and the following observation, we recover the general result (9) for balanced graphs.

Proposition 4. *For all balanced graphs H, we have that $\rho(H) = \lambda(H)$.*

To prove this result, we will require the technical fact, mentioned above, that if H is balanced then any witness graph W has at least $\lambda(H)(v(W) - 2) + 1$ edges. Let us briefly sketch a proof (see [9] for details) of this fact:

- First, we slow down the H-dynamics so that in each step a single copy of H is completed, and suppress the steps of WGA that do not contribute to W. This is called the *red edge algorithm (REA)* in [8].
- Then, at any given time, the edges in W (the black edges in REA) and those so far added to W (the red edges in REA) by the H-dynamics can be partitioned into a collection of edge-disjoint components, where the edges in two copies of H are in the same component if the copies share an edge.
- By induction, it can be shown that each component C has at least $\lambda(H)(v(C) - 2) + 1$ edges in W (i.e., black edges). This proves the claim, since by the end of the REA there is only one component, the witness graph W.
- In the inductive step, suppose that a copy H' of H merges together with some number of components $C_1, \dots, C_m$ to form a new component C. If C_i has $v_i \geqslant 2$ vertices in common with H' then, since H is balanced, they have at most $\lambda(v_i - 2) + 1$ edges in common (see (5) above), so the loss of vertices and edges caused by the overlap compensates for one another.

Proof of Proposition 4. Suppose that $(e,A) \in \mathscr{A}(H)$. The WGA gives an inclusion-minimal subgraph $W \subseteq A$ for which $(e,W) \in \mathscr{A}(H)$. Since H is balanced, $e(W) \geqslant \lambda(H)(v(W)-2)+1$. Since W contains the vertices in e but not the edge e itself, the graph $F = W \cup e$ satisfies

$$\frac{e(F)-1}{v(F)-2} = \frac{e(W)}{v(W)-2} \geqslant \lambda(H) + \frac{1}{v(W)-2}. \tag{10}$$

Since there are, e.g., arbitrarily large H-ladders for which (10) is an equality, it follows, taking an infimum over all $(e,A) \in \mathscr{A}(H)$, that $\rho(H) = \lambda(H)$. ■

1.6. **Questions.** The expression (2) above for $\rho(H)$ is rather abstract. At this level of generality, a simpler expression may not exist. Currently, the most interesting example of a unbalanced graph H for which the critical exponent has been explicitly identified is the result that $p_c(n,K_{2,4}) = \Theta(n^{-10/13})$ by Bidgoli, Mohammadian, and Tayfeh-Rezaie [12].

In recent work of Balister, Bollobás, Morris, and Smith [7] it has been announced that forthcoming work will show that the exponent in p_c for various monotone cellular automata on Euclidean grids/lattices is uncomputable in general. This leads us to ask the following question.

Question 5. *Is $\rho(H)$ computable for general H?*

Although we do not know if $\rho(H)$ can be defined effectively, we also ask the following question.

Question 6. *Is $\rho(H)$ rational for all H?*

A somewhat related question on possible values of the limit $\mathrm{wsat}(n,H)/n$ was asked by Terekhov and the sixth author [27]. In particular, it is known that this limit can be fractional (and therefore linear algebraic methods give suboptimal results; see [28] by the same authors), but it is not known whether the limit is always rational. Some progress towards describing the possible values of the limit was recently achieved by Ascoli and He [6].

Bartha and the first author [9] have shown that $p_c(n,H) = O(n^{-1/\rho(H)})$ for all strictly balanced graphs (in which case, by Proposition 4 above, $\rho(H) = \lambda(H)$). It is known (see [8, §5] and Section 4.1 below) that $p_c(n,H) = \Theta(n^{-1/\rho(H)})$ for all graphs H with a *leaf* (i.e., a vertex of degree 1). We are not aware of any graphs H for which $\rho(H) > 1$ and $p_c(n,H) \gg n^{-1/\rho(H)}$.

Question 7. *Does the upper bound*

$$p_c(n,H) = O(n^{-1/\rho(H)})$$

hold for all graphs H with $\rho(H) > 1$?

On the other hand, as shown in Section 4.4 below, if $\rho(H) = 1$ and the minimum degree $\delta(H) = 2$ then $p_c(n,H) \sim (\log n)/n$.

Finally, let us define

$$p_\varepsilon(n,H) = \inf\{p > 0 : \mathbb{P}(\langle \mathscr{G}_{n,p} \rangle_H = K_n) \geqslant \varepsilon\}. \tag{11}$$

Observe (see (1) above) that $p_c(n,H) = p_{1/2}(n,H)$. We say that the H-percolation threshold is *sharp* if, for all small $\varepsilon > 0$,

$$p_{1-\varepsilon}(n,H) - p_\varepsilon(n,H) \ll p_c(n,H), \tag{12}$$

and call it *coarse* otherwise. Problem 2 in [8] asks when the H-percolation threshold is sharp in this sense. We believe that sharpness is related to whether the critical H-activation density is ever attained for finite A.

Question 8. *Is the H-percolation threshold sharp if and only if*

$$\max_{e\subset F\subseteq A\cup e} \frac{e(F)-1}{v(F)-2} > \rho(H), \tag{13}$$

for all $(e,A) \in \mathscr{A}(H)$?

The K_r-percolation threshold is sharp for all cliques. It is shown in [8] that $p_c(n,K_4) = \Theta((n\log n)^{-1/2})$ and, as noted in [8], this result together with Friedgut [18, Theorem 1.4] implies sharpness. Further work [4,5,23] shows that $p_c(n,K_4) \sim (3n\log n)^{-1/2}$. Recent work by Bartha, the first author, Kronenberg, and Peled has shown, for all $r \geqslant 5$, that $p_c(n,K_r) \sim \gamma_r n^{-1/\lambda(K_r)}$, where γ_r is described in terms of the $(\binom{r}{2}-1)$th Fuss–Catalan numbers. Note that, for $r \geqslant 5$, there is no longer a polylogarithmic term in the asymptotics for $p_c(n,K_r)$, but nonetheless the threshold continues to be sharp. Since cliques are balanced it follows (see Proposition 4 and (10) above) that (13) holds when $H = K_r$, in line with Question 8.

On the other hand, the H-percolation threshold is coarse, e.g., for all H with a leaf; see Section 4.1 below.

1.7. **Acknowledgements.** This project began at the MATRIX event *Combinatorics of McKay and Wormald* in June 2025. We thank the organizers, Jane Gao, Catherine Greenhill, Mikhail Isaev, Anita Liebenau, Ian Wanless, and the MATRIX staff. This work is dedicated to the lasting influence of Brendan McKay and Nick Wormald on the authors. The fourth author's research was supported in part by NSF grant DMS2201590. The fifth author's research was supported in part by NSERC DG grant RGPIN-2022-03804.

2. Lower bound when $\rho(H) > 1$

In this section, we prove the following result; cf. [8, Proposition 8] for the case $H = K_r$.

Theorem 9. *Suppose that $\rho = \rho(H) > 1$ and $np^\rho \log^\rho n = \varepsilon$. Then, for any sufficiently small $\varepsilon > 0$, we have that $\langle \mathscr{G}_{n,p}\rangle_H \neq K_n$ with high probability.*

Throughout this section, we will assume that $\rho(H) > 1$. In particular, this implies that $e(H) \geqslant v(H)$.

For graphs $A \subset B$, let

$$\rho(A,B) = \frac{e(B)-e(A)}{v(B)-v(A)}$$

and put

$$\rho^{\max}(A,B) = \max_{A\subset B'\subseteq B} \rho(A,B').$$

Note that

$$\rho(H) = \inf_{(e,A)\in\mathscr{A}(H)} \rho^{\max}(e, A\cup e). \tag{14}$$

The following observations are straightforward. We omit the proof.

Lemma 10. *Let $A \subset B \subset C$. Then, for any $\alpha > 0$,*

(1) if $\rho(A,B) \geqslant \alpha$ and $\rho(B,C) \geqslant \alpha$ then $\rho(A,C) \geqslant \alpha$;

(2) if $\rho(A,B) < \alpha$ and $\rho(B,C) < \alpha$ then $\rho(A,C) < \alpha$.

Lemma 11. *Let A be a graph. For each edge e in A, put $W_e = F_e = e$. Recall that, for all other edges e in $\langle A\rangle_H \setminus A$, the WGA identifies a copy H_e of H completed by e and an inclusion-minimal subgraph*

$$W_e = \bigcup_{f\in E(H_e\setminus e)} W_f \subset A$$

for which $(e,W_e) \in \mathscr{A}(H)$. Then, for each such e, there is some $e \subset F_e \subseteq W_e \cup e$, for which $\rho(e,F_e) \geqslant \rho(H)$ and

$$v(F_e) - 2 \leqslant \sum_{f\in E(H_e\setminus e)} v(F_f). \tag{15}$$

The "-2" in (15) compensates for the vertices in e, as they may not be in any of the edges $f \in E(H_e \setminus e)$.

Proof. We will prove this claim by induction, on the time at which an edge is activated by the H-dynamics starting with A. In doing so, it will be helpful to also establish, for each edge e in $\langle A\rangle_H \setminus A$, the existence of some B_e such that either (1) $F_e = B_e$ or else (2) $F_e \subset B_e$ and $\rho^{\max}(F_e,B_e) < \rho(H)$, and in either case $(e, B_e \setminus e) \in \mathscr{A}(H)$. Let us stress that $B_e \setminus e$ is not necessarily a subgraph of A, but is rather a technical device that will be used in the proof.

For an edge e added at time $t = 1$, we simply take $B_e = H_e$ and let F_e be an inclusion-maximal subgraph $e \subset F_e \subseteq H_e$ for which $\rho(e,F_e) \geqslant \rho(H)$. We note that (15) holds, since $\rho(H) > 1$ and so, as noted above, $e(H) \geqslant v(H)$. If $F_e \neq B_e$ then, by Lemma 10(1), it follows by the choice of F_e that $\rho^{\max}(F_e,B_e) < \rho(H)$.

Next, if some e is added at time $t > 1$, then all $f \in E(H_e \setminus e)$ were added by time $t-1$, so the induction hypothesis applies to these edges. To begin, we put

$$F'_e = e \cup \bigcup_{f \in E(H_e \setminus e)} F_f \setminus f,$$

and obtain B_e from F'_e by attaching a disjoint copy of $B_f \setminus F_f$ to each $F_f \setminus f$ above. By construction, $(e, B_e \setminus e) \in \mathscr{A}(H)$.

First, note that if all the $B_f = F_f$ then $B_e = F'_e$. In this case, since $(e, B_e \setminus e) \in \mathscr{A}(H)$, we can take F_e to be an inclusion-maximal subgraph $e \subset F_e \subseteq F'_e$ satisfying $\rho(e, F_e) \geqslant \rho(H)$.

On the other hand, suppose that at least one of the $B_f \neq F_f$. Then, by construction, $\rho^{\max}(F'_e, B_e) < \rho(H)$. We claim that $\rho^{\max}(e, F'_e) \geqslant \rho(H)$. Indeed, otherwise it would follow by Lemma 10(2) that $\rho^{\max}(e, B_e) < \rho(H)$, in contradiction to $(e, B_e \setminus e) \in \mathscr{A}(H)$. To conclude, we let F_e be an inclusion-maximal subgraph $e \subset F_e \subseteq F'_e$ satisfying $\rho(e, F_e) \geqslant \rho(H)$. Once again, by Lemma 10(1), it follows that $\rho^{\max}(F_e, B_e) < \rho(H)$. ■

The previous lemma has the following useful consequence.

Corollary 12. *Let A be a graph. Consider the graphs F_e in Lemma 11. Suppose that, for some e in $\langle A \rangle_H \setminus A$, we have $v(F_e) > v(H)e(H)$. Then, for some f that is added to A before e during the WGA, we have*

$$v(F_e)/e(H) \leqslant v(F_f) < v(F_e).$$

Proof. This follows noting, by (15) in Lemma 11, that in each time step of the WGA the maximum size over all F_e defined so far increases by at most a factor of $e(H)$. Therefore, if the WGA eventually defines an F_e of size $v(F_e)$, then there must have previously been defined some smaller F_f of size at least $v(F_e)/e(H)$. ■

We are ready to prove the main result of this section. This proof is inspired by the proof of Proposition 8 in [8]. The crucial difference, however, is that we look at the dense parts F of witness graphs W, as given by Lemma 11 above, rather than the witness graphs W themselves.

Proof of Theorem 9. Suppose that $np^\rho \log^\rho n = \varepsilon$, for a sufficiently small constant $\varepsilon > 0$, as determined below.

We apply the WGA to $\mathscr{G}_{n,p}$ and use the graphs F_e in Lemma 11. Suppose that some given e is in $\langle \mathscr{G}_{n,p} \rangle_H$. Then either:

(1) $v(F_e) \leqslant e(H) \log n$; or else
(2) some edge f satisfies $\log n \leqslant v(F_f) \leqslant e(H) \log n$.

In the first case, there would be some graph F containing the vertices of e and $k \leqslant e(H) \log n$ many other vertices and at least $\rho k + 1$ edges. Taking a union bound, the probability of this event is at most

$$\sum_{k=0}^{e(H)\log n} n^k \binom{(k+2)^2}{\rho k+1} p^{\rho k+1} \leqslant O(p\log n) \sum_{k=0}^{e(H)\log n} (Cnp^\rho \log^\rho n)^k$$
$$= O(p\log n) \sum_{k=0}^{\infty} (C\varepsilon)^k,$$

for some constant C depending only on H. For any small enough $\varepsilon > 0$, the above expression is $O(p\log n) \to 0$ as $n \to \infty$.

Likewise, in the second case, the probability is bounded by

$$O(pn^2 \log n) \sum_{k=\log n}^{e(H)\log n} (C\varepsilon)^k = O(n^{2-1/\rho+\log(C\varepsilon)} \log^2 n) \to 0,$$

for any small enough $\varepsilon > 0$. ■

3. Upper bound when $\rho(H) > 1$

In this section, we will prove the following result.

Theorem 13. *Fix a graph H, with $\rho = \rho(H) > 1$. If ρ is rational, let a/b be its expression as an irreducible fraction. If ρ is irrational, let b be arbitrarily large. Suppose that $np^\rho = A\log^{2+2/b} n$. Then, for large enough A, with high probability, we have that $\langle \mathscr{G}_{n,p} \rangle_H = K_n$.*

Throughout this section, we fix some $(e^*, A^*) \in \mathscr{A}(H)$ for which

$$\rho^{\max}(e^*, A^* \cup e^*) < \rho + 1/\log n. \tag{16}$$

Such an A^* exists by the definition (14) of ρ. Furthermore, we let $W^* \subseteq A^*$ denote the witness graph for e^* found by the WGA, as discussed in Section 1.3 above. By (16), W^* is an efficient witness graph.

In proving Theorem 13, the overall plan is to use W^* as a basis to show that all missing edges in $\mathscr{G}_{n,p}$ are activated eventually. For instance, if W^* was small, with only $v(W^*) = O(\log n)$ many vertices, say, we might naturally try to embed a copy of W^* into $\mathscr{G}_{n,p}$ "based" at each missing edge e in $\mathscr{G}_{n,p}$ (i.e., with the vertices of e^* mapped to those of e). However, in appealing to (14), we have no control on $v(W^*)$. In particular, $v(W^*)$ could even be much larger than n, the size of $\mathscr{G}_{n,p}$ itself.

The main idea behind our proof of Theorem 13 is to use suitably "compressed" versions of W^* to activate all missing edges in $\mathscr{G}_{n,p}$. Roughly speaking, this will be achieved by "unfolding" W^* and then "folding" it into $\mathscr{G}_{n,p}$ by recursively embedding a series of small extensions. Crucially, the "unfolded" version of W^* preserves the local structure of W^*, so that activation patterns at all stages of the process are maintained. Indeed, the "unfolded" version of W^* can be viewed as its covering graph, and we will show that there exists a covering map from this "unfolded" version of W^* into $\mathscr{G}_{n,p}$ that is based at each missing edge of $\mathscr{G}_{n,p}$.

3.1. Extensions and cores.

The first step towards our proof of Theorem 13 is to formally introduce the notion of an extension.

Definition 14. Let $S \subset F$ be graphs and suppose that $V(F)$ is ordered. Let $S' \subset F' \subseteq G$ be some other graphs, where $V(G)$ is ordered. Then we call F' an (S,F)-*extension* of S' in G if there is an isomorphism of the graphs $F \setminus E(S)$ and $F' \setminus E(S')$ such that the ith largest vertex in S is mapped to the ith largest vertex in S', for each $1 \leqslant i \leqslant v(S)$.

In other words, F' extends S' in the same way as F extends S. Note that, in the definition above, F' is not necessarily an induced subgraph of G, and that the edges in $F[S]$ and $G[S']$ play no role.

Definition 15. For graphs $S \subset F$ and some $\alpha > 0$, we say that (S,F) is α-*dense* if

$$e(F) - e(F') \geqslant \alpha(v(F) - v(F')),$$

for *every* induced subgraph $S \subseteq F' \subset F$.

In other words, for any partial revealment $F' \supseteq S$ of F, we will need to add at least α edges per vertex in order to reveal the rest of F. In light of this, let us make the following observation.

Lemma 16. *Let $e \subset F_1, F_2$ and suppose that (e,F_1) and (e,F_2) are α-dense. Then $(e, F_1 \cup F_2)$ is α-dense.*

Proof. Indeed, let F' be an induced subgraph of $F_1 \cup F_2$ containing e. Let x be the number of vertices in F_1 that are not in F', and let y be the number of vertices in F_2 that are not in $F' \cup F_1$. Then there are at least αx many edges in F_1 with at most one endpoint in F', and at least αy many edges in F_2 with at most one endpoint in $F' \cup F_1$. Hence

$$e(F_1 \cup F_2) - e(F') \geqslant \alpha(x+y) = \alpha(v(F_1 \cup F_2) - v(F')),$$

as required; see Figure 2. ∎

Definition 17. Let $(e,A) \in \mathscr{A}(H)$. The α-*core* $C_e \subseteq A$ of (e,A) is defined as follows. If there is no subgraph $e \subset F \subseteq A \cup e$ such that (e,F) is α-dense, let C_e be the empty graph on the vertices of e. Otherwise, let $C_e = F_e \setminus e$ where F_e is the maximum such F.

Note that, by Lemma 16, C_e is well defined.

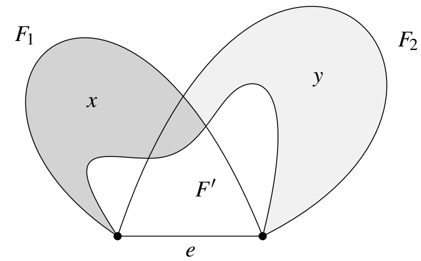

FIGURE 2. Illustration of the proof of Lemma 16. If there are x vertices (dark grey) in F_1 that are not in F' and y vertices (light grey) in F_2 that are not in $F' \cup F_1$, then there are at least αx edges in F_1 with at most one endpoint in F', and at least αy edges in F_2 with at most one endpoint in $F' \cup F_1$.

Definition 18. Let $S \subset F$. Naturally, we say that (S,F) is α-*sparse* if $\rho^{\max}(S,F) < \alpha$; that is, if

$$e(F') - e(S) < \alpha(v(F') - v(S)),$$

for *every* $S \subset F' \subseteq F$.

Lemma 19. *Let $(e,A) \in \mathscr{A}(H)$. Suppose that $C_e \subset A$ is the α-core of (e,A). Then (C_e,A) is α-sparse.*

Proof. For simplicity, put $C = C_e$. Indeed, towards a contradiction, let F be a minimal $C \subset F \subseteq A$ for which

$$e(F) - e(C) \geqslant \alpha(v(F) - v(C)). \tag{17}$$

Then we claim that (e,F) is α-dense, in contradiction to the fact that C is the α-core (the maximum such F). To see this, first note that, by the minimality of F, for any $C \subset F'' \subset F$ we have that

$$e(F'') - e(C) < \alpha(v(F'') - v(C)). \tag{18}$$

Next, consider some $e \subset F' \subset F \cup e$. Suppose there are x many vertices in C that are not in F', and y many vertices in F that are not in $F' \cup C$. Since C is the α-core, there are at least αx many edges in C with at most one endpoint in F'. By (17) and (18), there are at least αy many edges in F with at most one endpoint in $F' \cup C$. Hence

$$e(F) - e(F') \geqslant \alpha(x+y) = \alpha(v(F) - v(F')),$$

and so (e,F) is α-dense, yielding the desired contradiction; see Figure 3. ∎

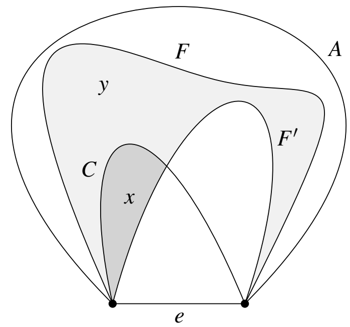

FIGURE 3. Illustration of the proof of Lemma 19. If there are x vertices (dark grey) in C that are not in F' and y vertices (light grey) in F that are not in $F' \cup C$, then there are at least αx edges in C with at most one endpoint in F', and at least αy edges in F with at most one endpoint in $F' \cup C$.

3.2. **Red and black edges.** As discussed in Section 1.3, the WGA applied to A^* begins by setting $W_e^* = e$ for each initially active (at time $t = 0$) edge e in A^*. Then, for each edge e activated at some time $t \geqslant 1$ it selects (arbitrarily if not unique) some copy H_e^* completed by e at time t and sets

$$W_e^* = \bigcup_{f \in E(H_e^* \setminus e)} W_f^*. \tag{19}$$

Since all f in $H_e^* \setminus e$ have been activated by time $t-1$, all W_f^* have already been defined, and so W_e^* is well defined. Note that the witness graph $W^* \subseteq A^*$ discussed above is $W^* = W_{e^*}^*$.

The following definition is similar to the *red edge algorithm (REA)* in [8], where we run the WGA but suppress all witness graphs that do not contribute to the formation of the witness graph W^* of interest.

Definition 20. For each e in A^*, put $\mathscr{R}_e^* = \emptyset$ and recall that $W_e^* = e$. For edges e activated at time $t \geqslant 1$, we put

$$\mathscr{R}_e^* = \{e\} \cup \bigcup_{f \in E(H_e^* \setminus e)} \mathscr{R}_f^*.$$

We call the edges in $\mathscr{R}_e^*$ the *red edges* of W_e^*. The edges in W_e^* itself are its *black edges*. Note that if the black edges of W_e^* are initially active then the H-dynamics will activate each red edge of W_e^*, towards eventually activating e itself. We let $\mathscr{R}^* = \mathscr{R}_{e^*}^*$ denote the set of red edges of W^*.

3.3. **Small cores.** Throughout the remainder of this section, we let

$$\alpha_* = \rho + (\rho + 2)/\log n.$$

Our next key observation is that, although we have no control over the size of W^* itself, it is however the case that W_f^* has a small α_*-core for each red edge f of W^*. As we will see, this will allow us to embed a "folded" version of W^*, one small extension at a time, based at each missing edge of $\mathscr{G}_{n,p}$.

Definition 21. For each red edge $f \in \mathscr{R}^*$ of W^*, we let C_f^* denote the α_*-core of (f, W_f^*). If $f \in E(W^*)$ is a black edge of W^*, we put $C_f^* = f$.

Lemma 22. *For all large n, we have $v(C_f^*) \leqslant 2\log n$, for all $f \in \mathscr{R}^*$.*

Proof. Suppose that, for some $f \in \mathscr{R}^*$, there is some subgraph $f \subset F \subset W_f^* \cup f$ for which

$$\frac{e(F)-1}{v(F)-2} \geqslant \rho + \frac{\rho+2}{\log n}.$$

Recall that $W_f^* \subset W^*$. Therefore, $e^* \subset F' \subset W^* \cup e^*$, where $F' = (F \setminus f) \cup e^*$. Also recall that $W^* \subseteq A^*$. Hence, by (16), and since $e(F') = e(F)$ and $v(F') \leqslant v(F) + 2$, it follows that

$$\frac{e(F)-1}{v(F)} \leq \frac{e(F')-1}{v(F')-2} \leqslant \rho + \frac{1}{\log n}.$$

Altogether, we find that

$$\frac{e(F)-1}{v(F)} - \frac{1}{\log n} \leqslant \rho \leqslant \frac{e(F)-1}{v(F)-2} - \frac{\rho+2}{\log n},$$

and so

$$\frac{\rho+1}{\log n} \leqslant \frac{2}{v(F)} \cdot \frac{e(F)-1}{v(F)-2} \leqslant \frac{2}{v(F)-2}\left(\rho + \frac{1}{\log n}\right).$$

Hence

$$v(F) \leqslant \frac{2\rho}{\rho+1}\log n + O(1) \leqslant 2\log n,$$

for all large n, and the result follows. ∎

3.4. **Modified cores.** As discussed, we aim to embed a "folded" version of W^* based at each missing edge e in $\mathscr{G}_{n,p}$, and we plan to do so one small extension at a time. In order to allow for such a recursive construction, it will be useful to consider the following modified version of the α-core.

Definition 23. For each red edge $f \in \mathscr{R}^*$ in W^*, we let

$$M_f^* := \bigcup_{g \in E(H_f^* \setminus f)} C_g^* \tag{20}$$

denote the *modified α_*-core* of f. In particular, we let $M^* = M_{e^*}^*$ denote the *modified α_*-core* of e^*.

Note that, by Lemma 22, for all large n,

$$v(M_f^*) \leqslant \sum_{g\in E(H_f^*\setminus f)} v(C_g^*) \leqslant 2e(H)\log n, \tag{21}$$

for all $f \in \mathscr{R}^*$.

3.5. **Extensions in $\mathscr{G}_{n,p}$.** The following technical result will be used in our proof below of Theorem 13. Informally, it says that, with high probability, for *all* sufficiently small and sparse (S,F)-extensions, *all* subgraphs $S' \subset \mathscr{G}_{n,p}$ have an (S,F)-extension in $\mathscr{G}_{n,p}$. This fact will allow us to recursively build a "folded" version of W^* starting from each missing edge in $\mathscr{G}_{n,p}$.

Lemma 24. *Let ρ and $np^\rho = A\log^{2+2/b} n$ be as in Theorem 13. Then, for large enough A, with high probability, for every (S,F) such that*

- *$V(F)$ is ordered,*
- *$v(F) \leqslant 2e(H)\log n$, and*
- *(S,F) is α_*-sparse,*

every subgraph $S' \subset \mathscr{G}_{n,p}$ with $v(S') = v(S)$ has an (S,F)-extension in $\mathscr{G}_{n,p}$ (where $V(S')$ is ordered by the natural order of the integers).

Proof. For technical convenience, we will restrict our attention to *ordered* (S,F)-extensions, such that,

- as above, the ith largest vertex in S is mapped to the ith largest vertex in S', for each $1 \leqslant i \leqslant v(S)$, but that also
- the jth largest vertex in $V(F)\setminus V(S)$ is mapped to the jth largest vertex in $V(F')\setminus V(S')$, for each $1 \leqslant j \leqslant v(F)-v(S)$.

As such, we essentially ignore the symmetries of F outside of S, and consider one particular way of extending S. This simplifies various calculations in our analysis.

Fix some (S,F), as above, and some $V \subset [n]$ of size $|V| = v(S)$. Let S' denote the subgraph of $\mathscr{G}_{n,p}$ induced by V, and let X denote the number of ordered (S,F)-extensions of S' in $\mathscr{G}_{n,p}$.

For ease of notation, put $B = 2e(H)$ and note that $v(F) \leqslant B\log n$. Also note that, since (S,F) is α_*-sparse, it follows that

$$\rho^{\max}(S,F) \leqslant \rho + C\rho/\log n,$$

where $C = 1+2/\rho$.

Applying Janson's inequality, we will show that, for all large n,

$$-\log \mathbb{P}(X=0) > B^2\log^2 n. \tag{22}$$

Taking a union bound, this will prove the lemma, since the number of relevant (S,F) and V is at most

$$2^{B\log n+(B\log n)^2/2} n^{B\log n} \ll e^{B^2\log^2 n}.$$

Recall (see, e.g., [3]) that Janson's inequality implies that

$$-\log \mathbb{P}(X=0) \geqslant \frac{\mu^2}{\mu+\Delta} = \frac{1}{1/\mu+\Delta/\mu^2},$$

where $\mu = \mathbb{E}X$ is the expected number of ordered (S,F)-extensions of S' in $\mathscr{G}_{n,p}$, and Δ is the expected number of pairs of distinct ordered (S,F)-extensions of S' in $\mathscr{G}_{n,p}$ that share at least one edge outside of S'. To prove (22), we will show that $\mu \gg \log^2 n$ and $\mu^2/\Delta > 2B^2 \log^2 n$.

Let $s = v(S)$ and $t = v(F) - v(S)$.

First, to see that $\mu \gg \log^2 n$, we note (using $\binom{n}{k} \geqslant (n/k)^k$) that

$$\mu = \binom{n-s}{t} p^{e(F)-e(S)} \geqslant (1-o(1)) \left(\frac{np^{\rho^{\max}(S,F)}}{t} \right)^t,$$

since there are at most $t\rho^{\max}(S,F)$ many edges in F that are not in S. Next, we observe that

$$\begin{aligned} np^{\rho^{\max}(S,F)} &\geqslant \left(\frac{A \log^{2+2/b} n}{n} \right)^{C/\log n} A \log^{2+2/b} n \\ &= (1+o(1)) e^{-C} A \log^{2+2/b} n. \end{aligned}$$

Since $(np^{\rho^{\max}(S,F)}/t)^t$ increases in t, we find that $\mu \gg \log^2 n$, as claimed.

Finally, we turn to μ^2/Δ. Let $(F_i, i \geqslant 1)$ be an enumeration of all possible ordered (S,F)-extensions of S' in $\mathscr{G}_{n,p}$. We let $T_i \subset F_i$ denote the subgraph induced by the edges of F with at least one endpoint outside of V. Let us write $i \sim j$ if T_i, T_j share at least one edge. Then

$$\Delta = \sum_{i \sim j} \mathbb{P}(T_i, T_j \subset \mathscr{G}_{n,p}) = \mu \sum_{j \sim 1} \mathbb{P}(T_j \subset \mathscr{G}_{n,p} | T_1 \subset \mathscr{G}_{n,p}).$$

Therefore, to show that $\mu^2/\Delta > 2B^2 \log^2 n$, it suffices to verify that

$$\frac{1}{\mu} \sum_{j \sim 1} \mathbb{P}(T_j \subset \mathscr{G}_{n,p} | T_1 \subset \mathscr{G}_{n,p}) < \frac{1}{2B^2 \log^2 n}. \tag{23}$$

By considering F_j with x vertices in common with F_1, we claim that

$$\frac{1}{\mu} \sum_{j \sim 1} \mathbb{P}(T_j \subset \mathscr{G}_{n,p} | T_1 \subset \mathscr{G}_{n,p}) \leqslant \sum_x \frac{\binom{t}{x}\binom{n}{t-x}}{\binom{n-s}{t}} p^{-\lfloor x\rho + xC\rho/\log n \rfloor}. \tag{24}$$

Indeed, this follows since, if F_j, F_1 have x vertices in common, then they can share at most $\lfloor x\rho^{\max}(S,F) \rfloor$ many edges.

Next, since the relevant s and t are of order $\log n$, we note (using $(n-k)^k/k! \leqslant \binom{n}{k} \leqslant n^k/k!$) that

$$\frac{\binom{t}{x}\binom{n}{t-x}}{\binom{n-s}{t}} \leqslant \frac{n^t}{(n-s-t)^t}(t/n)^x\binom{t}{x} \leqslant (1+o(1))\left(\frac{et^2}{xn}\right)^x.$$

In particular, if $x \geqslant 2b+1$, then

$$\frac{\binom{t}{x}\binom{n}{t-x}}{\binom{n-s}{t}} p^{-\lfloor x\rho + xC\rho/\log n\rfloor} \leqslant \left(\frac{O(t^2)}{np^\rho}\right)^x \leqslant \left(\frac{O(1)}{\log^{2/b} n}\right)^x$$
$$= O(\log^{-4} n)\left(\frac{O(1)}{\log^{2/b} n}\right)^{x-2b}.$$

On the other hand, if $x \leqslant 2b$ and $\rho x \notin \mathbb{Z}$, then

$$\frac{\binom{t}{x}\binom{n}{t-x}}{\binom{n-s}{t}} p^{-\lfloor x\rho + xC\rho/\log n\rfloor} \leqslant \left(\frac{O(1)}{\log^{2/b} n}\right)^x n^{-\Theta(x)} = n^{-\Theta(x)}.$$

Also note that

$$\sum_{x\in\{b,2b\}} \frac{\binom{t}{x}\binom{n}{t-x}}{\binom{n-s}{t}} p^{-\lfloor x\rho + xC\rho/\log n\rfloor} \leqslant (1+o(1)) \sum_{x\in\{b,2b\}} \left(\frac{e^{C+1}B^2}{A\log^{2/b} n}\right)^x$$
$$= (1+o(1))\frac{e^{C+1}B^2}{A\log^2 n} < \frac{1}{2B^2\log^2 n},$$

for all large n, provided that $A > 2e^{C+1}B^4$.

Altogether, for any such A, summing over all relevant x in (24), we find that (23) holds for all large n, and this completes the proof. ■

3.6. **The construction.** We are ready to prove Theorem 13.

Proof of Theorem 13. Let $np^\rho = A\log^{2+2/b} n$, where A is sufficiently large. Then, by Lemma 24, with high probability, for *every* $S \subset F$ such that

- $V(F)$ is ordered,
- $v(F) \leqslant 2e(H)\log n$, and
- (S,F) is α_*-sparse,

every subgraph $S' \subset \mathscr{G}_{n,p}$ with $v(S') = v(S)$ has an (S,F)-extension in $\mathscr{G}_{n,p}$. By Lemma 22 and (21), for all large n, we have $v(M_f^*) \leqslant 2e(H)\log n$ for all red edges $f \in \mathscr{R}^*$ of W^*. Also, by Lemma 19, note that (C_f^*, M_f^*) is α_*-sparse for all $f \in \mathscr{R}^*$. For the rest of this proof, we will assume that these statements hold. Using these, we will show how to activate an arbitrary edge g missing from $\mathscr{G}_{n,p}$.

Base step. Consider the modified α_*-core $M^* = M^*_{e^*}$ of $W^* = W^*_{e^*}$. Since $M^* \subseteq W^* \subseteq A^*$, it follows by (16) that

$$\rho^{\max}(e^*, M^* \cup e^*) \leqslant \rho^{\max}(e^*, A^* \cup e^*) < \rho + 1/\log n < \alpha_*.$$

Hence $(e^*, M^* \cup e^*)$ is α_*-sparse, and so, there is an $(e^*, M^* \cup e^*)$-extension of g in $\mathscr{G}_{n,p}$. In other words, we can embed a copy of M^* in $\mathscr{G}_{n,p}$ based at g. We call g the *copy* of e^* in $\mathscr{G}_{n,p}$, and say that g has been *processed*.

Next step. Recall that M^* is associated with a copy $H^* = H^*_{e^*}$ of H that e^* completes. Specifically,

$$M^* = \bigcup_{f \in E(H^* \setminus e^*)} C^*_f.$$

Let $f_1, \ldots, f_k$ be the red edges in $H^* \setminus e^*$ (for which $C^*_f \neq f$). Each f_i has an *unprocessed* copy g_i in $\mathscr{G}_{n,p}$, and a copy C_{g_i} of $C^*_{f_i}$ embedded in $\mathscr{G}_{n,p}$ based at g_i. Therefore, since $(C^*_{f_i}, M^*_{f_i})$ is α_*-sparse, there is an $(C^*_{f_i}, M^*_{f_i})$-extension of C_{g_i} in $\mathscr{G}_{n,p}$. We process the edges $g_1, \ldots, g_k$ in turn, one at a time, embedding such extensions in $\mathscr{G}_{n,p}$. In each step, for each red edge f in $H^*_{f_i}$, an unprocessed copy g' in $\mathscr{G}_{n,p}$ is created.

Let us stress that red edges in $\mathscr{R}^*$ can have multiple unprocessed copies in $\mathscr{G}_{n,p}$, as a single red edge may appear in different $H^*_{f_i}$; see Figure 4 for an illustrative example.

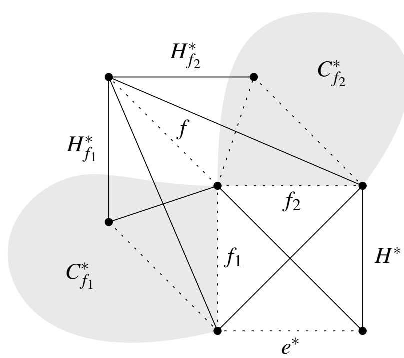

FIGURE 4. There are two red edges f_1, f_2 in the copy H_* of H that e_* completes in W^*. There is a red edge f in both $H^*_{f_i}$ that is not contained in either of the α_*-cores $C^*_{f_i}$. In the next step of the construction, $(C^*_{f_i}, M^*_{f_i})$-extensions of $C^*_{f_i}$ will be added to $\mathscr{G}_{n,p}$. Since both of these extensions involve the red edge f, these extensions can place copies of f at different locations in $\mathscr{G}_{n,p}$. Each copy will need to be processed separately later on in the construction.

Inductive step. We continue recursively, until there is no unprocessed copy of any $f \in \mathscr{R}^*$ remaining in $\mathscr{G}_{n,p}$. (Note that if M^*_f is a black copy of H minus an edge in W^*, then processing a copy of f in $\mathscr{G}_{n,p}$ does not create any further edges to process.) In this way, by taking the union of all the extensions embedded during this recursive procedure, we obtain some subgraph $A_g \subseteq \mathscr{G}_{n,p}$.

Activation. Finally, we claim that $(g, A_g) \in \mathscr{A}(H)$, which completes the proof of the theorem. Indeed, by induction, we claim that every copy of each red edge $f \in \mathscr{R}^*$ in $\mathscr{G}_{n,p}$ is activated.

For the base case, we first consider those f for which all edges in $H^*_f \setminus f$ are black (i.e., the red edges that are activated at time $t = 1$). Trivially, every copy of such an f in $\mathscr{G}_{n,p}$ is activated by the embedding of $H^*_f \setminus f$ in $\mathscr{G}_{n,p}$, since clearly $M^*_f = H^*_f \setminus f$.

On the other hand, for the other red edges f that are activated at some time $t > 1$, we recall that f completes some copy H^*_f of H. By construction, for each copy g' of such an f in $\mathscr{G}_{n,p}$, an extension M^*_f of f has been added to $\mathscr{G}_{n,p}$, which contains a copy of H^*_f based at g'. By induction, the copy in $\mathscr{G}_{n,p}$ of each red edge in $H^*_f \setminus f$ is activated. Hence, g' is also activated. Finally, recalling that e^* is the final edge red edge in $\mathscr{R}^*$ to be activated, and since g is the copy of e^* in $\mathscr{G}_{n,p}$, we see that g is activated. ■

4. HITTING TIME RESULTS

In this section, to further demonstrate the utility of formula (2) for $\rho(H)$, we discuss hitting time results for all H with $\rho(H) \leqslant 1$. We will also prove such results for all H with a *leaf* (i.e., a vertex of degree 1). The main results are stated as Theorems 25 and 28 below.

We consider the usual random graph process $(\mathscr{G}_m, 0 \leqslant m \leqslant \binom{n}{2})$, where $\mathscr{G}_{m+1}$ is obtained from $\mathscr{G}_m$ by adding a uniformly random edge (from amongst those missing from $\mathscr{G}_m$). In discussing the *hitting time* for some graph property, we mean the first time that $\mathscr{G}_m$ has this property.

4.1. Graphs H with a leaf.

We recall that in [8, Proposition 26], for graphs H with a leaf, it is shown that

$$p_c(n,H) = \Theta(n^{-1/\beta(H)}), \tag{25}$$

where

$$\beta(H) = \min_{e \in E(H)} \rho^{\max}(\emptyset, H \setminus e) = \min_{e \in E(H)} \max_{\emptyset \subset F \subset H \setminus e} \frac{e(F)}{v(F)}. \tag{26}$$

In particular, this gives a large family of graphs H for which $p_c(n,H) = O(n^{-1/\rho(H)})$ holds, as in Question 7 above.

Let $\mathscr{H}_\beta^-$ be the set of all graphs $H^- = H \setminus e$ obtained by removing an edge e from H for which $\rho^{\max}(\emptyset, H^-) = \beta(H)$ is minimal.

Theorem 25. *Suppose that H has a leaf. Then, with high probability, the hitting time for H-percolation in the random graph process $(\mathscr{G}_m, 0 \leqslant m \leqslant \binom{n}{2})$ coincides with the first time that $\mathscr{G}_m$ contains a copy of some $H^- \in \mathscr{H}_\beta^-$.*

The main idea is to note that the first copy H^- of H minus an edge to emerge in the random graph process will, with high probability, belong to $\mathscr{H}_\beta^-$. In the following proof, we consider some vertex x that is neighbors with a leaf, and take cases with respect to whether H^- has subgraphs F with density $e(F)/v(F) = \beta(H)$ that contain x.

Proof. Suppose that H has a leaf.

Recall that $\mathscr{H}_\beta^-$ is the set of all $H^- = H \setminus e$, with $e \in E(H)$, such that $\rho^{\max}(\emptyset, H^-)$ attains the minimal possible value $\beta = \beta(H)$, as in (26) above. We also let $\mathscr{H}^-$ denote the set of all $H^- = H \setminus e$, with $e \in E(H)$. Recall (see, e.g., [20, Theorem 3.4]) that the hitting time τ^-, in the random graph process $(\mathscr{G}_m, 0 \leqslant m \leqslant \binom{n}{2})$, for the appearance of a copy of some graph in $\mathscr{H}^-$ coincides with the hitting time τ_β^- for the appearance of a copy of some graph in $\mathscr{H}_\beta^- \subseteq \mathscr{H}^-$. Of course, if $G \neq \langle G \rangle_H$, then G must contain at least one copy of some $H^- \in \mathscr{H}^-$. As such, it remains to prove that, with high probability, $\langle \mathscr{G}_{\tau^-} \rangle_H = K_n$. To this end, we claim that $\langle G \rangle_H = K_n$ for *any* graph $G \subseteq K_n$ that contains:

- some H_β^- that is a copy of a graph in $\mathscr{H}_\beta^-$;
- at least $2v(H)$ disjoint copies of every graph F on at most $v(H)$ vertices with $\rho^{\max}(\emptyset, F) < \beta$;
- at least $2v(H)$ disjoint (H_1, H_2)-extensions of every set of vertices U of size $v(H_1)$ for every β-sparse pair (H_1, H_2) with $v(H_2) \leqslant v(H)$.

We note that, with high probability, $\mathscr{G}_{\tau^-}$ has all of these properties; see [20, Remark 3.7], [26, Theorem 3], and Lemma 24 above. Hence the claim implies the theorem.

In proving the claim, we take three cases. Let x be the neighbor of some leaf in H. In what follows, we say that a subgraph $F \subseteq H_\beta^-$ is β-dense if $e(F)/v(F) = \beta$.

Case 1. Suppose that no β-dense subgraph $F \subseteq H_\beta^-$ contains x. Let H^* be an inclusion-maximal β-dense subgraph of H_β^-. Along similar lines as Lemma 19 above, it can be seen that the pair (H^*, H_β^-) is β-sparse. Hence, we can find $v(H) - 2$ disjoint (H^*, H_β^-)-extensions of H^*. Let $x_1, \ldots, x_{v(H)-2}$ be the copies of x in these extensions. To see that $\langle G \rangle_H = K_n$, note that we can activate:

- the missing edge e in H_β^-;
- the copy of e in each (H^*, H_β^-)-extension of H^*, if the endpoints of e are not both in H^* (otherwise there is no such copy);
- all edges that contain one of the vertices $x_1, \ldots, x_{v(H)-2}$;
- and then finally any other missing edge $\{u,v\}$, by completing a copy of H with vertices $u, v, x_1, \ldots, x_{v(H)-2}$.

In the next two cases, we assume there is a β-dense subgraph of H_β^- that contains x, and let H_x^* be an inclusion-minimal subgraph with these properties.

Case 2. Suppose that all β-dense subgraphs of H_β^- contain x. Then all such subgraphs contain H_x^*. Hence, for any vertex $z \neq x$ in H_x^*, we have that $\rho^{\max}(\emptyset, H_\beta^- \setminus z) < \beta$. Therefore, we can find $v(H)-2$ disjoint copies of $H_\beta^- \setminus z$ in G that are also disjoint from H_β^-. Let $x_1, \ldots, x_{v(H)-2}$ be the copies of x in these subgraphs. To see that $\langle G \rangle_H = K_n$, we activate:

- the missing edge e in H_β^-;
- all edges that contain x (and note that x may now play the role of the missing vertex z in the $v(H)-2$ copies of $H_\beta^- \setminus z$);
- the copy of e in every extended (with x playing the role of z) copy of $H_\beta^- \setminus z$, if e is not incident to x;
- all edges that contain one of the vertices $x_1, \ldots, x_{v(H)-2}$;
- and then finally any missing edge $\{u,v\}$ by completing a copy of H with vertices $u, v, x_1, \ldots, x_{v(H)-2}$.

Case 3. Finally, we assume that, while H_β^- has a β-dense subgraph that contains x, not all such subgraphs contain x. In this case, we will use the fact that an intersection of any β-dense subgraph $F \subseteq H_\beta^-$ with H_x^* is either empty or β-dense. This follows by the next claim (taking $A = H_\beta^-$, $A_1 = F$ and $A_2 = H_x^*$).

Claim 26. *Suppose that $\rho^{\max}(\emptyset, A) = \beta$. Let A_1, A_2 be β-dense subgraphs of A such that $A_1 \cap A_2 \neq \emptyset$. Then $A_1 \cap A_2$ is β-dense.*

Proof. Since A_1, A_2 are β-dense, it follows that

$$e(A_1 \cup A_2) \geqslant \beta(v(A_1) + v(A_2)) - e(A_1 \cap A_2).$$

Therefore, since $A_1 \cup A_2$ is β-dense, it follows that $e(A_1 \cap A_2) = \beta(v(A_1 \cap A_2))$, and so $A_1 \cap A_2$ is β-dense, as claimed. ■

If there is no β-dense subgraph of H_β^- that does not contain x and have a non-empty intersection with H_x^*, then we put $H^* = \emptyset$. Otherwise, let H^* be an inclusion-maximal β-dense subgraph of H_x^* that does not contain x. Let H_-^* be an inclusion-minimal β-dense subgraph of H^* (where $H_-^* = \emptyset$

if $H^* = \emptyset$). Also, let H^*_+ be an inclusion-maximal β-dense subgraph of $H^* \setminus H^*_-$, if such a subgraph exists; otherwise, if $\rho^{\max}(\emptyset, H^* \setminus H^*_-) < \beta$, we put $H^*_+ = \emptyset$.

Next, we let Y be the (possibly empty) union of all β-dense subgraphs of H^-_β that do not contain x and are vertex-disjoint with H^*. Note that $Y \cap H^*_x = \emptyset$, and in particular $x \notin Y$. There are no edges between Y and H^*_x, as otherwise the density of $Y \cup H^*_x$ would be strictly greater than β.

We claim that every β-dense subgraph of H^-_β that contains $Y \cup H^*_+$ as a *proper* subgraph must contain all of H^*_-. Indeed, if such a subgraph:

- contains x but not all of H^*_-, then this would contradict the minimality of H^*_x;
- does not contain x, and some but not all of H^*_-, then this would contradict the minimality of H^*_-;
- does not contain x, or any of H^*_-, then this would contradict the maximality of H^*_+ (as then, by the choice of Y and since $Y \cup H^*_+$ is a proper subset, it would contain some other vertex $H^* \setminus H^*_-$ that is not already in H^*_+).

To conclude, take any vertex z in H^*_- if $H^* \neq \emptyset$, and take any $z \neq x$ in H^*_x otherwise. The pair $(Y \cup H^*_+, H^-_\beta \setminus z)$ is β-sparse, so we can find $v(H) - 2$ disjoint $(Y \cup H^*_+, H^-_\beta \setminus z)$-extensions of $Y \cup H^*_+$ in G that also disjoint from H^-. Since z does not have neighbors in $Y \cup H^*_+$, each such extension can be extended to a copy of H^- by replacing z with x, as soon as all edges between x and vertices outside of H^- have been activated. Hence, using the copies $x_1, \ldots, x_{v(H)-2}$ of x in these $v(H) - 2$ resulting copies of H^-, the remaining edges of K_n can be activated in the same way as before.

The proof is complete. ■

4.2. **A characterization.** Our next result characterizes all graphs H with $\rho(H) \leqslant 1$. Recall (as discussed in Section 1) that we always assume that H has no isolated vertices.

Lemma 27. *If H has at least two cycles then $\rho(H) \geq 1$. Moreover, $\rho(H) = 1$ if and only if H*

- *is a cycle, or*
- *has at least two cycles and some edge e for which each connected component of $H \setminus e$ has at most one cycle.*

Proof. If H has at least two cycles, then no edge of H is in all of its cycles. Hence, for any $(e, A) \in \mathscr{A}(H)$, there is a cycle in A, and so $\rho(H) \geqslant 1$.

Next, let us assume $\rho(H) = 1$ and that H is not a cycle.

Towards a contradiction, let us assume that H has only one cycle but is not itself a cycle. Then there is some vertex x that is neighbors with a leaf in H. To obtain a contradiction with our assumption that $\rho(H) = 1$, we will construct a pair $(f, A) \in \mathscr{A}(H)$ for which $\rho^{\max}(f, A \cup f) < 1$ as follows: Obtain H^- from H by deleting an edge f in its cycle. We obtain A by taking a disjoint union of copies $H_1^-, \dots, H_{v(H)-2}^-$ of H^- together with two isolated vertices y, z. Let f be the missing edge $\{y, z\}$. Clearly, $\rho^{\max}(f, A \cup f) < 1$. To see that A activates f, we first activate the copies f_i of the missing edge f in each of the H_i^-. Let x_i be the copies of x in the H_i^-, and note that, by the choice of x, every missing edge containing some x_i can be activated. In particular, all edges of the clique on $x_1, \dots, x_{v(H)-2}, y, z$ except f can be activated in this way; and then f can also be activated.

Towards another contradiction, suppose that H does not have an edge whose deletion leaves at most one cycle in every connected component. Then, for every $(f, A) \in \mathscr{A}(H)$, there exists a connected subgraph $F \subseteq A$ on at most $v(H)$ vertices with at least two cycles, and hence $\rho^{\max}(f, A \cup f) \geqslant 1 + 1/v(H)$, in contradiction to out assumption that $\rho(H) = 1$.

Finally, let us assume that H is either a cycle or has at least two cycles and an edge e whose deletion leaves at most one cycle in every connected component. It remains only to prove that $\rho(H) = 1$.

Case 1. The case that $H = C_m$ is a cycle is the simplest. For instance, as noted in [8, Proposition 24], if we let A be the union of a triangle and a path of length at least m that share exactly one vertex, then $\langle A \rangle_H$ is the clique on $V(A)$. (Other constructions are possible, such as a long path and an extra edge between a pair of vertices at distance $m - 1$.) In particular, the edge f between the endpoints of the path is activated. Since the path can be arbitrarily long, this shows that $\rho(C_m) = 1$.

Case 2. In the second case, H is a disjoint union of a graph H_2 that has at most one cycle in each of its connected components, and a connected graph H_1 that contains either:

- a single cycle,
- two cycles that share exactly one vertex,
- two disjoint cycles connected by a path, or
- three paths that share their endpoints (i.e., two cycles that share a path).

Case 2a. If H_1 is a cycle then we can argue along the same lines as in Case 1, taking A to be the disjoint union of a copy of H_2 together with a triangle attached to an arbitrarily long path, as before. If H_1 contains a single cycle, we proceed in the same way, but attach large enough perfect trees to all vertices along the path and triangle.

In the remaining cases, for simplicity, let us assume that $H = H_1$, since otherwise, in all of our constructions of A, we can simply add a disjoint copy of H_2 to A. Likewise, let us assume that the minimum degree $\delta(H) = 2$, since otherwise we can simply add perfect trees to all vertices in A.

Case 2b. Suppose that H is the union of two cycles, of lengths ℓ and m, that share a single vertex x. Then to obtain A, take a copy of C_m and attach an arbitrarily long path of length $(\ell-1)+k(\ell-2)$, for some $k \geqslant 0$. By induction, the edge f between the endpoints of the path is activated. Taking k arbitrarily large, we see that $\rho(H) = 1$ in this case.

Case 2c. If H is the union of two cycles, of lengths ℓ and m, that are connected by a path of length p, the argument is essentially the same as in Case 2b. We obtain A, we take a copy of C_m, attach a path of length p from C_m to some other vertex x, and then attach a path of length $(\ell-1)+k(\ell-2)$ to x. By induction, the edge f from x to the end of the path is activated.

Case 2d. Finally, suppose that H is the union of three paths, of lengths ℓ_1, ℓ_2 and ℓ_3, that share their endpoints. To construct A, we start with two paths of lengths ℓ_1 and ℓ_2 that share their endpoints x and y. Next, we attach a path P_x to x of length $k(\ell_2+\ell_3-2)+(\ell_2-1)$, and a path P_y to y of length $2k(\ell_1-1)$. These paths are otherwise disjoint with the two paths between x and y, and are disjoint with each other. Let f be the missing edge between the endpoints between these paths. Note that the edge between y and the vertex along P_x at distance ℓ_3-1 from x can be activated. Thereafter, the edge between x and the vertex along P_y at distance ℓ_1-1 from y can be activated. In this way, by induction, it can be seen that f is eventually activated. ■

4.3. **Graphs H with $\rho < 1$.** It follows immediately by Lemma 27 that $\rho(H) < 1$ if and only if H has at most one cycle and a leaf. For such a graph H and an edge $e \in E(H)$, let $f(e,H)$ be the maximum number of edges in a tree in $H \setminus e$. Let $f(H) = \min_e f(e,H)$, minimizing over all $e \in E(H)$ for which $H \setminus e$ is a forest. Then, it can be seen that $p_c(n,H) = \Theta(n^{-(f+1)/f})$. Indeed, when $p \ll n^{-(f+1)/f}$, with high probability $\mathscr{G}_{n,p}$ has no trees of size f (see, e.g., [20, Theorem 3.4]), and therefore no edge can be activated. On the other hand, when $p \gg n^{-(f+1)/f}$, with high probability $\mathscr{G}_{n,p}$ has many copies of any tree with at most f edges (see, e.g., [20, Remark 3.7]), and this is enough to activate all of K_n.

In fact, by Theorem 25 we obtain the following hitting time result. Let $\mathscr{F}_H$ be the family of all forests F with the following property: For each $F \in \mathscr{F}_H$, there is some $e \in E(H)$ such that F consists of trees of size $f(H)$ in the forest $H \setminus e$. Then, by Theorem 25, the hitting time for H-percolation coincides with the appearance of a copy of some forest $F \in \mathscr{F}_H$.

4.4. **Graphs H with $\rho = 1$.** We note that, if $\rho(H) = 1$ and H has a leaf, then the hitting time result Theorem 25 implies that the H-percolation threshold is coarse. Our next result shows that the threshold is sharp for all other H with $\rho(H) = 1$.

Theorem 28. *Let H be a graph with $\rho(H) = 1$ and minimum degree $\delta(H) = 2$. Then, with high probability, the hitting time in the random graph process $(\mathscr{G}_m, 0 \leqslant m \leqslant \binom{n}{2})$ for H-percolation coincides with that of connectivity.*

We note that this result extends the result in [8, Proposition 24] for the cases $H = C_k$, $k \geqslant 3$, and $H = K_{2,3}$.

Proof. If H has minimum degree $\delta(H) = 2$ then, according to Lemma 27, H consists of connected components $H_1, \ldots, H_k$, where $H_2, \ldots, H_k$ are cycles, and H_1 is either a cycle, or a union of two cycles that are either connected by a path or share a path (see the proof of Lemma 27).

Let $\mathscr{C}$ be the property of being connected. Recall that, with high probability, the hitting time $\tau_{\mathscr{C}}$ for connectivity coincides with that of the disappearance of isolated vertices, and $\tau_{\mathscr{C}} = \frac{\ln n + O_p(1)}{n} \cdot \binom{n}{2}$; see [14, Theorems 7.4, 7.7]. Since $\delta(H) = 2$, note that $\langle G \rangle_H \neq K_n$ for any $G \subset K$ with an isolated vertex. Therefore, it suffices to find a graph property $\mathscr{P}$ such that:

- $\mathscr{C} \cap \mathscr{P}$ is increasing,
- with high probability $\tau_{\mathscr{C}} = \tau_{\mathscr{C} \cap \mathscr{P}}$, and
- deterministically any graph with $\mathscr{C} \cap \mathscr{P}$ is H-percolating.

Recall that, in the proof of Lemma 27, we constructed a pair $(f, A) \in \mathscr{A}(H)$, where A is a union of a cycle and one or two long paths, and possibly also a disjoint copy of some graph H_2 that plays no essential role (and note that perfect trees are note required in constructing A here, as we currently assume that $\delta(H) = 2$.) We let $\mathscr{P}$ be the property that every pair of non-isolated vertices has an (f, A)-extension, for some fixed $v(A) = \Theta(\log n)$.

To show that with high probability $\tau_{\mathscr{C}} = \tau_{\mathscr{C} \cap \mathscr{P}}$, consider $\mathscr{G}_{n,p}$ with let $p = (1 \pm o(1)) \log n / n$. With high probability, every non-isolated vertex in $\mathscr{G}_{n,p}$ has a neighor of degree at least $0.5 \log n$. Take a pair of vertices v_1, v_2, and reveal their neighborhoods. If at least one of them is isolated, there is nothing to prove, so assume both vertices have neighbors u_1, u_2 with degrees at least $0.5 \log n$. Let X be the number of (f, A)-extensions of $\{v_1, v_2\}$. The expected number of such extensions equals $\Theta(n^{v(A)} p^{v(A)+1}) = n^{\omega(1)}$. By Janson's inequality, $\mathbb{P}(X = 0) \leqslant \exp(-\mu^2/(2\Delta))$, where

$$\Delta \leqslant \max\{\mu, 2\mu^2/np \log n, \mu^2/(np)^3\} \leqslant \mu^2/0.4 \log^2 n.$$

Therefore, $\mathbb{P}(X = 0) \leqslant e^{-0.4 \log^2 n} = n^{-\omega(1)}$, and so by taking a union bound over all pairs $\{v_1, v_2\}$ we complete the proof. ∎

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, *An extremal problem for sets with applications to graph theory*, J. Combin. Theory Ser. A **40** (1985), no. 1, 82–89.
3. N. Alon and J. H. Spencer, *The probabilistic method*, fourth ed., Wiley Series in Discrete Mathematics and Optimization, John Wiley & Sons, Inc., Hoboken, NJ, 2016.
4. O. Angel and B. Kolesnik, *Sharp thresholds for contagious sets in random graphs*, Ann. Appl. Probab. **28** (2018), no. 2, 1052–1098.
5. ______, *Large deviations for subcritical bootstrap percolation on the Erdős–Rényi graph*, J. Stat. Phys. **185** (2021), no. 2, Paper No. 8, 16.
6. R. Ascoli and X. He, *Rational values of the weak saturation limit*, preprint (2026), available at arXiv:2501.15686.
7. P. Balister, B. Bollobás, R. Morris, and P. Smith, *Universality for monotone cellular automata*, J. Eur. Math. Soc. (JEMS), to appear, available at arXiv:2203.13806.
8. J. Balogh, B. Bollobás, and R. Morris, *Graph bootstrap percolation*, Random Structures Algorithms **41** (2012), no. 4, 413–440.
9. Z. Bartha and B. Kolesnik, *Weakly saturated random graphs*, Random Structures Algorithms **65** (2024), no. 1, 131–148.
10. Z. Bartha, B. Kolesnik, and G. Kronenberg, *H-percolation with a random H*, Electron. Commun. Probab. **29** (2024), Paper No. 33, 5.
11. Z. Bartha, B. Kolesnik, G. Kronenberg, and Y. Peled, *Sharp Fuss–Catalan thresholds in graph bootstrap percolation*, preprint (2025), available at arXiv:2510.26724.
12. M. Bidgoli, A. Mohammadian, and B. Tayfeh-Rezaie, *On $K_{2,t}$-bootstrap percolation*, Graphs Combin. **37** (2021), no. 3, 731–741.
13. B. Bollobás, *Weakly k-saturated graphs*, Beiträge zur Graphentheorie (Kolloquium, Manebach, 1967), Teubner, Leipzig, 1968, pp. 25–31.
14. ______, *Random graphs*, second ed., Cambridge Studies in Advanced Mathematics, vol. 73, Cambridge University Press, Cambridge, 2001.
15. J. Chalupa, P. L. Leath, and G. R. Reich, *Bootstrap percolation on a bethe lattice*, J. Phys. C **21** (1979), L31–L35.
16. P. Erdős, A. Hajnal, and J. W. Moon, *A problem in graph theory*, Amer. Math. Monthly **71** (1964), 1107–1110.
17. P. Frankl, *An extremal problem for two families of sets*, European J. Combin. **3** (1982), no. 2, 125–127.
18. E. Friedgut, *Sharp thresholds of graph properties, and the k-sat problem*, J. Amer. Math. Soc. **12** (1999), no. 4, 1017–1054, With an appendix by Jean Bourgain.
19. M. Gardner, *Mathematical games: The fantastic combinations of John Conway's new solitaire game "life"*, Sci. Am. **223** (1970), no. 4, 120–123.
20. S. Janson, T. Łuczak, and A. Rucinski, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000.
21. G. Kalai, *Weakly saturated graphs are rigid*, Convexity and graph theory (Jerusalem, 1981), North-Holland Math. Stud., vol. 87, North-Holland, Amsterdam, 1984, pp. 189–190.
22. ______, *Hyperconnectivity of graphs*, Graphs Combin. **1** (1985), no. 1, 65–79.
23. B. Kolesnik, *The sharp K_4-percolation threshold on the Erdős–Rényi random graph*, Electron. J. Probab. **27** (2022), Paper No. 13, 23.

24. L. Lovász, *Flats in matroids and geometric graphs*, Combinatorial surveys (Proc. Sixth British Combinatorial Conf., Royal Holloway Coll., Egham, 1977), Academic Press, London-New York, 1977, pp. 45–86.
25. R. Morris, *Monotone cellular automata*, Surveys in combinatorics 2017, London Math. Soc. Lecture Note Ser., vol. 440, Cambridge Univ. Press, Cambridge, 2017, pp. 312–371.
26. J. Spencer, *Counting extensions*, J. Combin. Theory Ser. A **55** (1990), no. 2, 247–255.
27. N. Terekhov and M. Zhukovskii, *Weak saturation in graphs: a combinatorial approach*, J. Combin. Theory Ser. B **172** (2025), 146–167.
28. ______, *Weak saturation rank: a failure of the linear algebraic approach to weak saturation*, Combinatorica **45** (2025), no. 5, Paper No. 45, 18.
29. S. Ulam, *Random processes and transformations*, Proceedings of the International Congress of Mathematicians, Vol. 2, Cambridge, Mass., 1950, pp. 264–275.
30. J. von Neumann, *Theory of self-reproducing automata*, University of Illinois Press, Champaign, IL, USA, 1966.

UNIVERSITY OF WARWICK, DEPARTMENT OF STATISTICS
Email address: brett.kolesnik@warwick.ac.uk

LMU MUNICH, DEPARTMENT OF MATHEMATICS
Email address: makai@math.lmu.de

UNIVERSITY OF CANTERBURY, SCHOOL OF MATHEMATICS AND STATISTICS
Email address: rajko.nenadov@canterbury.ac.nz

UNIVERSITY OF NEBRASKA–LINCOLN, DEPARTMENT OF MATHEMATICS
Email address: xperez@unl.edu

TORONTO METROPOLITAN UNIVERSITY, DEPARTMENT OF MATHEMATICS
Email address: pralat@torontomu.ca

UNIVERSITY OF SHEFFIELD, SCHOOL OF COMPUTER SCIENCE
Email address: m.zhukovskii@sheffield.ac.uk