Electron. Commun. Probab. **29** (2024), article no. 33, 1–5.
https://doi.org/10.1214/24-ECP602
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

H-percolation with a random H*

Zsolt Bartha† Brett Kolesnik‡ Gal Kronenberg§

Abstract

In H-percolation, we start with an Erdős–Rényi graph $\mathcal{G}_{n,p}$ and then iteratively add edges that complete copies of H. The process percolates if all edges missing from $\mathcal{G}_{n,p}$ are eventually added. We find the critical threshold p_c when $H = \mathcal{G}_{k,1/2}$ is uniformly random, solving a problem of Balogh, Bollobás and Morris. In this sense, we find p_c for most graphs H.

Keywords: bootstrap percolation; cellular automaton; critical threshold; phase transition; random graph; weak saturation.
MSC2020 subject classifications: 05C35; 05C80; 05C99; 60K35; 68Q80.
Submitted to ECP on January 10, 2024, final version accepted on June 25, 2024.

1 Introduction

Following Balogh, Bollobás and Morris [1], we fix a graph H, and begin with an Erdős–Rényi graph $\mathcal{G}_0 = \mathcal{G}_{n,p}$. Then, for $t \geqslant 1$, we obtain $\mathcal{G}_t$ from $\mathcal{G}_{t-1}$ by adding every edge that creates a new copy of H. We let $\langle \mathcal{G}_{n,p} \rangle_H = \bigcup_{t \geqslant 0} \mathcal{G}_t$ denote the graph containing all eventually added edges. When $\langle \mathcal{G}_{n,p} \rangle_H = K_n$, we say that the process *H-percolates,* or equivalently, in the terminology of Bollobás [4], that $\mathcal{G}_{n,p}$ is *weakly H-saturated*.

H-percolation generalizes bootstrap percolation [7, 5], which is one of the most well-studied of all cellular automata [8, 9]. Such processes model evolving networks, in which sites update their status according to the behavior of their neighbors. Although the dynamics are local, they can lead to global behavior, emulating various real-world phenomena of interest, such as tipping points, super-spreading, self-organization and collective decision-making.

The *critical H-percolation threshold* is defined to be the point

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

at which $\mathcal{G}_{n,p}$ becomes likely to H-percolate. Problem 1 in [1] asks for

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

*ZB is supported by the ERC Consolidator Grant 772466 “NOISE”. GK is supported by the Royal Commission for the Exhibition of 1851. BK is supported by a Florence Nightingale Bicentennial Fellowship (Oxford Statistics) and a Senior Demyship (Magdalen College).
†Alfréd Rényi Institute of Mathematics. E-mail: bartha@renyi.hu
‡University of Oxford, Department of Statistics. E-mail: brett.kolesnik@stats.ox.ac.uk
§University of Oxford, Department of Mathematics. E-mail: gal.kronenberg@maths.ox.ac.uk

for *every* graph H. Finding ℓ corresponds to locating $p_c = n^{-\ell+o(1)}$. The authors state that "Problem 1 is likely to be hard." For instance, even $\ell(K_{2,t})$ remains open for $t \geqslant 5$; see Bidgoli et al. [3].

In this note, we answer Problem 1 for *random* graphs H. We find that typically $\ell(H) = 1/\lambda(H)$, where

$$\lambda(H) = \frac{e_H - 2}{v_H - 2}$$

is the adjusted edge per vertex ratio in [1] for graphs H with v_H vertices and e_H edges. We recall that Theorem 1 in [1] shows that $\ell(K_k) = 1/\lambda(K_k)$ for cliques. This corresponds to the case $\alpha = 1$ in the following result, which shows that $\ell = 1/\lambda$ is "stable."

Theorem 1.1. *Fix $0 < \alpha \leqslant 1$. With high probability, $\ell(\mathcal{G}_{k,\alpha}) = 1/\lambda(\mathcal{G}_{k,\alpha})$ as $k \to \infty$.*

The case $\alpha = 1/2$ answers Problem 6 in [1], that asks for "bounds on $p_c(n, \mathcal{G}_{k,1/2})$ which hold with high probability as $k \to \infty$." We find that $p_c = n^{-1/\lambda+o(1)}$, except with exponentially small probability. By the Borel–Cantelli lemma, this holds almost surely for all large k. Note that $\mathcal{G}_{k,1/2}$ is uniformly random over simple labelled graphs. In this sense, we find $\ell(H)$ for "most" graphs H.

In fact, our proof works for $\alpha \geqslant A(\log k)/k$, for some sufficiently large constant A. See Section 3.1 below for more on this.

In closing, we note that it is natural to ask about k growing with n. For instance, $\lambda \approx k/4$ when $\alpha = 1/2$, in which case $k = o(\log n)$ appears to be the region of interest.

2 Background

As defined[1] in [1], a graph H is *balanced* if

$$\frac{e_F - 1}{v_F - 2} \leqslant \lambda(H), \tag{2.1}$$

for all subgraphs $F \subset H$ with $3 \leqslant v_F < v_H$. Otherwise, we call H *unbalanced*. In [1], it is shown that $\ell \geqslant 1/\lambda$ for all balanced H. A sharper upper bound (on p_c) is proved in [2] for *strictly balanced* graphs H, which satisfy the above condition, with $\leqslant$ replaced by $<$. Specifically, it is shown that $p_c = O(n^{-1/\lambda})$ for all such H, replacing $n^{o(1)}$ with a constant.

On the other hand, in [2] it is shown that $\ell \leqslant 1/\lambda_*$ for all H, with $v_H \geqslant 4$ and minimum degree $\delta_H \geqslant 2$, where

$$\lambda_*(H) = \min \frac{e_H - e_F - 1}{v_H - v_F},$$

minimizing over all subgraphs $F \subset H$ with $2 \leqslant v_F < v_H$. The quantity λ_* is related to the "cost" of adding an edge, via the H-percolation dynamics, as depicted in Figure 1.

As observed in [2], we have that $\lambda_* \leqslant \lambda$ in general, and $\lambda_* = \lambda$ if and only if H is balanced. Moreover, when H is balanced, single edges F (when $v_F = 2$) attain the minimum λ_*. When H is strictly balanced, these are the only minimizers.

3 Proof

To prove Theorem 1.1, we show that $\mathcal{G}_{k,\alpha}$ is (strictly) balanced, with high probability. By the results from [1, 2] discussed above, the result follows.

Let us give some intuition for why a random graph $\mathcal{G}_{k,\alpha}$, with sufficiently large edge probability α, is likely to be balanced. If (2.1) fails for some $F \subset \mathcal{G}_{k,\alpha}$, then the edge density of F is larger than that of $\mathcal{G}_{k,\alpha}$. For instance, if F includes all but j vertices,

[1] In [1], balanced graphs are assumed to have $e_H \geqslant 2v_H - 2$. In fact, this assumption is only needed in the proof of their Lemma 6. In [2, Lemma 16], a stronger result is proved, without this assumption.

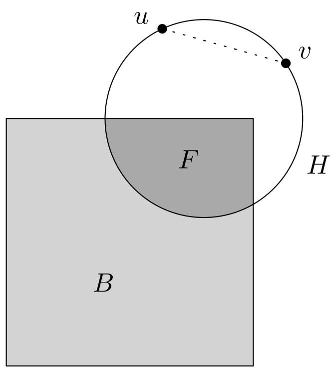

Figure 1: Suppose that edges in a "base" graph B have already been added. Next, we add an edge $\{u, v\}$ using a copy of H. The "price" is $e_H - e_F - 1$ edges and $v_H - v_F$ vertices, where $F = H \cap B$. Hence the edge per vertex "cost" is at least λ_*.

then the number of edges between these j vertices and F is less than (roughly) j times half the average degree in F. By the law of large numbers (when k is large) this large deviation event is most likely to occur when $j = 1$, and even then, it is quite unlikely.

We will use the following, standard Chernoff tail estimates. Recall that if X is binomial with mean μ then, for $0 < \delta < 1$,

$$\mathbb{P}(|X/\mu - 1| \geqslant \delta) \leqslant 2\exp[-\delta^2\mu/3]. \tag{3.1}$$

On the other hand, for $\delta \geqslant 1$,

$$\mathbb{P}(X \geqslant (1+\delta)\mu) \leqslant \exp[-\delta\mu/3]. \tag{3.2}$$

Proposition 3.1. *Suppose that $A(\log k)/k \leqslant \alpha \leqslant 1$, for some sufficiently large constant A. Then, almost surely, $\mathcal{G}_{k,\alpha}$ is strictly balanced for all large k. In particular, $\ell(\mathcal{G}_{k,\alpha}) = 1/\lambda(\mathcal{G}_{k,\alpha})$ for all large k.*

Proof. If $\mathcal{G}_{k,\alpha}$ is not strictly balanced, then there is a subset $S \subset [k]$ of size $3 \leqslant s < k$, such that

$$\frac{e_S - 1}{s-2} \geqslant \frac{e(\mathcal{G}_{k,\alpha}) - 2}{k-2}, \tag{3.3}$$

where e_S is the number of edges in $\mathcal{G}_{k,\alpha}$ with both endpoints in S. Let $e'_S = e(\mathcal{G}_{k,\alpha}) - e_S$ be the number of all other edges in $\mathcal{G}_{k,\alpha}$, i.e., those with at most one endpoint in S. For any given S, the random variables e_S and e'_S are independent. Since

$$\frac{1}{s-2} - \frac{1}{k-2} = \frac{k-s}{(s-2)(k-2)},$$

(3.3) implies that

$$\frac{e_S - 1}{s-2} \geqslant \frac{e'_S - 1}{k-s}. \tag{3.4}$$

Using the Chernoff bounds (3.1) and (3.2), we will estimate the probability that (3.4) occurs for some S of size $3 \leqslant s < k$. Note that

$$\mathbb{E}(e_S) = \frac{s(s-1)}{2}\alpha, \qquad \mathbb{E}(e'_S) = \frac{(k-s)(k+s-1)}{2}\alpha. \tag{3.5}$$

Therefore, if $e_S = (1+\delta)\mathbb{E}(e_S)$, then in order for (3.4) to hold, we would require $e'_S \leqslant (1 - \delta' + o(1))\mathbb{E}(e'_S)$, where

$$\delta' = \frac{k - \delta s}{k+s}. \tag{3.6}$$

Let $P_s(\delta)$ denote the probability that e_S and e'_S take such values. (We assume $\delta > -1$ throughout, since clearly the right hand side of (3.3) will be positive, with high probability.)

Case 1. First, we consider the case that $\delta < 1$. By (3.1) and (3.5), for any given S, it follows that e_S and e'_S take such values with probability at most $\exp[-\varepsilon\alpha\vartheta(\delta)]$, for some $\varepsilon > 0$, where

$$\vartheta(\delta) = \delta^2 s^2 + (\delta')^2(k^2 - s^2).$$

Note that $\vartheta(\delta)$ is convex and minimized at $\delta_* = (k-s)/2s$, at which point $\vartheta(\delta_*) = (k-s)k/2$. Therefore, when $\delta < 1$,

$$\binom{k}{k-s} P_s(\delta) \leqslant (ke^{-\varepsilon\alpha k})^{k-s}, \tag{3.7}$$

for some $\varepsilon > 0$.

Case 2. Next, we consider $\delta \geqslant 1$. We claim that

$$\binom{k}{s} P_s(\delta) \leqslant (ke^{-\varepsilon\alpha k})^{s}, \tag{3.8}$$

for some $\varepsilon > 0$. We will consider the cases that $k/2s$ is smaller/larger than δ separately. In the former case, we consider only the large deviation by e_S, and in the latter case, only the deviation by e'_S.

Case 2a. If $k/2s \leqslant \delta$, then consider the event $e_S = (1+\delta)\mathbb{E}(e_S)$. By (3.2) and (3.5), this occurs with probability at most $\exp[-\varepsilon\alpha s^2(k/s)]$, for some $\varepsilon > 0$. Therefore (3.8) holds in this case.

Case 2b. If $k/2s > \delta$, then consider the event $e'_S \leqslant (1 - \delta' + o(1))\mathbb{E}(e'_S)$, with δ' as in (3.6). Since $1 \leqslant \delta < k/2s$, we have $\delta' > 1/3$. By (3.1) and (3.5), for large k this occurs with probability at most $\exp[-\varepsilon\alpha(k^2 - s^2)]$, for some $\varepsilon > 0$. Then, since $s < k/2$, we find once again that (3.8) holds.

Finally, we take a union bound, summing over all relevant values $3 \leqslant s < k$ for the size of S. For each such S, there are $O(s^2)$ possible values of e_S. Therefore, combining (3.7) and (3.8) above, we find that (3.4) holds for some such S with probability at most $O(k^3 e^{-\varepsilon\alpha k})$, for some $\varepsilon > 0$. This is $o(1)$ for $\alpha \geqslant A(\log k)/k$, for large A, in which case $\mathcal{G}_{k,\alpha}$ is strictly balanced with high probability. Furthermore, for large A, this probability is summable. Therefore, by the Borel–Cantelli lemma, almost surely, we have that $\ell = 1/\lambda$, for all large k. □

3.1 Smaller A

For ease of exposition (and since, in our view, $\alpha = 1/2$ is the most interesting case) we have not pursued the smallest possible constant A in the proof above. However, we expect that more technical arguments can show that, with high probability, $\mathcal{G}_{k,\alpha}$ is balanced, and so $\ell = 1/\lambda$, provided that $\alpha = A(\log k)/k$ with $A > 2/\log(e/2)$.

Let us give some brief intuition in this direction. Recall that, in the proof above, the case $s = k-1$ is critical. This corresponds (roughly speaking) to the existence of a vertex v with degree less than half the average degree in the subgraph $F \subset \mathcal{G}_{k,\alpha}$ induced by the other $k-1$ vertices. Applying the sharper, relative entropy tail bound for the binomial (see, e.g., Diaconis and Zabell [6, Theorem 1]) when $\alpha = A(\log k)/k$, we have that

$$\mathbb{P}(\mathrm{Bin}(k,\alpha) \leqslant \alpha k/2) = O\left(\frac{k^{-(A/2)\log(e/2)}}{\sqrt{\log k}}\right).$$

Large deviations of e_F are significantly less likely, as the expected number of edges in F is of a larger order $O(\alpha k^2)$ than that $O(\alpha k)$ of the expected degree of v. Therefore, when $A = 2/\log(e/2)$, it can be seen that such a vertex v exists with probability at most $O(1/\sqrt{\log k}) = o(1)$. Therefore, if (2.1) fails, it is due to some smaller F, with $s \leqslant k-2$ vertices, however, the existence of such an F is increasingly less likely as s decreases. On the other hand, when $A < 2/\log(e/2)$ it can be shown, by the second moment method, that with high probability such a vertex v exists, in which case $\mathcal{G}_{k,\alpha}$ is unbalanced.

In closing, we remark that it might be of interest to study the behavior of ℓ as α decreases to the point $\alpha \sim (\log k)/k$ of connectivity. Note that $2/\log(e/2) \approx 6.518$. When the minimum degree of $\mathcal{G}_{k,\alpha}$ is at least 2, the general upper bound $\ell \leqslant 1/\lambda_*$ from [2] holds. In the extreme case that $\mathcal{G}_{k,\alpha}$ has a leaf, the value of ℓ is given by Proposition 26 in [1].

References

1. J. Balogh, B. Bollobás, and R. Morris, *Graph bootstrap percolation*, Random Structures Algorithms **41** (2012), no. 4, 413–440. MR2993128
2. Z. Bartha and B. Kolesnik, *Weakly saturated random graphs*, Random Structures Algorithms **65** (2024), no. 1, 131–148. MR4760769
3. M. R. Bidgoli, A. Mohammadian, and B. Tayfeh-Rezaie, *On $K_{2,t}$-bootstrap percolation*, Graphs Combin. **37** (2021), no. 3, 731–741. MR4249197
4. B. Bollobás, *Weakly k-saturated graphs*, Beiträge zur Graphentheorie (Kolloquium, Manebach, 1967), Teubner, Leipzig, 1968, pp. 25–31. MR0244077
5. J. Chalupa, P. L. Leath, and G. R. Reich, *Bootstrap percolation on a Bethe lattice*, J. Phys. C **21** (1979), L31–L35.
6. P. Diaconis and S. Zabell, *Closed form summation for classical distributions: variations on a theme of de Moivre*, Statist. Sci. **6** (1991), no. 3, 284–302. MR1144242
7. M. Pollak and I. Riess, *Application of percolation theory to 2d-3d Heisenberg ferromagnets*, Physica Status Solidi (b) **69** (1975), no. 1, K15–K18.
8. S. Ulam, *Random processes and transformations*, Proceedings of the International Congress of Mathematicians, Vol. 2, Cambridge, MA, 1950, pp. 264–275. MR0045334
9. J. von Neumann, *Theory of self-reproducing automata*, University of Illinois Press, Champaign, IL, USA, 1966.

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)
- Secure publication (LOCKSS[1])
- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4], ProjectEuclid[5]
- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

[1] LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2] EJMS: Electronic Journal Management System: https://vtex.lt/services/ejms-peer-review/
[3] IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4] BS: Bernoulli Society http://www.bernoulli-society.org/
[5] Project Euclid: https://projecteuclid.org/
[6] IMS Open Access Fund: https://imstat.org/shop/donation/