Bootstrap Percolation in Random Geometric Graphs

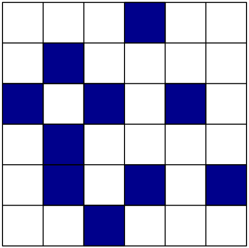
Amites Sarkar

Western Washington University

2 November 2022

Joint work with Victor Falgas-Ravry (Umeå University)

Bootstrap percolation (Chalupa, Leath, Reich 1979)



Start with an $n \times n$ grid S_n

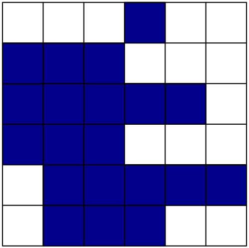
$A_0 \subset S_n$: sites initially infected independently with probability p

$A_0 \subset A_1 \subset S_n$: simultaneously, sites with at least two out of four neighbors in A_0 become infected

Repeat to get $A_2, A_3, \ldots$

What is $A_\infty = \bigcup_{t \geq 0} A_t$? Is it S_n?

Bootstrap percolation (Chalupa, Leath, Reich 1979)



Start with an $n \times n$ grid S_n

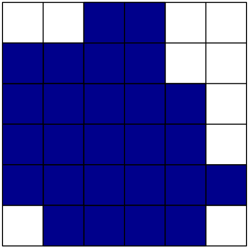
$A_0 \subset S_n$: sites initially infected independently with probability p

$A_0 \subset A_1 \subset S_n$: simultaneously, sites with at least two out of four neighbors in A_0 become infected

Repeat to get $A_2, A_3, \ldots$

What is $A_\infty = \bigcup_{t \geq 0} A_t$? Is it S_n?

Bootstrap percolation (Chalupa, Leath, Reich 1979)



Start with an $n \times n$ grid S_n

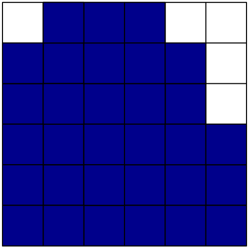
$A_0 \subset S_n$: sites initially infected independently with probability p

$A_0 \subset A_1 \subset S_n$: simultaneously, sites with at least two out of four neighbors in A_0 become infected

Repeat to get $A_2, A_3, \ldots$

What is $A_\infty = \bigcup_{t \geq 0} A_t$? Is it S_n?

Bootstrap percolation (Chalupa, Leath, Reich 1979)



Start with an $n \times n$ grid S_n

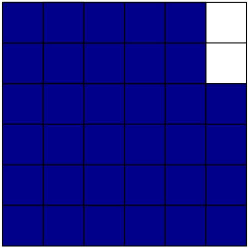
$A_0 \subset S_n$: sites initially infected independently with probability p

$A_0 \subset A_1 \subset S_n$: simultaneously, sites with at least two out of four neighbors in A_0 become infected

Repeat to get $A_2, A_3, \ldots$

What is $A_\infty = \bigcup_{t \geq 0} A_t$? Is it S_n?

Bootstrap percolation (Chalupa, Leath, Reich 1979)



Start with an $n \times n$ grid S_n

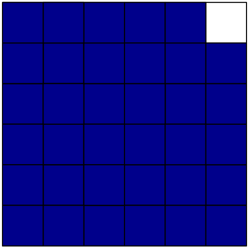
$A_0 \subset S_n$: sites initially infected independently with probability p

$A_0 \subset A_1 \subset S_n$: simultaneously, sites with at least two out of four neighbors in A_0 become infected

Repeat to get $A_2, A_3, \ldots$

What is $A_\infty = \bigcup_{t \geq 0} A_t$? Is it S_n?

Bootstrap percolation (Chalupa, Leath, Reich 1979)



Start with an $n \times n$ grid S_n

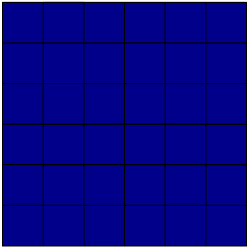
$A_0 \subset S_n$: sites initially infected independently with probability p

$A_0 \subset A_1 \subset S_n$: simultaneously, sites with at least two out of four neighbors in A_0 become infected

Repeat to get $A_2, A_3, \ldots$

What is $A_\infty = \bigcup_{t \geq 0} A_t$? Is it S_n?

Bootstrap percolation (Chalupa, Leath, Reich 1979)



Start with an $n \times n$ grid S_n

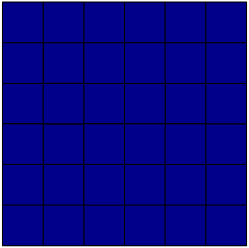
$A_0 \subset S_n$: sites initially infected independently with probability p

$A_0 \subset A_1 \subset S_n$: simultaneously, sites with at least two out of four neighbors in A_0 become infected

Repeat to get $A_2, A_3, \ldots$

What is $A_\infty = \bigcup_{t \geq 0} A_t$? Is it S_n?

Bootstrap percolation (Chalupa, Leath, Reich 1979)



Start with an $n \times n$ grid S_n

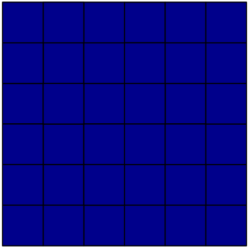
$A_0 \subset S_n$: sites initially infected independently with probability p

$A_0 \subset A_1 \subset S_n$: simultaneously, sites with at least two out of four neighbors in A_0 become infected

Repeat to get $A_2, A_3, \ldots$. What is $A_\infty = \bigcup_{t \geq 0} A_t$? Is it S_n?

Holroyd (2003) Sharp metastability threshold at $p \log n = \pi^2/18$

Bootstrap percolation (Chalupa, Leath, Reich 1979)



Start with an $n \times n$ grid S_n

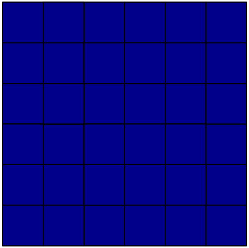
$A_0 \subset S_n$: sites initially infected independently with probability p

$A_0 \subset A_1 \subset S_n$: simultaneously, sites with at least two out of four neighbors in A_0 become infected

Repeat to get $A_2, A_3, \ldots$. What is $A_\infty = \bigcup_{t \geq 0} A_t$? Is it S_n?

Holroyd's results were extended to other dimensions d and thresholds s by Balogh, Bollobás, Duminil-Copin and Morris (2012)

Bootstrap percolation (Chalupa, Leath, Reich 1979)



Start with an $n \times n$ grid S_n

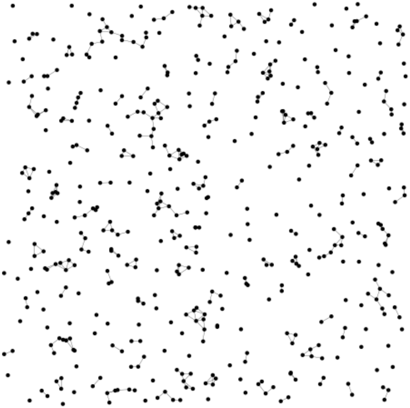
$A_0 \subset S_n$: sites initially infected independently with probability p

$A_0 \subset A_1 \subset S_n$: simultaneously, sites with at least two out of four neighbors in A_0 become infected

Repeat to get $A_2, A_3, \ldots$. What is $A_\infty = \bigcup_{t \geq 0} A_t$? Is it S_n?

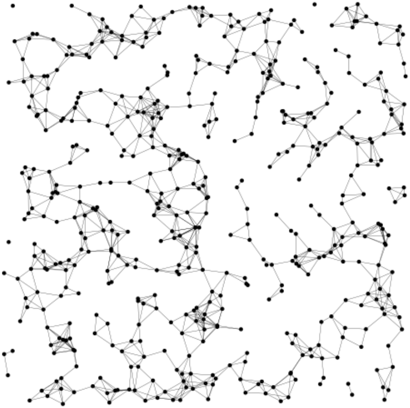
Janson, Łuczak, Turova and Vallier (2012) studied the model on Erdős-Rényi random graphs $G(n, q)$

Random geometric graphs (Gilbert 1961)



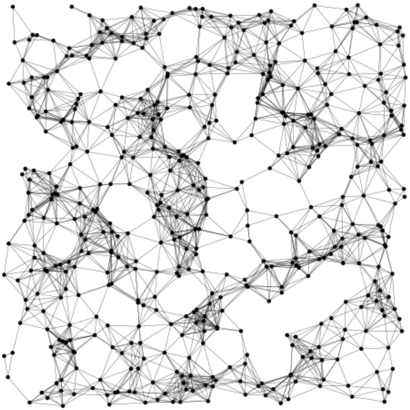
Vertices (nodes) are a Poisson process of intensity 1
Edges join vertices at distance less than r
Gilbert's motivation: communications networks

Percolation



Hall (1985) $0.833 < r_{\mathrm{perc}} < 1.836$
Balister, Bollobás and Walters (2005) $1.1978 < r_{\mathrm{perc}} < 1.1989$
- semi-rigorous, high confidence result

Connectivity



Penrose (1997) $\pi r_{\mathrm{conn}}^2(n) = \log n$

Obstruction to connectivity: isolated vertices

At the threshold, $\mathbb{E}(\textit{isolated vertices}) = 1$

The Bradonjić-Saniee model (2014)

Start with the Gilbert model in a torus T_n of area n, above the connectivity threshold

$$\pi r^2 = a \log n \text{ with } a > 1$$

Initially infect vertices independently with probability p: this is A_0
Each vertex expects

$a \log n$ neighbors

$ap \log n$ infected neighbors

$A_t :=$ set of infected vertices at time t
In each discrete time step ($t = 1, 2, \ldots$)

For each $v \notin A_t$ (i.e. each uninfected v)

If v has at least $a\theta \log n$ infected neighbors

- v becomes infected (and stays infected forever)

Repeat for each vertex v to get A_{t+1}

Repeat for each t to get A_∞

What proportion $|A_\infty|/n$ of the graph eventually becomes infected?

Motivation

- Activation of neurons
- Economic networks
- Social networks
- Spread of viruses

Mathematical motivation

Extend methods developed to study connectivity in:

- the Gilbert model
- the k-nearest neighbor model

The percolative regime

What if $\pi r^2 = a$ and the infection threshold is k?

Theorem (Whittemore 2021)

Define

$$p^* = \frac{1}{n^{1/k} a^{1-1/k}}$$

Then for $1 \ll a \ll n$ and $p/p^* \to 0$ **whp** no initially inactive vertex becomes infected, but for $1 \ll a \ll n$ and $p/p^* \to \infty$ **whp** almost every initially uninfected vertex becomes infected.

Theorem (Whittemore 2021)

For $1 \ll a \ll n$ and $p/p^* = \gamma$, there exists $\alpha = \alpha(\gamma)$ such that

$$\alpha \leq \mathbb{P}(A_1 \neq A_0) \leq 1 - \alpha$$

Theorem (Bradonjić and Saniee 2014)

For $x > 0$, define

$$J(x) = \log x - 1 - 1/x$$

and write J_R^{-1} for the inverse of J on $[1, \infty]$. Then if

$$p < p' = \theta / J_r^{-1}(1/a\theta)$$

then no initially uninfected vertex becomes infected.

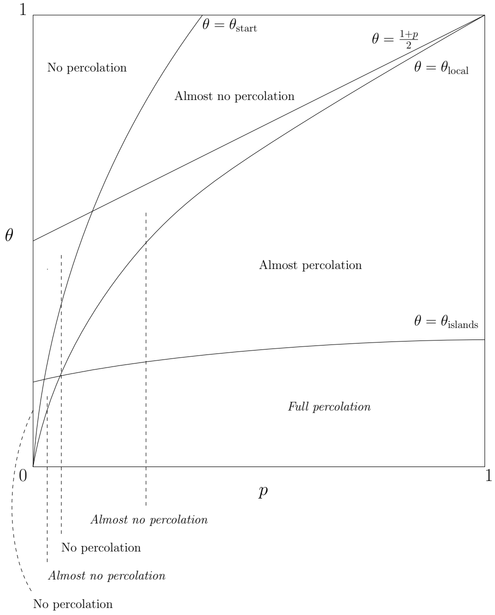
Theorem (Bradonjić and Saniee 2014)

If

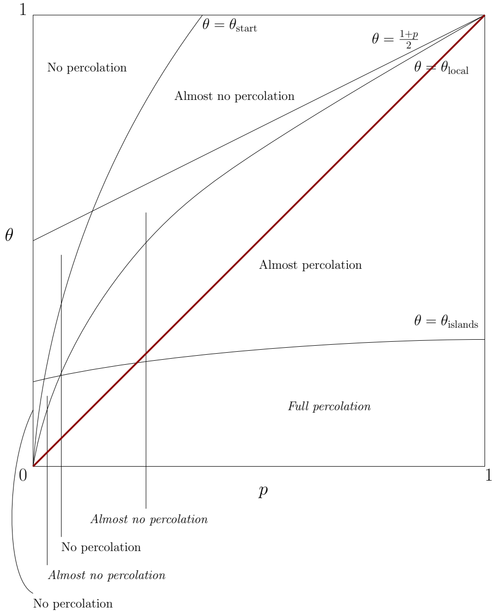
$$p > p'' = \min \left\{ \theta, \frac{5\pi\theta}{J_r^{-1}(1/a\theta)} \right\}$$

then every initially uninfected vertex becomes infected.

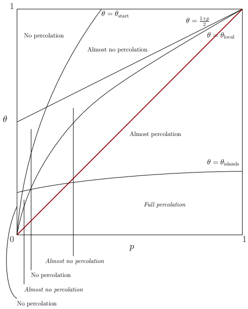
Theorems (Falgas-Ravry and S 2022+)



Basic orientation - the threshold $\theta = p$



Basic orientation - the threshold $\theta = p$



If $\theta < p$, almost everything becomes infected immediately.

If $\theta > p$, almost no new infections occur initially.

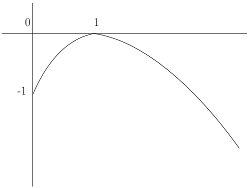
A useful lemma

Let $A \subset R^d$ be measurable, and let $\rho \geq 0$ be a real number such that $\rho|A| \in \mathbb{Z}$. Then the probability that a Poisson process in $\mathbb{R}^d$ with intensity 1 has precisely $\rho|A|$ points in the region A is given by

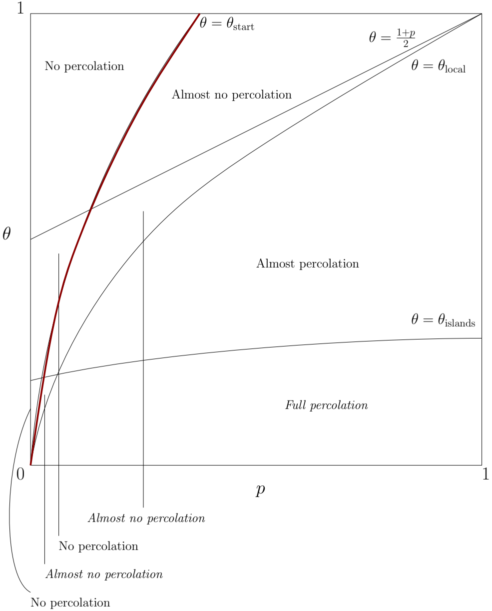
$$\exp\left\{(\rho - 1 - \rho\log\rho)|A| + O(\log_+ \rho|A|)\right\}$$

with the convention that $0\log 0 = 0$, and $\log_+ x = \max(\log x, 1)$. We will usually apply this lemma when $|A| = C\log n$, so that the relevant probability will be approximated by

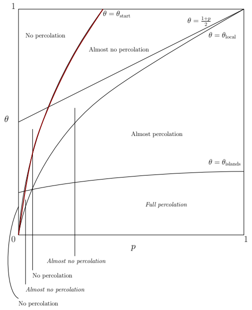
$$n^{-C(\rho-1-\rho\log\rho)}$$



The starting threshold $\theta = \theta_{\text{start}}(p)$



The starting threshold $\theta = \theta_{\text{start}}(p)$

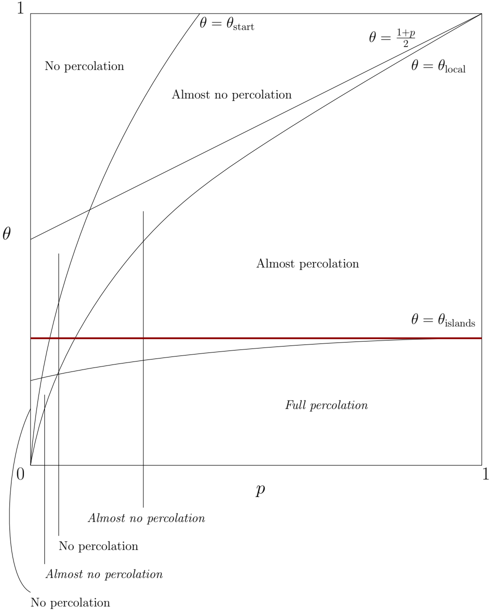


Sometimes, even when the threshold θ is much greater than p, some uninfected vertices will see $a\theta \log n$ infected neighbors, despite only expecting to see only $ap \log n$.

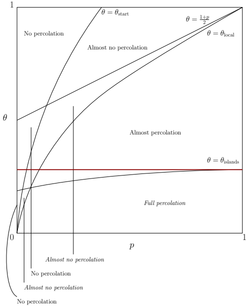
This will happen when

$$f_{\text{start}}(a, p, \theta) = a(p - \theta + \theta \log(\theta/p)) < 1$$

The simple stopping threshold $\theta = \theta_{\text{stop}}$



The simple stopping threshold $\theta = \theta_{\mathrm{stop}}$

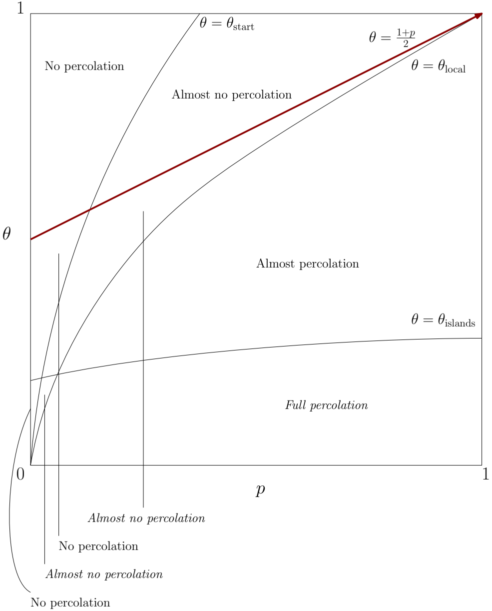


On the other hand, some initially uninfected vertices will not even have $a\theta \log n$ neighbors, despite only expecting to see $a \log n$. These vertices can never become infected.

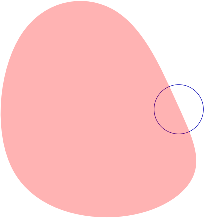
This will happen when

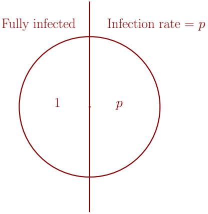
$$f_{\mathrm{stop}}(a, \theta) = a(1 - \theta + \theta \log \theta) < 1$$

The growing threshold $\theta = \frac{1+p}{2}$



The growing threshold $\theta = \frac{1+p}{2}$





When infections have broken the logarithmic barrier, they will grow as long as

$$\theta < \frac{1+p}{2}$$

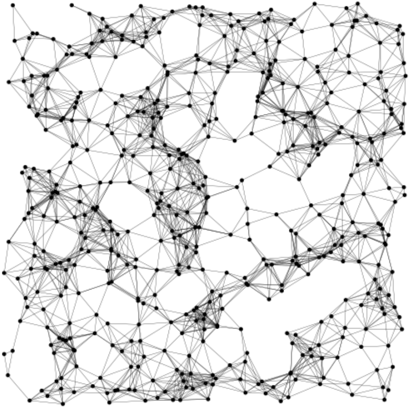
Though intuitive, this is nontrivial to prove.

Main theorem

Let a, θ, p be fixed. Then the following hold.

1. If $\theta < \frac{1+p}{2}$, then there exists a constant $C = C(a, \theta, p)$ such that w.h.p. if any ball B in T_n^2 of radius Cr is infected (either artificially or as a result of the bootstrap percolation process), then all but $o(n)$ vertices of $G_{n,r}^2$ eventually become infected. Furthermore, when the infection stops, all connected components of uninfected vertices in $G_{n,r}^2[\mathcal{P} \setminus A_\infty]$ have Euclidean diameter $O(\sqrt{\log n})$ in T_n^2.
2. If $\theta > \frac{1+p}{2}$, then for every constant $C > 0$, w.h.p. even if one adversarially selects a ball B in T_n^2 of radius Cr and infects all the vertices it contains, only $o(n)$ additional vertices of $G_{n,r}^2$ become infected in the bootstrap percolation process starting from the initially infected set $A_0 \cup (B \cap \mathcal{P})$.
   Furthermore, all components of $G_{n,r}^2[A_\infty \setminus (A_0 \cup B)]$ have Euclidean diameter $O(\sqrt{\log n})$ in T_n^2.

Connectivity



Penrose (1997) $\pi r_{\text{conn}}^2(n) = \log n$

Obstruction to connectivity: isolated vertices

At the threshold, $\mathbb{E}(\textit{isolated vertices}) = 1$

The isoperimetric argument for the Gilbert model in T_n

Suppose $\pi r^2 = \log n$.

Why are there no two large components in $G_r(n)$?

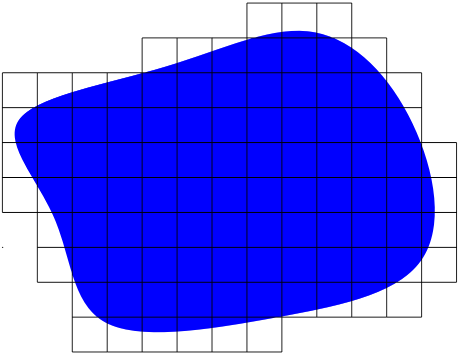
Two vertices x and y of $G_r(n)$ are joined iff $||x - y|| < r$.

- Edges from different components of $G_r(n)$ do not cross.
- Edges from different components of $G_r(n)$ are separated by $r/2$.

Tessellate T_n with squares of side length $r/\sqrt{20}$.
Points in neighboring squares lie at distance at most $r/2$.

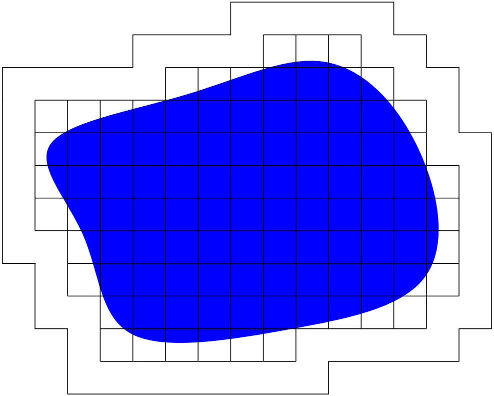
Color squares blue if they intersect an edge of a fixed large component of $G_r(n)$.

The isoperimetric argument for the Gilbert model



Two large components in G_r are separated by a long boundary B.

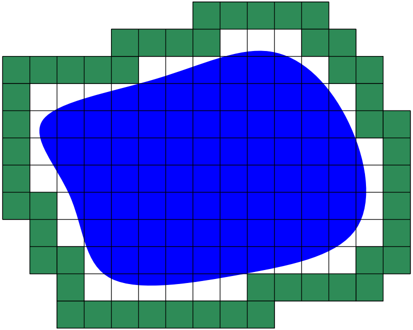
The isoperimetric argument for the Gilbert model



Two large components in G_r are separated by a long boundary B.

B yields a long *empty* external vertex boundary B_T.

The isoperimetric argument for the Gilbert model



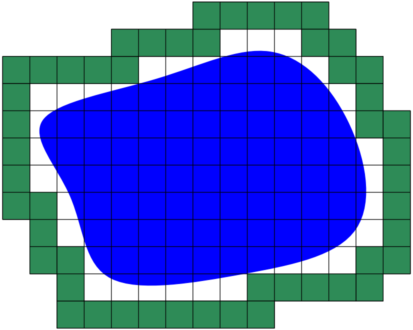
Suppose B_T consists of $K \gg 1$ squares.

There are at most $n(8e)^K$ choices for B_T.

Each square in B_T is empty with probability n^{-C}.

The expected number of such configurations is at most $n(8e)^K n^{-CK} \to 0$ as $n \to \infty$, if K is sufficiently large.

The Bradonjić-Saniee model when $\theta < \frac{1+p}{2}$



This time, the (frozen) boundary of a growing infection is not so well-defined.

On the boundary, there will be a mixture of infected and uninfected points.

Also, there need not be any vacant squares, just low-density regions.

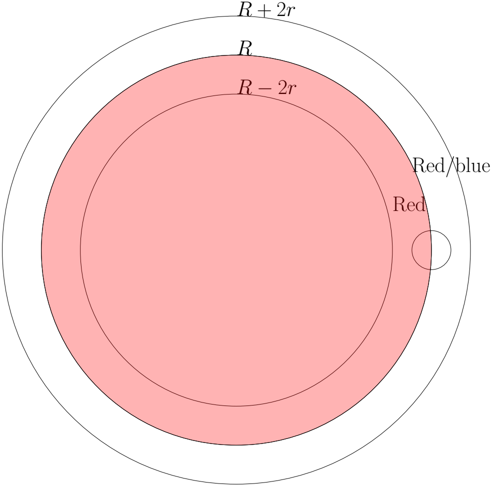
A tale of two tilings

Fine tiling $\mathcal{F}$: tiles of side length $r/K, K \gg 1$
A fine tile $T \in \mathcal{F}$ is coloured **white** if either it contains fewer than $(1-\eta)p|T|$ initially infected points, or fewer than $(1-\eta)|T|$ points in total. Otherwise, we colour T **red** if all its points are infected by the end of the bootstrap percolation process, and **blue** if this is not the case.

Rough tiling $\mathcal{R}$: tiles of side length $Kr, K \gg 1$
A tile in $\mathcal{R}$ is colored **white** if one of its subtiles in $\mathcal{F}$ is coloured white, **red** if all its subtiles in $\mathcal{F}$ are coloured red, and **blue** otherwise. $\mathbb{P}(\text{rough tile is white}) = O(n^{-C})$.

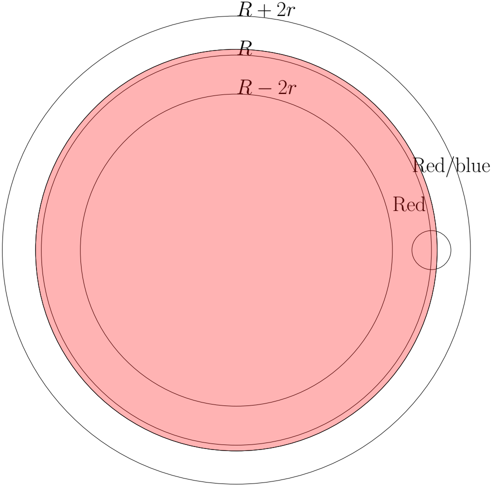
The fine tiling $\mathcal{F}$

If a large circular component of fine red tiles has no fine white tiles in its vicinity



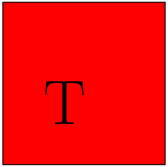
The fine tiling $\mathcal{F}$

If a large circular component of fine red tiles has no fine white tiles in its vicinity its radius expands by at least δr.



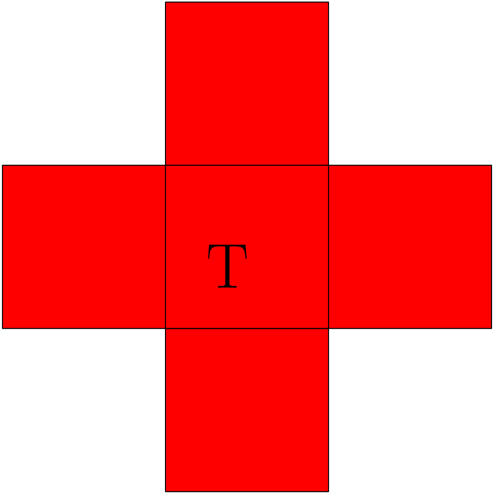
The rough tiling $\mathcal{R}$

Suppose a rough tile T is red.



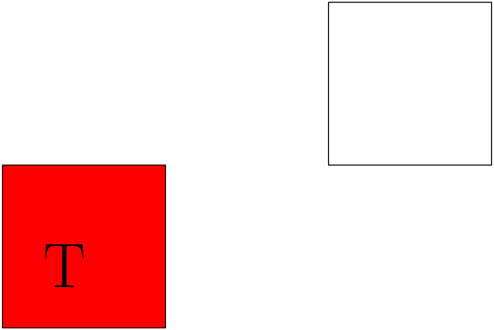
The rough tiling $\mathcal{R}$

Then either all its neighbors are red...



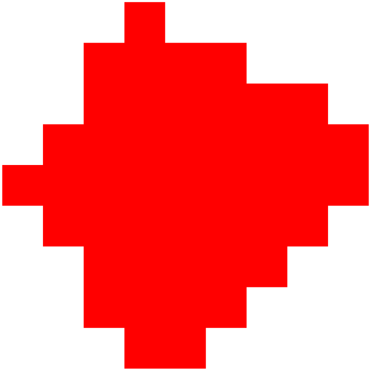
The rough tiling $\mathcal{R}$

or there is a white rough tile at graph distance at most 3 from T.



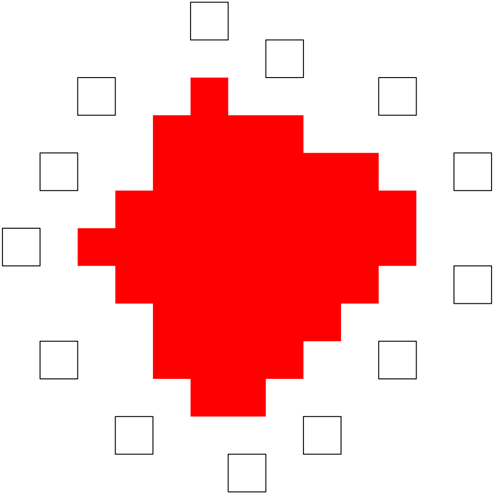
The rough tiling $\mathcal{R}$

Consequently, any large component of red rough tiles



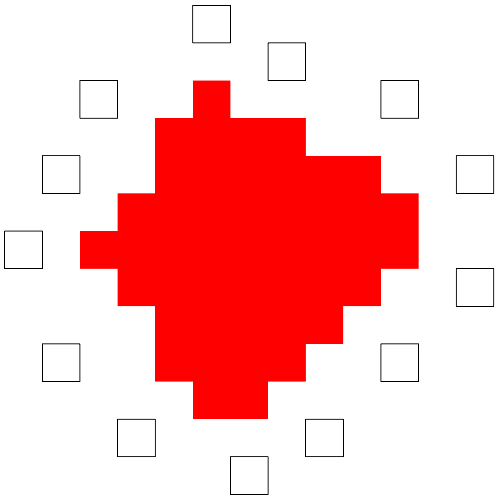
The rough tiling $\mathcal{R}$

Consequently, any large component of red rough tiles must be associated with a long cycle of white rough tiles in H^8.

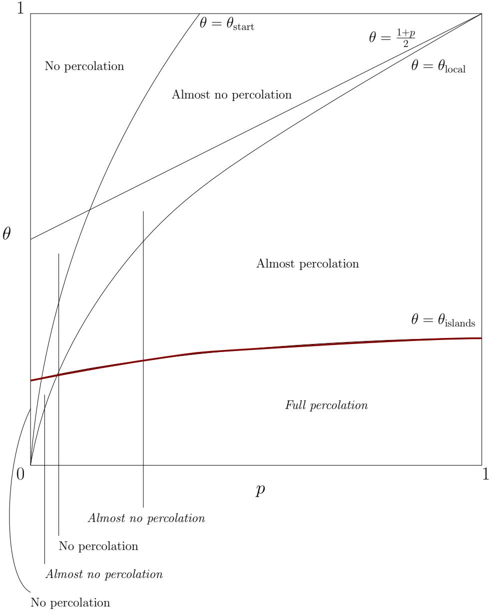


The rough tiling $\mathcal{R}$

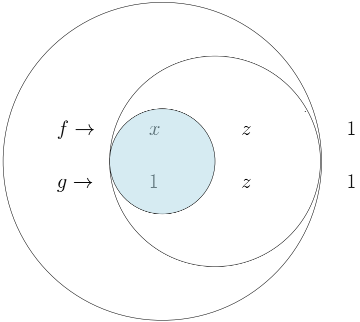
Consequently, any large component of red rough tiles must be associated with a long cycle of white rough tiles in H^8. This has probability $o(n^{-1})$, and so is unlikely to occur anywhere in T_n.



The threshold for full percolation $\theta = \theta_{\text{islands}}(p)$



An obstruction to full percolation with diameter r

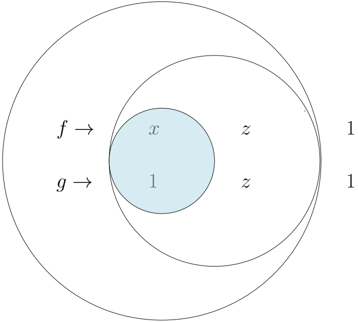


Condition for non-infectability: $G(x,z) = 3z + px < 4\theta$

Probability of configuration: $q = \exp\left\{\frac{a \log n}{4} F(x,z)\right\}$

$F(x,z) = 8(z - 1 - z\log z) + p(x - 1 - x\log x)$

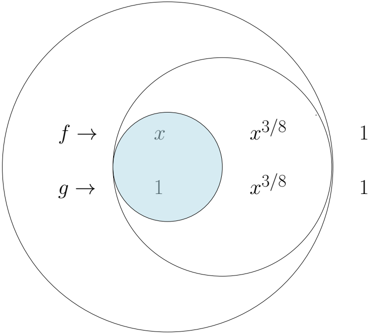
An obstruction to full percolation with diameter r



Maximize: $F(x, z) = 8(z - 1 - z \log z) + p(x - 1 - x \log x)$

subject to: $G(x, z) = 3z + px = 4\theta$

An obstruction to full percolation with diameter r

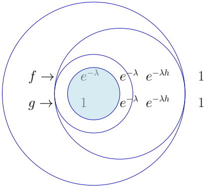


Solution: $z = x^{3/8}$ (recall $3z + px = 4\theta$)

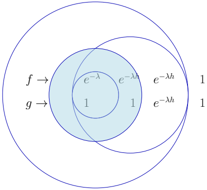
Threshold:

$$4 + a\left\{8(x^{3/8} - 1 - \tfrac{3}{8}x^{3/8}\log x) + p(x - 1 - x\log x)\right\} = 0$$

The threshold for full percolation $\theta = \theta_{\mathrm{islands}}(p)$



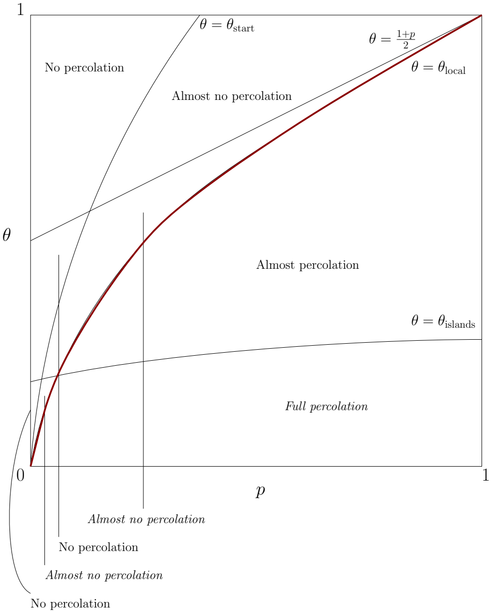
$\tau \leq 1/2$



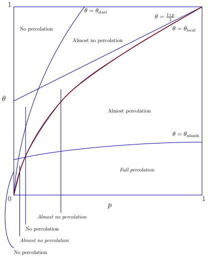
$1/2 \leq \tau \leq 1$

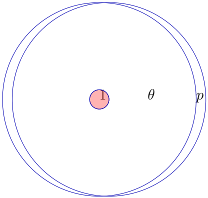
- Vary diameter $D = 2\tau r$ and let densities f, g vary continuously
- For fixed τ, optimize f, g
- Then optimize over τ

The local growth threshold $\theta = \theta_{\text{local}}(p)$



The local growth threshold $\theta = \theta_{\text{local}}(p)$



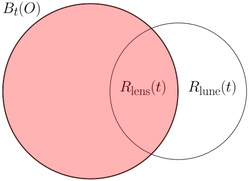


To break the logarithmic barrier, infections need to do more than just start.

They need to be able to expand beyond each radius τr.

This yields a Lagrange multiplier problem with **infinitely many conditions**.

The local growth threshold $\theta = \theta_{\mathrm{local}}(p)$



Find functions f, g which maximize

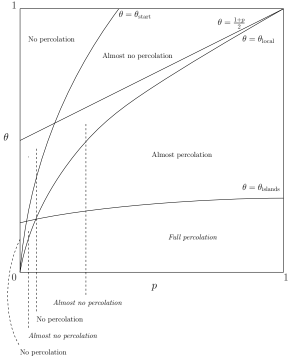
$$q(f,g) := \int_{x \in \mathbb{R}^2} p\left(f(\|x\|) - 1 - f(\|x\|)\log[f(\|x\|)]\right) + (1-p)\left(g(\|x\|) - 1 - g(\|x\|)\log[g(\|x\|)]\right) dx$$

subject to

$$I(f,g)(t) := \int_{x \in R_{\mathrm{lens}}(t)} \left(pf(\|x\|) + (1-p)g(\|x\|)\right) dx + \int_{x \in R_{\mathrm{lune}}(t)} pf(\|x\|) dx > \theta$$

for all $t \geq 0$.

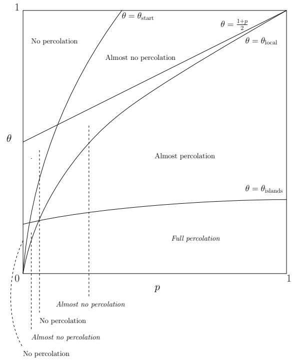
What did we actually prove?



The local growth threshold is only a sufficient condition for local growth, and the islands threshold is only a necessary condition for full percolation.

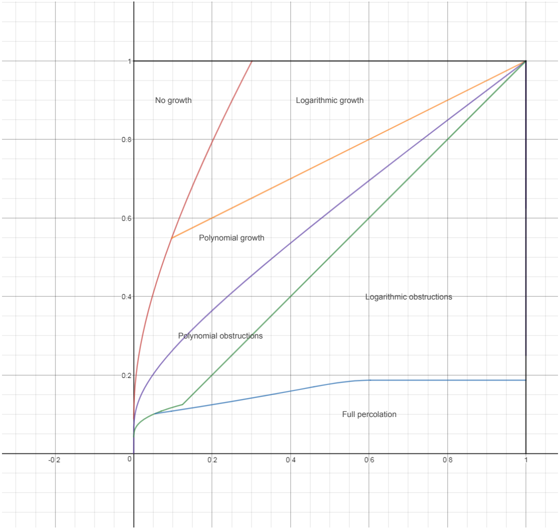
Accordingly, these thresholds only provide a lower bound for local growth and an upper bound for full percolation.

The main open questions



- Prove that the local growth threshold is also a necessary condition for local growth, so that the solution to the above optimization problem yields the correct threshold for local growth.
- Prove that the symmetric islands described above are in fact the last obstacles to full percolation, so that the islands threshold is the correct threshold for full percolation.

It’s more complicated in one dimension



Thank you for your attention!