Percolation on Finite Transitive Graphs

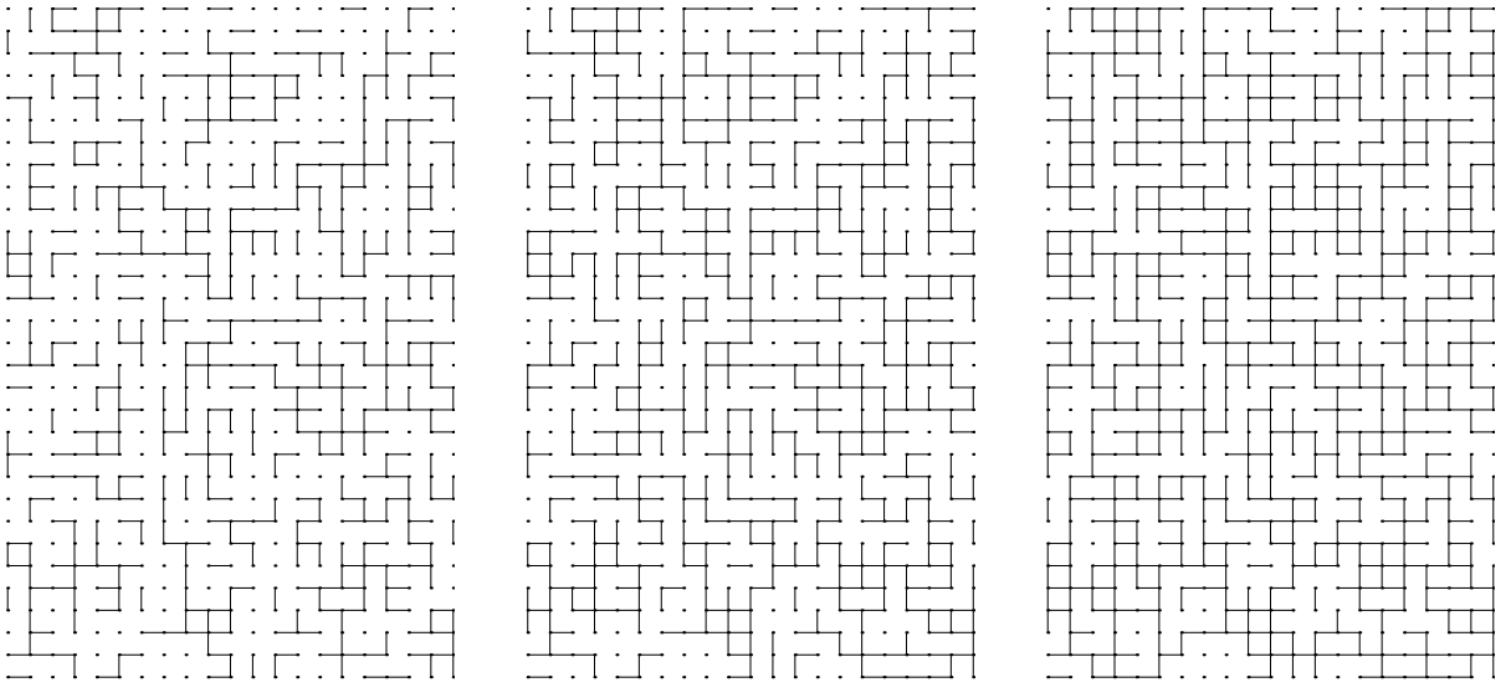
Tom Hutchcroft

California Institute of Technology

Joint work with Philip Easo

Percolation basics

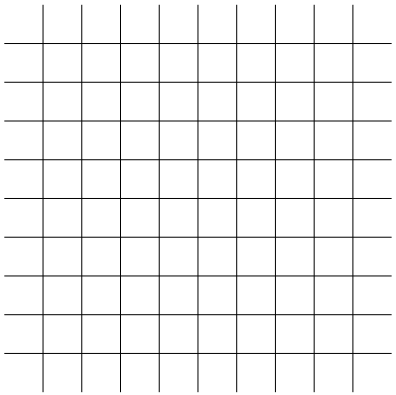
- Every edge of a graph G is deleted or retained independently with retention probability p.

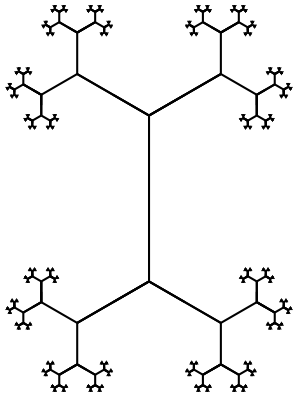


- Retained edges are **open**, deleted edges are **closed**. Open connected components are **clusters**.

Transitive graphs

We will restrict attention to **(vertex-)transitive** graphs, i.e., graphs for which for any two vertices $u, v \in V$ there is a graph automorphism (symmetry of the graph) mapping u to v.





This is partly because there is almost no good *general* theory without this assumption!

We will also take all our graphs to be **locally finite** (all degrees finite), **simple** (no self-loops or multiple edges), and **connected**.

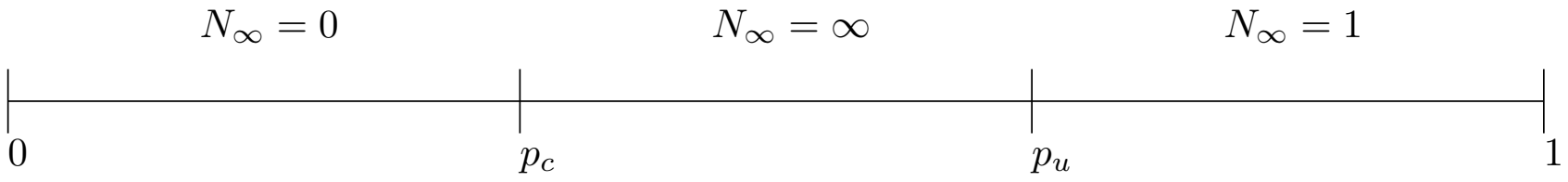
The phase transition in infinite graphs

- $\theta(p)$ probability that the origin is in an infinite cluster. **Critical probability** $p_c(G) = \inf\{p \in [0,1] : \theta(p) > 0\}$.
- **There is a phase transition:** Typically $0 < p_c < 1$.
- For the square grid $p_c = 1/2$ (Kesten, 1980's). Usually we don't expect p_c to have a nice or interesting value.
- Duminil-Copin, Goswami, Raoufi, Severo, and Yadin 2018: Every infinite transitive graph that is *not one-dimensional* in the sense that its graph-distance balls grow faster than linearly has $p_c < 1$.

How many infinite clusters are there?

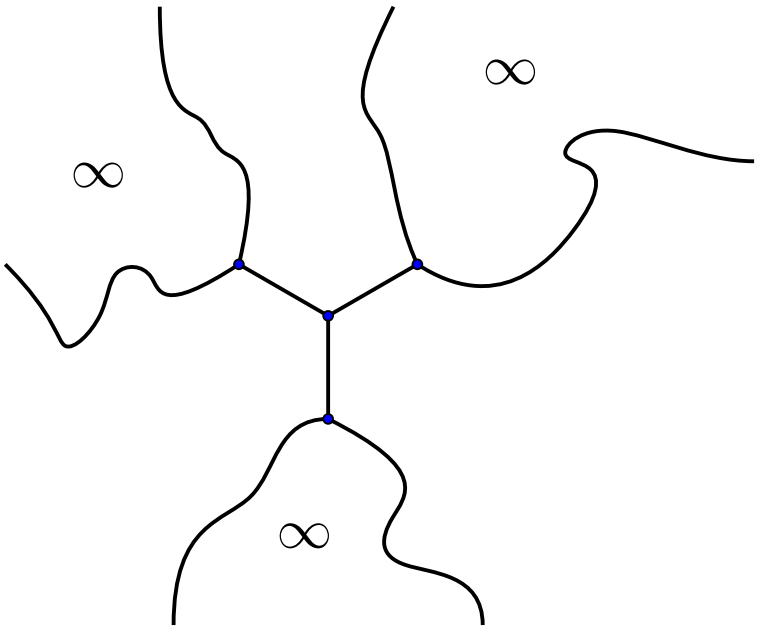
We define the **uniqueness threshold** to be

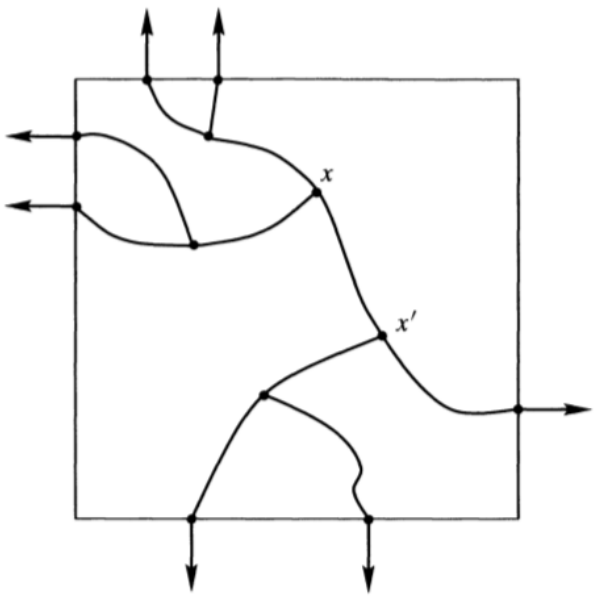
$$p_u(G) = \inf\{p \in [0,1] : \exists \text{ a } \textit{unique} \text{ infinite cluster a.s.}\}$$



- Newman and Schulman 1981: If G is transitive then $G[p]$ has either $0, 1$ or infinitely many infinite clusters.
- Häggströmm, Peres, and Schonmann 1999: If $p \in (p_u, 1]$ then $G[p_u]$ has a unique infinite cluster a.s.

- Aizenman, Kesten, and Newman 1987, Burton and Keane 1989: For each $p \in [0,1]$, $\mathbb{Z}^d[p]$ either has one infinite cluster or no infinite clusters almost surely. In particular $p_c = p_u$.





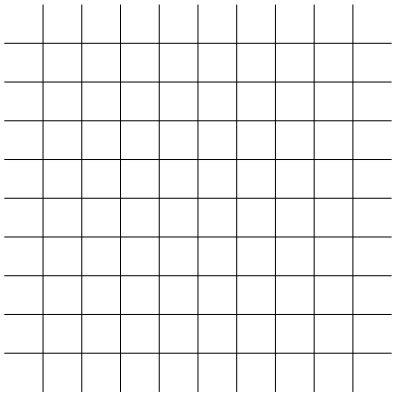
Amenability and nonamenability

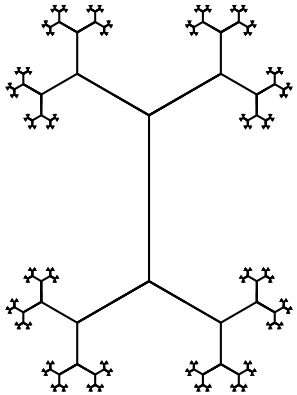
These proofs readily generalize to any **amenable** transitive graph.

Here, a locally finite graph $G = (V, E)$ is said to be **nonamenable** if its **Cheeger constant**

$$h(G) = \inf\left\{ \frac{\#\{\text{edges in the boundary of } W\}}{\#\{\text{edges with both endpoints in } W\}} : W \subset V \text{ finite} \right\}$$

is positive.

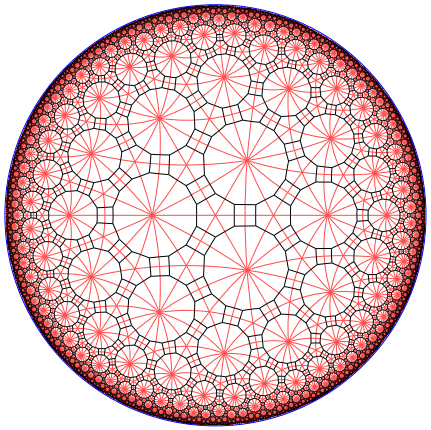


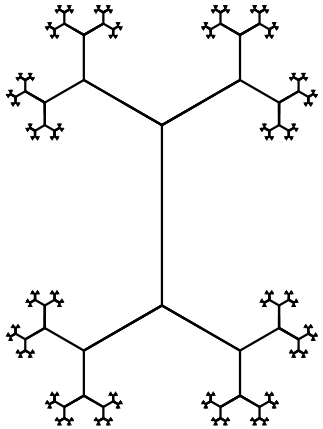


Conjecture (Benjamini and Schramm 1996)

For a transitive graph G, $p_c(G) = p_u(G)$ if and only if G is amenable.

Aizenman-Kesten-Newman/Burton-Keane proofs extend to any amenable transitive graph, so the problem is to prove the 'only if' direction.





Note that for the k-regular tree we have that $p_c = 1/(k-1)$, but p_u is trivially 1, so $p_c < p_u$.

The Erdős-Rényi random graph

Percolation on the complete graph C_n with parameter p_n.

- If $\limsup np_n < 1$ then every cluster has size $O(\log n)$ with high probability.
- If $\liminf np_n > 1$ then there exists a *unique* **giant** cluster of size $\Omega(n)$ with high probability.
- If $p_n = (1 - o(1))/n$ lots of interesting things happen. E.g. the largest cluster has size $\Theta(n^{2/3})$ with high probability when $p_n = 1/n$ (Bollobas 1984, Luczak 1994).

The Erdős-Rényi random graph

Percolation on the complete graph C_n with parameter p_n.

- If $\limsup np_n < 1$ then every cluster has size $O(\log n)$ with high probability.
- If $\liminf np_n > 1$ then there exists a *unique* **giant** cluster of size $\Omega(n)$ with high probability.
- If $p_n = (1 - o(1))/n$ lots of interesting things happen. E.g. the largest cluster has size $\Theta(n^{2/3})$ with high probability when $p_n = 1/n$ (Bollobas 1984, Luczak 1994).

Several other examples including the hypercube and the hamming graph are also well-studied by combinatorialsts: Similar results hold with the critical sequence $1/n$ replaced by $1/\deg$.

Probabilists have extensively studied the torus $(\mathbb{Z}/n\mathbb{Z})^d$, where a giant cluster emerges at $p_c(\mathbb{Z}^d)$. (Highly non-trivial for $d \geq 3$!)

The setting of finite transitive graphs interpolates between two classical settings:

- Infinite transitive graphs – usually studied by probabilists, often using ergodic-theoretic tools.
- The complete graph, hypercube, etc. – usually studied by combinatorialists, often using counting arguments.

Some natural questions to consider:

- When is there a non-trivial phase in which a giant cluster exists?
- When a giant cluster exists, must it be unique? Does this depend on the geometry?
- Is there a phase transition with an associated "critical sequence" like $p_n = 1/n$ is critical for Erdős-Rényi?
- Can we read off the properties of percolation on large finite transitive graphs by passing to appropriate infinite-volume limits?

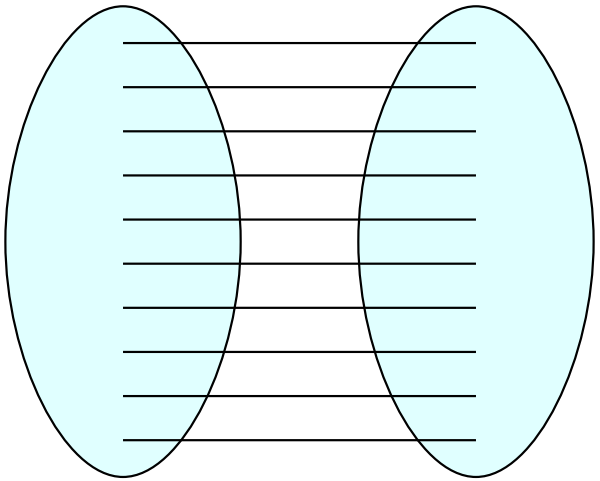
Unfortunately, neither set of tools is very well-adapted to this setting!

- Arguments from infinite volume often generalize in the "wrong way" – they often tell you something, but not what you want to know!
- For example, Burton-Keane only tells you that a typical small ball will not intersect many large clusters, and arguments about infinite clusters usually tell you about *divergently large* clusters than rather *giant* clusters.
- Proofs relying too heavily on qualitative, ergodic-theoretic techniques might tell you nothing at all in the finite case.
- Counting arguments typically don't work.

Unfortunately, neither set of tools is very well-adapted to this setting!

- Arguments from infinite volume often generalize in the "wrong way" – they often tell you something, but not what you want to know!
- For example, Burton-Keane only tells you that a typical small ball will not intersect many large clusters, and arguments about infinite clusters usually tell you about *divergently large* clusters than rather *giant* clusters.
- Proofs relying too heavily on qualitative, ergodic-theoretic techniques might tell you nothing at all in the finite case.
- Counting arguments typically don't work.
- Many things simply aren't true without further assumptions!

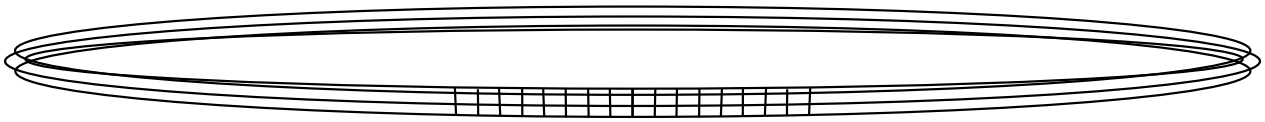
Counterexamples



$C_n \times C_2$

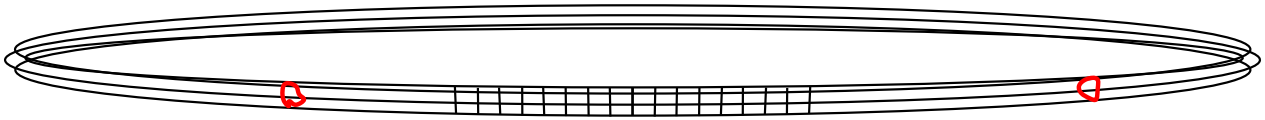
- If $p = 2/n$, each complete graph contains a giant cluster with high probability, but number of edges between two copies of C_n is approximately Poisson(2) distributed, so no such edges with good probability.
- $\Rightarrow$ Either one or two giant clusters, both with good probability.

Counterexamples



- Consider torus $(\mathbb{Z}/n\mathbb{Z}) \times (\mathbb{Z}/N\mathbb{Z})$ with $N \gg n$.
- Looks locally like $\mathbb{Z}^2$ when n large $\Rightarrow$ Origin in *divergently large* cluster with good probability when $p > 1/2$.
- For $p > 1/2$, each edge belongs to a dual cycle of closed edges cutting the torus with probability of order $\exp(-c_p n + o(n))$. The constant c_p is small for p close to $1/2$.
- $\Rightarrow$ If $N = 2^n$ and p is only slightly larger than $1/2$, so that $c_p < \log 2$, there are many large clusters but no giant cluster!

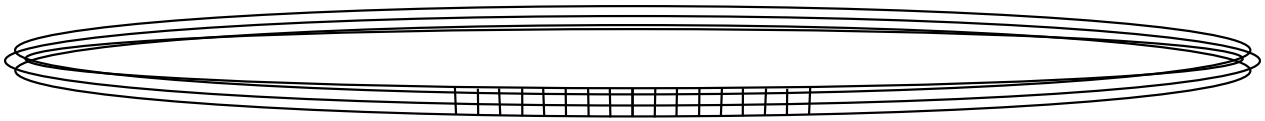
Counterexamples



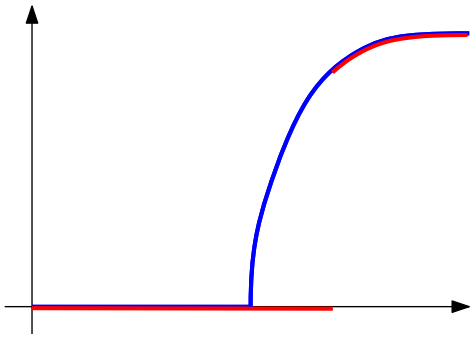
- Consider torus $(\mathbb{Z}/n\mathbb{Z}) \times (\mathbb{Z}/N\mathbb{Z})$ with $N \gg n$.
- If $N = 2^n$, we can choose p_n bounded away from zero and one so that "set of dual cuts" converges to a Poisson process on the circle.
- At this p_n the number of giant clusters is distributed asymptotically as $\max\{X - 1, 0\}$ where X is a Poisson random variable – any number of giants can occur with good probability!

(Similar things happen in the cycle $(\mathbb{Z}/n\mathbb{Z})$ when $p = 1 - \lambda/n$. In the long torus they can happen for p bounded away from 1.)

Counterexamples



The density of the giant is not always correctly predicted by the density of the infinite cluster in the limit!



If N grows faster than exponentially in n then there are never giant clusters at any fixed $p < 1$.

Conjectures: Existence

Alon, Benjamini, and Stacey conjectured that the 'long torus' example is worst case among bounded degree graphs, and that we should be able to rule out related pathological behaviour by putting very weak conditions on our graphs to exclude this example.

Conjecture (Benjamini)

If (G_n) has bounded degrees and

$$\mathrm{diam}(G_n) = O(|V_n|/\log|V_n|)$$

then there exists $\varepsilon > 0$ such that giant clusters exist with high probability for $p \geq 1 - \varepsilon$ as $n \to \infty$.

Conjectures: Existence

Theorem (H. & Tointon 2021)

If (G_n) has bounded degrees and

$$\mathrm{diam}(G_n) = O(|V_n|/(\log |V_n|)^{500})$$

then there exists $\varepsilon > 0$ such that giant clusters exist with high probability for $p \geq 1 - \varepsilon$ as $n \to \infty$.

Proof builds on the methods of Duminil-Copin, Goswami, Raoufi, Severo and Yadin plus **quantitative structure theory of low-growth transitive graphs** due to Breulliard, Green and Tao (for Cayley graphs) and Tessera and Tointon (for general transitive graphs).

Conjectures: Uniqueness

Alon, Benjamini, and Stacey conjectured that the 'long torus' example is worst case among bounded degree graphs, and that we should be able to rule out related pathological behaviour by putting very weak conditions on our graphs to exclude this example.

Conjecture (Alon, Benjamini, Stacey)

If (G_n) *has bounded degrees and*

$$\mathrm{diam}(G_n) = o(|V_n|/\log|V_n|)$$

then there is at most one giant cluster with high probability for every p.

Theorem (Alon, Benjamini, and Stacey)

If (G_n) is an **expander sequence** *then there is at most one giant cluster with high probability for every p.*

Moreover, if (G_n) converges locally to an infinite transitive graph G then there exists a giant cluster on G_n with high probability when $p > p_c(G)$, while the largest cluster is $O(\log |V_n|)$ with high probability when $p < p_c(G)$.

Note that non-uniqueness of the infinite cluster in the limit is not an obstruction to the uniqueness of the giant!

Conjectures: Uniqueness

Note that in the long torus with $N = 2^n$, we see multiple giant components *as a feature of a discontinuous phase transition*, rather than in the supercritical phase *per se*. As such, we might hope that the giant will be unique *in the supercritical regime* without any geometric assumptions on our graphs.

Conjecture (Benjamini)

If (G_n) has bounded degrees then there is at most one giant cluster with high probability for every 'supercritical' value of p.

This is the perspective we will take.

Conjectures: Uniqueness

Note that in the long torus with $N = 2^n$, we see multiple giant components *as a feature of a discontinuous phase transition*, rather than in the supercritical phase *per se*. As such, we might hope that the giant will be unique *in the supercritical regime* without any geometric assumptions on our graphs.

Conjecture (Benjamini)

If (G_n) has bounded degrees then there is at most one giant cluster with high probability for every 'supercritical' value of p.

This is the perspective we will take.

We will see that much more than this conjecture is true!

Let $(G_n) = (V_n, E_n)$ be a sequence of finite transitive graphs. Given a subset $A \subseteq V_n$ we write $\|A\| = |A|/|V_n|$ for the density of A.

Given $n \geq 1$, we write $K_1, K_2, \ldots$ for the clusters of percolation on G_n in size order (breaking ties arbitrarily).

Definition

We say that a sequence of parameters (p_n) is **critical** if

$$\lim_{n\to\infty} \mathbb{P}^{G_n}_{(1-\varepsilon)p_n}(\|K_1\| \geq c) = 0 \text{ for every } \varepsilon, c > 0$$

and for every $\varepsilon > 0$ there exists $\alpha > 0$ such that

$$lim_{n\to\infty}\mathbb{P}^{G_n}_{(1+\varepsilon)p_n}(\|K_1\| \geq \alpha) = 1.$$

Note that if (p_n) and (p'_n) both critical then $p_n \sim p'_n$.

Problem: Not clear that critical sequences always exist!

Supercritical sequences

Definition

We say that a sequence of parameters (p_n) is **supercritical** if there exists $\varepsilon > 0$ and $N < \infty$ such that

$$\lim_{n\to\infty} \mathbb{P}^{G_n}_{(1-\varepsilon)p_n}(\|K_1\| \geq \varepsilon) \geq \varepsilon.$$

Note that *if a critical sequence* (p_n) *exists*, then (q_n) is supercritical if and only if $\limsup q_n/p_n > 1$.

Supercritical uniqueness

Definition

We say that (G_n) has the **supercritical uniqueness property** if

$$\lim_{n\to\infty} \mathbb{P}_{p_n}(\|K_2\| \geq c) = 0$$

for every $c > 0$ and every supercritical sequence (p_n).

Benjamini's conjecture can now be stated precisely: Every sequence of bounded degree vertex-transitive graphs has the supercritical uniqueness property.

Sublinearity suffices

Theorem (Easo & H.)

Every **sparse** *sequence of vertex-transitive graphs has the supercritical uniqueness property.*

Here a graph sequence is said to be sparse if $\deg(G_n) = o(|V_n|)$.

The dense case

A graph sequence is dense if $\deg(G_n) = \Omega(|V_n|)$. We have seen that dense graph sequences such as $C_n \times C_2$ need not have the supercritical uniqueness property.

There are lots of simple variations on the same counterexample, with the same property.

Theorem (Easo & H.)

Every counterexample is of this form.

Molecular sequences

Given $m \geq 2$, we say a sequence of finite transitive graphs (G_n) is m-**molecular** if the following hold:

- (G_n) is dense.
- For each n there exists an automorphism-invariant set of edges F_n such that $G_n \setminus F_n$ has m connected components and $|F_n| = O(|V_n|)$.

We say (G_n) is molecular if it is molecular for some $m \geq 2$.

Theorem (Easo & H. 2021)

(G_n) has the supercritical uniqueness property if and only if it does not have any molecular subsequences.

Existence of critical sequences

Philip has more recently applied our results to show that critical sequences almost always exist.

Theorem (Easo 2022)

(G_n) admits a critical sequence if and only if it does not contain m-molecular subsequences for arbitrarily large m.

Interestingly, the proof of this theorem relies on our uniqueness theorem, which establishes properties of the supercritical phase without assuming that critical sequences exist.

Concentration and locality of the density

Theorem (Easo & H. 2022+)

If (G_n) does not have any molecular subsequences and (p_n) is supercritical then $\|K_1\|$ is concentrated:

$$\mathrm{Var}_{p_n}^{G_n}(\|K_1\|) \to 0.$$

If G_n has bounded degrees and converges locally to an infinite transitive graph G and $p_n \to p$ then $\|K_1\|$ converges to $\theta_G(p)$.

Note that both claims fail without the assumption that (p_n) is supercritical as we saw in the long torus.

Proof strategy (bounded degree case)

- Prove that if (p_n) is supercritical then $\inf \mathbb{P}(x \leftrightarrow y)$ bounded away from zero.
- Use **sharp threshold theory** to prove that there exists a supercritical sequence (q_n) with $q_n \leq p_n$ such that $\|K_1\|$ is concentrated at q_n.
- Deduce that the giant is unique *at the well-chosen parameter* q_n.
- **The sandcastle argument**: By analyzing the monotone coupling between q_n- and p_n-percolation, show that non-uniqueness of the giant at p_n contradicts concentration and uniqueness of the giant at q_n.

2nd and 4th steps are the most interesting.

Sharp threshold theory

An event is in **increasing** if adding more edges to the configuration can only help the event occur.

A sequence of increasing events has a **sharp threshold** if $\mathbb{P}_p(A)$ changes from close to 0 to close to 1 over an interval that is much smaller than the location of the threshold. This happens e.g. if $p\mathbb{P}_p(A) \gg 1$ over the interval of p for which $\mathbb{P}_p(A) \in [\varepsilon, 1-\varepsilon]$ for each fixed $\varepsilon > 0$.

Theorem (Talagrand)

Sufficiently symmetric increasing events depending on n bits always have sharp thresholds provided that the threshold occurs at a value of p subpolynomially small in n.

Consider the increasing event $\{\|K_1\| \geq \alpha\}$.

In our context, Talagrand implies that this event always has a sharp threshold, with probability changing from close to 0 to close to 1 over an interval of width $O(1/\log|V_n|)$.

Consider the function $\theta(p) = \sup\{\alpha : \mathbb{P}_p(\|K_1\| \geq \alpha) \geq \frac{1}{2}\}$.

Since (p_n) is supercritical, there exists $\varepsilon > 0$ such that $(1-\varepsilon)p_n$ is supercritical.

If we split $[(1-\varepsilon)p_n, p_n]$ into $\sqrt{\log|V_n|}$ equal-length pieces, there must be a piece where θ increments by at most $1/\sqrt{\log|V_n|}$.

Take q_n to be the midpoint of this interval. At this q_n we must have that $\|K_1\|$ is concentrated in an interval of width $O(1/\sqrt{\log|V_n|})$!

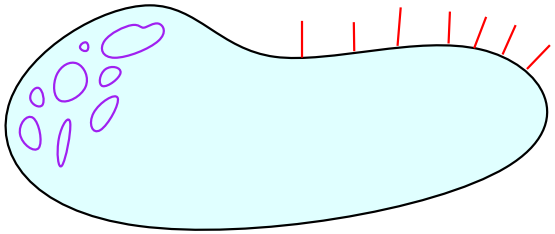
Uniqueness at q_n

Since $\|K_1\|$ is concentrated at q_n, by passing to a subsequence if necessary we may assume $\|K_1\| \to \alpha$ for some $\alpha > 0$.

If there was a good probabaility of a large K_2 at q_n, we could "glue" this cluster to K_1 with good probability to get that $\|K_1\|$ has a constant-order fluctuation with good probability, contradicting concentration.

Sandcastles

- We can couple p_n and q_n percolation by performing q_n/p_n percolation independently on each cluster of p_n percolation.
- If there is uniqueness at q_n but not at p_n, all but one of the giant p_n clusters must disintegrate into small clusters under this q_n/p_n percolation.



- If we condition on this cluster, the edges that do not touch it are still distributed as independent percolation, and must contain a largest cluster of size $\sim \alpha|V_n|$ with high probability.

Sandcastles

It follows that there exists a set of vertices S such that both

- $|S| = \Omega(|V_n|)$ and
- q_n percolation on $G_n \setminus S$ has good probability to contain a cluster of size $\sim \alpha|V_n|$.
- Since every vertex belongs to a q_n-giant cluster with probability $\sim \alpha$, we have by Markov that S contains at least $\frac{\alpha}{2}|S|$ vertices belonging to giant clusters with good probability.
- It follows from FKG that, with good probability, at least $\alpha|V_n| + \frac{\alpha}{2}|S|$ vertices belong to giant clusters with good probability. But this contradicts uniqueness and concentration of the giant at q_n!

The high degree case

The argument just sketched shows very generally that supercritical uniqueness holds whenever the events $\{\|K_1\| \geq \alpha\}$ all have sharp thresholds (in a sufficiently uniform way).

When $\deg(G_n) = |V_n|^{o(1)}$ this comes 'for free' from Talagrand. For higher degrees we must use some specific property of the event beyond symmetry.

This relies in part on characterizations of sharp thresholds for increasing events with polynomially small threshold functions due to Bourgain and Hatami; we show that the only obstruction to $\{\|K_1\| \geq \alpha\}$ having a sharp threshold is for the graph to be molecular.

An open problem

In the bounded degree case, our proof yields a quantitative bound: If (p_n) is supercritical then

$$\|K_2\| = O\left(\frac{|V_n|}{\sqrt{\log|V_n|}}\right) \qquad \text{with high probability.}$$

On the other hand, the worst known example maximizing the size of the second largest supercritical component is the 2d torus $(\mathbb{Z}/n\mathbb{Z})^2$, where the second largest component has volume of order $(\log n)^2$ with high probability.

Problem

How large can the second largest component be?