

# Marrying Graph Convergence and Epidemics

---

Yeganeh Alimohammadi  
Stanford University

# Motivation

Infection spreads over a contact network.

**Question:** What information about the network do we need to forecast an outbreak?

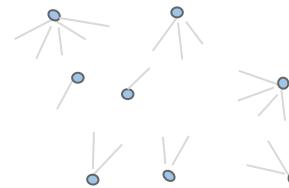
- So far, we saw model-dependent answers.

# Network Model and Estimation of Epidemics

1. Model the interaction between people with a network model:  
Erdos Renyi, Configuration Model, Preferential Attachment, Stochastic Block Model, Household Models, etc.
2. Estimate the relevant model parameters.



Erdős-Rényi: average degree



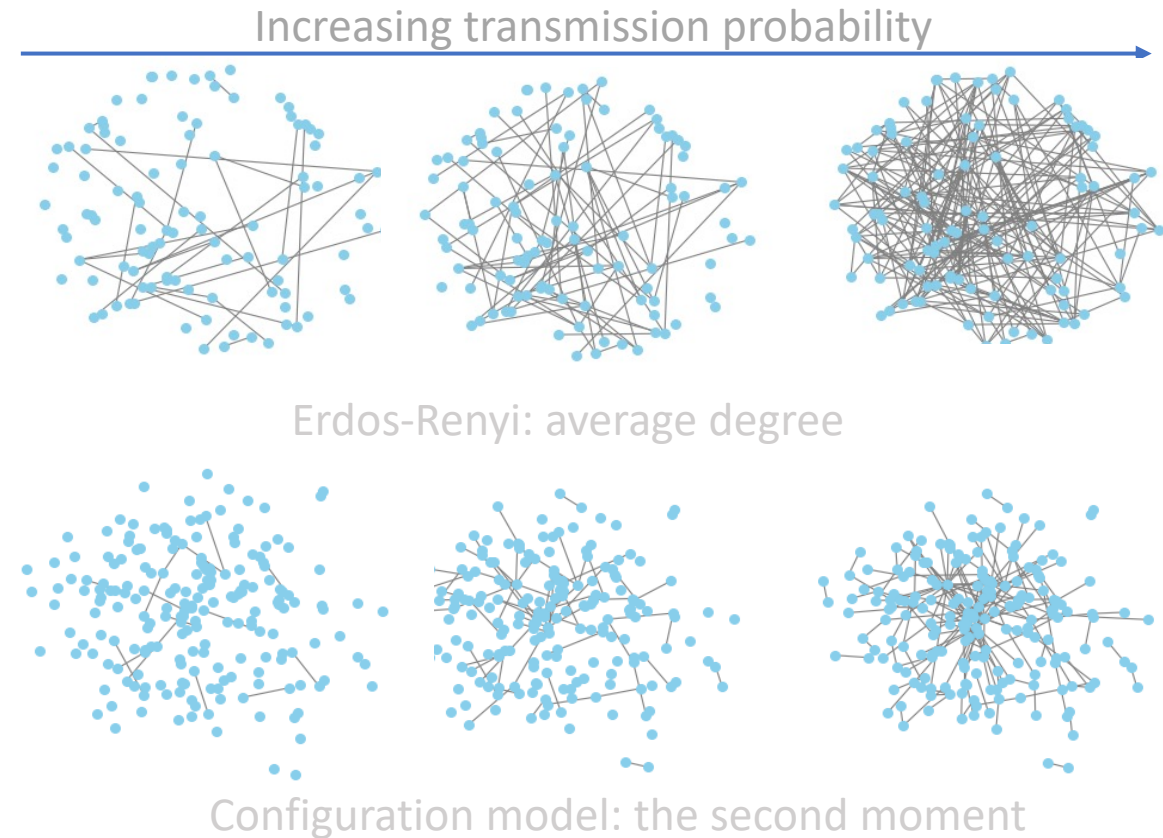
Configuration model: degree sequence

Question: Can we have a model-free estimation on different **properties** of epidemics?

# Example Properties of Epidemics

Different models share similar qualitative properties:

- Critical probability/phase transition of emergence of the outbreak (giant)
- Uniqueness of the outbreak (giant)
- Convergence of the size of the outbreak (giant)



Is there a meta theorem without assuming the underlying **model** or **full knowledge of the graph**?

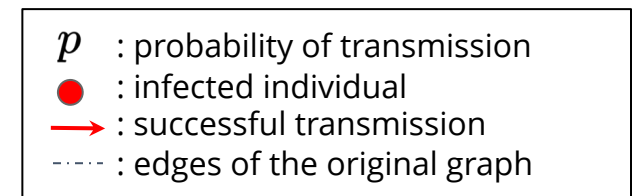
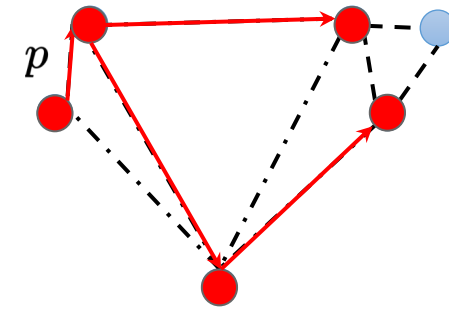
# Recap: Simple Model of Epidemics (Percolation)

Initially, one node (chosen uniformly at random) is infected.

An infected node transmits the disease to each neighbor independently with probability  $p$ , and then recovers (and will be immune to re-infection).

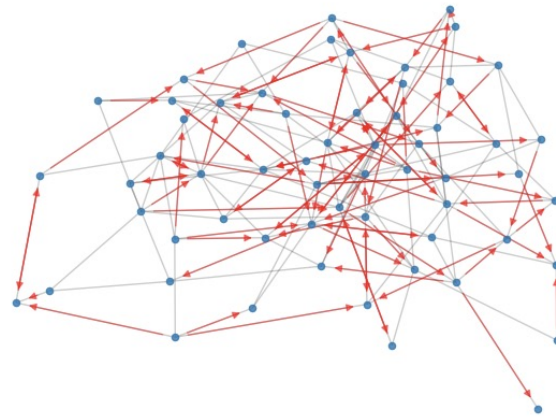
Equivalent to SIR with Constant Recovery time.

Percolation: keep each edge with probability  $p$  (call this graph  $G(p)$ ).



# Definition. Directed Percolation

- Replace each edge  $\{i, j\}$  by directed edges  $i \rightarrow j$  and  $j \rightarrow i$ ,
- Keep directed edges independently with probability  $p$ .



# This Talk in a Nutshell

Under some assumption (expansion) on converging graphs:

Critical probability converges to its limit.

Giant is unique, and its size converges to its limit.

We give an algorithm to estimate the limit.

The directed percolation on convergent sequence of expanders has:

- the same critical probability as the undirected percolation.
- a bow-tie structure.

# Results: Epidemics on Expanders



# Critical Probability

## Definition. (Critical Probability)

Given an infinite graph  $G$ , the critical  $p_c(G)$  is defined as

$$p_c(G) = \inf\{p \in [0,1]: \mathbb{P}_{G(p)}(\exists \text{ an infinite component in } G(p)) > 0\}.$$

## Theorem 1. [Benjamini, Nachmias, Peres '09]

Let  $G_n$  be a sequence of  $\alpha$ -expanders, with a uniform bounded degree  $d$ , and local weak limit  $G$ . If  $p < p_c(G)$ , then for any constant  $\beta > 0$ ,

$$\mathbb{P}(\exists \text{ a component of size at least } \beta n \text{ in } G_n(p)) \rightarrow 0 \text{ as } n \rightarrow \infty$$

and if  $p > p_c(G)$ , then there exists a constant  $\beta > 0$  such that

$$\mathbb{P}(\exists \text{ a component of size at least } \beta n \text{ in } G_n(p)) \rightarrow 1 \text{ as } n \rightarrow \infty.$$

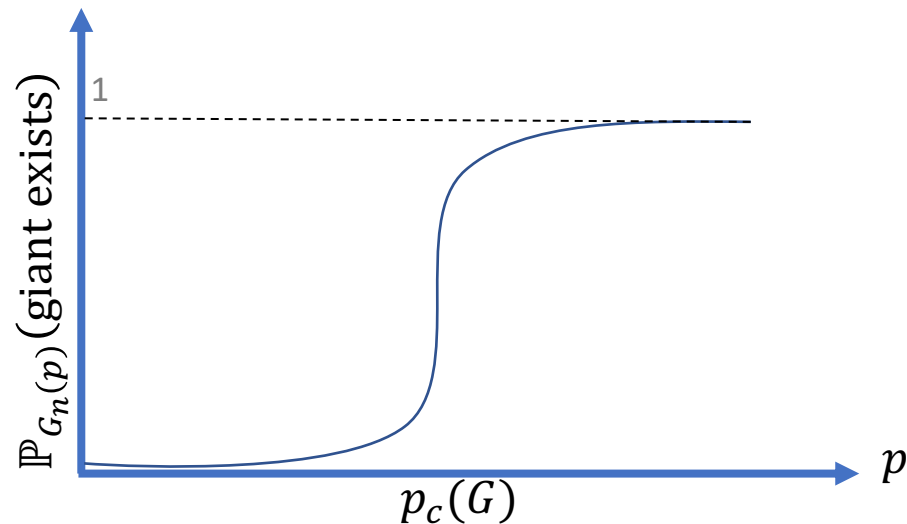
**Takeaway:** Critical probability in convergent expanders is local, and there's a phase transition at  $p_c(G)$ .

# Critical Probability

## Definition. (Critical Probability)

Given an infinite graph  $G$ , the critical  $p_c(G)$  is defined as

$$p_c(G) = \inf\{p \in [0,1]: \mathbb{P}_{G(p)}(\exists \text{ an infinite component in } G(p)) > 0\}.$$



**Takeaway:** Critical probability in convergent expanders is local, and there's a phase transition at  $p_c(G)$ .

# Recap: Local Convergence

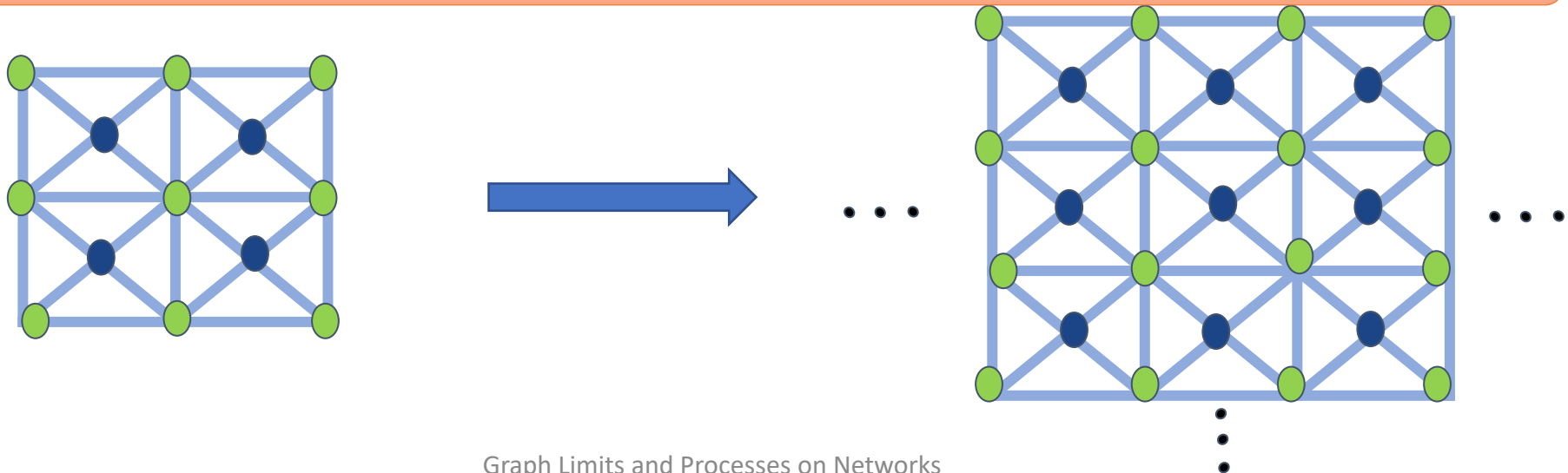
## Definition. (Local Convergence in Probability [Benjamini, Schramm '01])

A sequence of finite graphs  $\{G_n\}_{n \in \mathbb{N}}$  converges locally in probability to  $\mu$  if for any bounded continuous function  $f: \mathcal{G}_* \rightarrow \mathbb{R}$ ,

$$\mathbb{E}_{\mathcal{P}_n}[f|G_n] \xrightarrow{\mathbb{P}} \mathbb{E}_\mu[f],$$

where in  $\mathbb{E}_{\mathcal{P}_n}[f|G_n]$ , we take expectation with respect to the uniform random root in  $G_n$ .

**Takeaway:** the distribution of the neighborhood of a typical node converges.



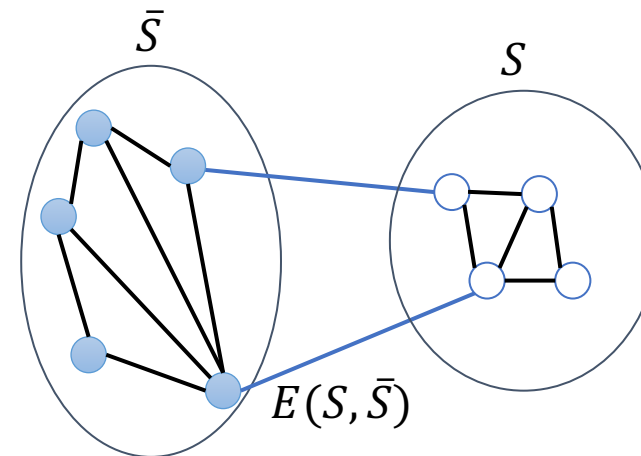
# Expanders

## Definition. (Expanders)

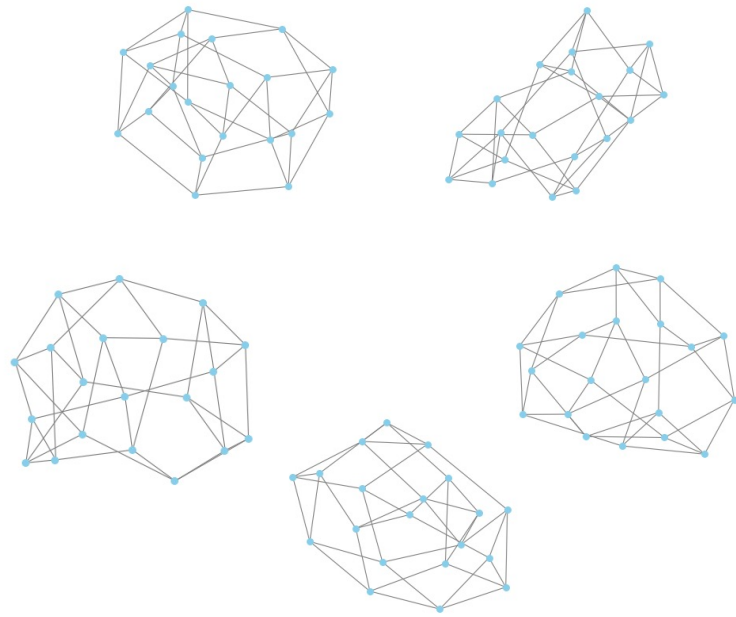
$G$  is  $\alpha$ -expander if  $\phi(G) \geq \alpha$ , where

$$\phi(G) = \min_{S \subseteq V(G)} \frac{E(S, \bar{S})}{\min(|S|, |\bar{S}|)}$$

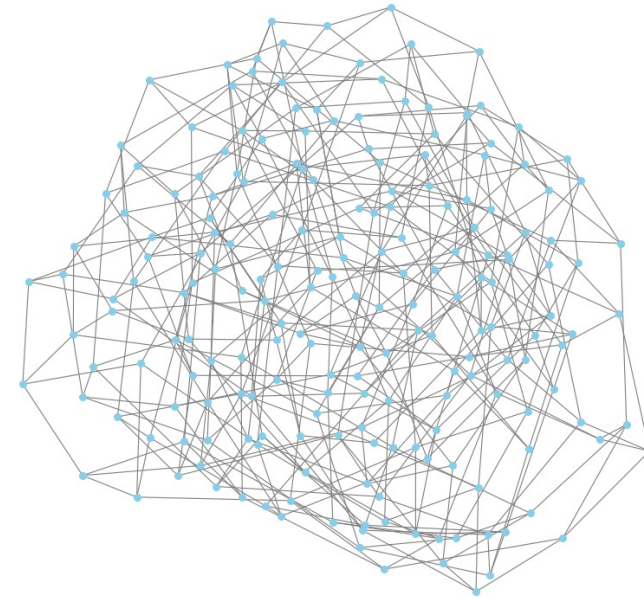
**Takeaway:** If you want to isolate a large community from the rest of the town, you need to remove many connections.



# Necessity of Expansion: Same Graph Limit but Different Epidemics



A collection of  $\frac{n}{\log n}$  4-regular random graphs, each of size  $\log n$



A 4-regular random graph of size  $n$

# Uniqueness of the Giant

## Theorem. [Alon, Benjamini, Stacey '04]

Let  $\{G_n\}_{n \in \mathbb{N}}$  be a sequence of (possibly random) expanders of size  $n$  with bounded maximum degree. Let  $\beta > 0$ , and  $p_n \in [0,1]$ . Then

$$\mathbb{P}_{G_n(p_n)}(\exists \text{ more than one component of size at least } \beta n \text{ in } G_n(p_n)) \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

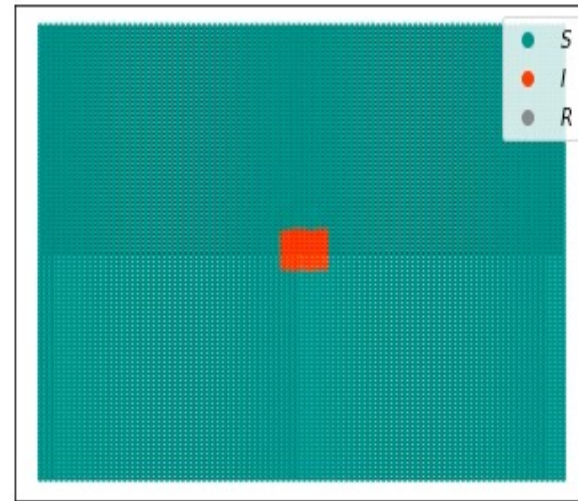
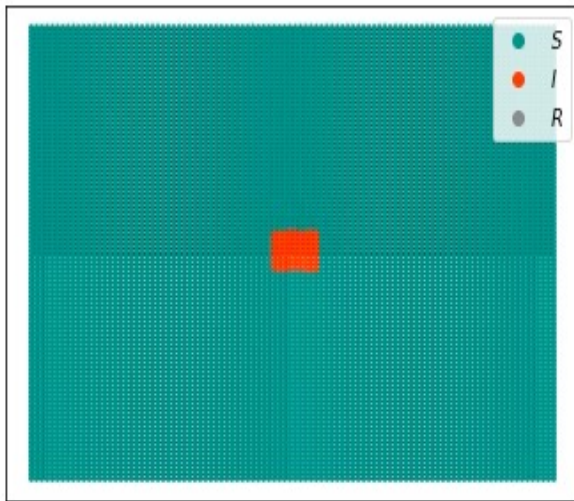
**Takeaway:** Giant in expanders is unique.

Previous two theorems show that the existence and uniqueness of the giant.

But what about its size?

# Relative Size of the Giant

Two copies of a network, with two runs of the same infection led to an outbreak. Can we predict the number of infected?



# Relative Size of the Giant in Expanders

## Theorem 2. [A., Borgs, Saberi '21]

Let  $\{G_n\}_{n \in \mathbb{N}}$  be a sequence of (possibly random) **large-set expanders** with bounded average degree converging locally in probability to  $(G, o) \in \mathfrak{G}_*$  with non-random distribution  $\mu$ . Let  $C_i$  be the  $i^{\text{th}}$  largest component. If  $p \neq p_c(\mu)$ ,

$$\frac{|C_1|}{n} \xrightarrow{\mathbb{P}} \zeta(p).$$

Further, for all  $p \in [0, 1]$ ,  $\frac{|C_2|}{n} \xrightarrow{\mathbb{P}} 0$ .

$\xrightarrow{\mathbb{P}}$ : convergence in probability in percolation and  $\mu$ .

$\zeta(p) := \mathbb{E}_{(G, o) \sim \mu} [\mathbb{P}_{G(p)}(|\text{connected component of } o| = \infty)]$ .

**Takeaway 1:** Giant in convergent expanders is unique, and its size converges to its limit.

**Corollary:** With high probability, the final infection size is either either  $O(1)$  or  $\Theta(n)$ .



# Algorithmic Implication

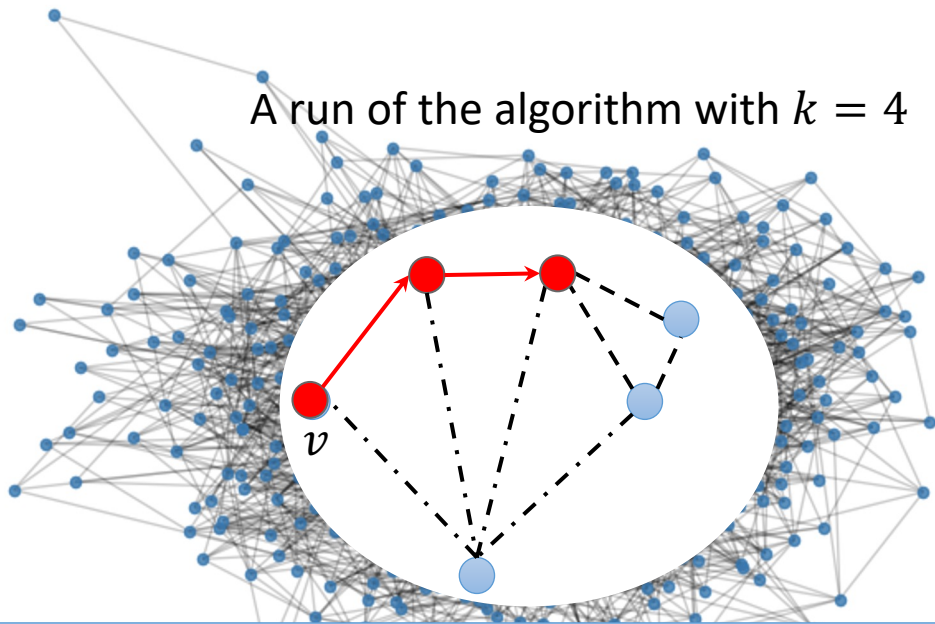
**Input:** a constant  $k$ .

1. Draw a uniform random node  $v$ .
2. Simulate an infection starting from  $v$ .
3. If  $v$  can lead to infecting  $k$  others:

**return 1.**

otherwise:

**return 0.**



## Theorem 3. [A., Borgs, Saberi '22]

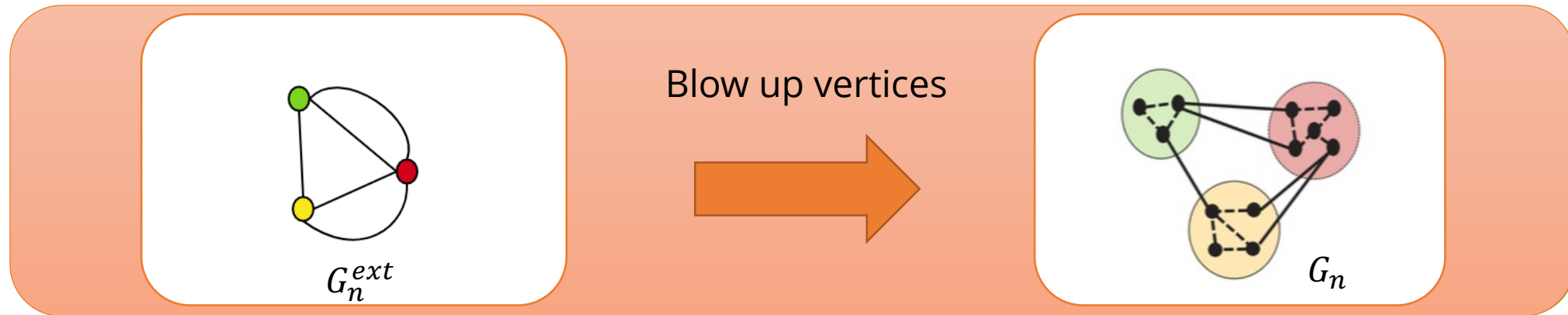
Let  $\{G_n\}_{n \in \mathbb{N}}$  be a sequence of (possibly random) graphs converging locally in probability to  $(G, o) \in \mathfrak{G}_*$  with distribution  $\mu$ , such that  $\frac{|C_1|}{n} \xrightarrow{\mathbb{P}} \zeta(p)$ .

Then for any  $\epsilon > 0$ , there exist constants  $q_\epsilon, k_\epsilon \geq 0$ , such that whp  $q_\epsilon$  queries to the above algorithm with input  $k_\epsilon$  (denoted by  $\tilde{N}(q_\epsilon, k_\epsilon)$ ) is a  $(1 - \epsilon)$ -approximation of  $\zeta(p)$ .

Formally, there exists  $n_\epsilon > 0$ , such that for all  $n > n_\epsilon$ ,  $\mathbb{P}(|\tilde{N}(q_\epsilon, k_\epsilon) - \zeta(p)| \geq \epsilon) \leq \epsilon$ .

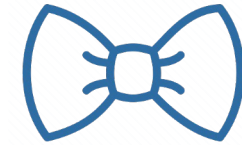
# Examples of Convergent Large-set Expanders

- Configuration Model [Molloy, Reed, Newman, Barabasi, Watts '11]
- Preferential Attachment [Bollobás, Riordan '03]
- Household models [Ball, Sirl, Trapman. 2009, Hofstad, Leeuwaarden, Stegehuis. '15 -- for configuration model]



**Informal Lemma.** If  $\{G_n^{ext}\}_{n \in \mathbb{N}}$  are convergent large-set expanders then  $\{G_n\}_{n \in \mathbb{N}}$  is as well.

# Directed Percolation Creates a Bow-Tie



## Theorem 4. [A., Borgs, Saberi]

For a sequence of large-set expanders as in Theorem 2,  $\frac{|SCC_2|}{n} \xrightarrow{\mathbb{P}} 0$ .

Also, if  $p > p_c(G)$ :

- $\liminf_{n \rightarrow \infty} \frac{|SCC_1|}{n} \geq \zeta^2(p)$  and  $\frac{|SCC_1|}{\mathbb{E}|SCC_1|} \xrightarrow{\mathbb{P}} 1$ .

Critical probability is the same as undirected.

- $\frac{1}{n} |SCC_1^+| \xrightarrow{\mathbb{P}} \zeta(p)$  and  $\frac{1}{n} |SCC_1^-| \xrightarrow{\mathbb{P}} \zeta(p)$

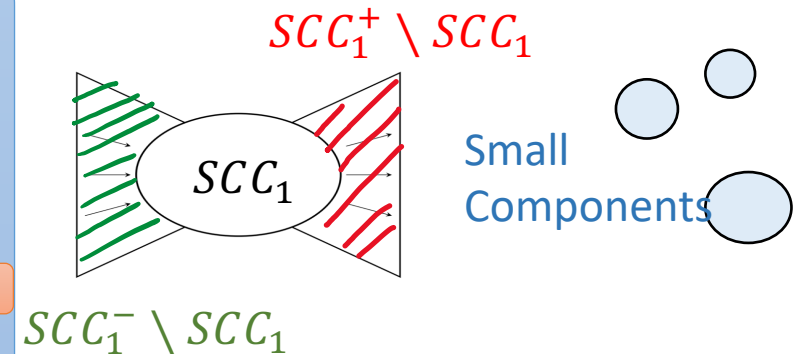
The number of super-spreaders and infected is local and always the same.

- For a uniform random node  $v$  whp  $|out(v) \setminus SCC_1^+| = o(n)$ , and  $|in(v) \setminus SCC_1^-| = o(n)$ .

An outbreak is inevitable when  $\epsilon n$  people are infected.

If  $p < p_c(G)$ , for a uniform random node  $v$  whp

$$\frac{|out(v)|}{n} \xrightarrow{\mathbb{P}} 0, \frac{|in(v)|}{n} \xrightarrow{\mathbb{P}} 0, \text{ and } \frac{|SCC_1|}{n} \xrightarrow{\mathbb{P}} 0.$$



## Recall. Strongly Connected Component (SCC)

A directed graph is SCC if there exists a directed path between any pairs of nodes.

# Bow-tie: From Undirected to Directed Cascade

## **Theorem. [A., Borgs, Saberi]**

Directed cascade on any sequence of possibly random graphs  $\{G_n\}$  satisfying:

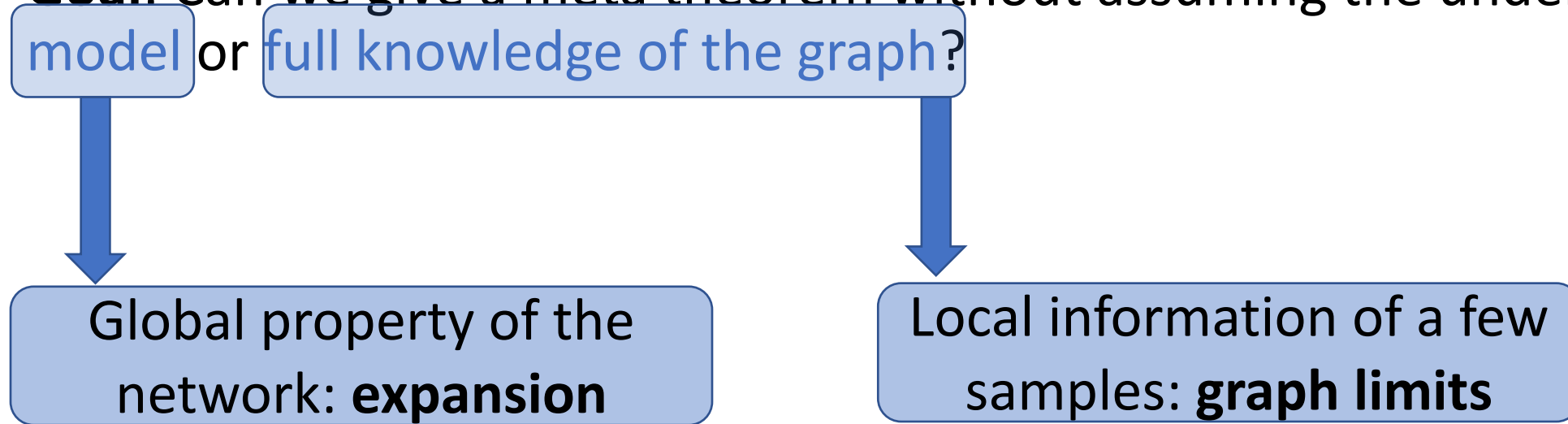
- a. there exists  $q \in (0, p]$  and a function  $\zeta: [p - q, p] \rightarrow [0, 1]$  that is left-continuous at  $p$  such that  $\frac{|C_1|}{n} \xrightarrow{\mathbb{P}} \zeta(p')$  for all  $p' \in [p - q, p]$ ;
- b.  $\frac{|C_2|}{n} \xrightarrow{\mathbb{P}} 0$  uniformly in  $[p - q, q]$ .

results in a bow-tie structure (as in Theorem 3).

**Takeaway :** The bow-tie structure holds when the giant in the undirected percolation is unique, and its relative size converges to its limit.

# Recap: What information about the network do we need to forecast an outbreak?

**Goal:** Can we give a meta theorem without assuming the underlying model or full knowledge of the graph?



# Proofs

# Critical Probability is Local

## Definition. (Critical Probability)

Given an infinite graph  $G$ , the critical  $p_c(G)$  is defined as

$$p_c(G) = \inf\{p \in [0,1]: \mathbb{P}_{G(p)}(\exists \text{ an infinite component in } G(p)) > 0\}.$$

## Theorem 1. [Benjamini, Nachmias, Peres '09]

Let  $G_n$  be a sequence of  $\alpha$ -expanders, with a uniform bounded degree  $d$ , and local weak limit  $G$ . If  $p < p_c(G)$ , then for any constant  $\beta > 0$ ,

$$\mathbb{P}(\exists \text{ a component of size at least } \beta n \text{ in } G_n(p)) \rightarrow 0 \text{ as } n \rightarrow \infty$$

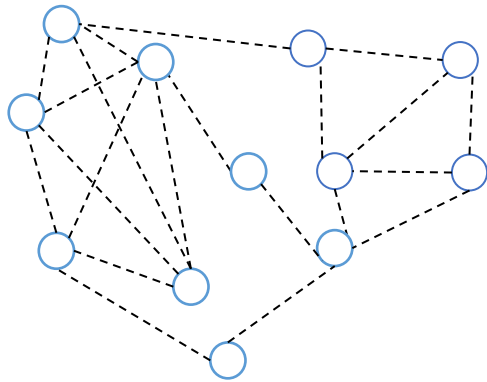
and if  $p > p_c(G)$ , then there exists a constant  $\beta > 0$  such that

$$\mathbb{P}(\exists \text{ a component of size at least } \beta n \text{ in } G_n(p)) \rightarrow 1 \text{ as } n \rightarrow \infty.$$

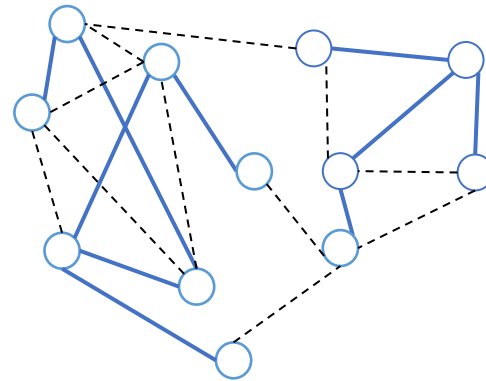
**Takeaway:** Critical probability in convergent expanders is local, and there's a phase transition at  $p_c(G)$ .

# Proof: Super Critical Case $p > p_c(G)$

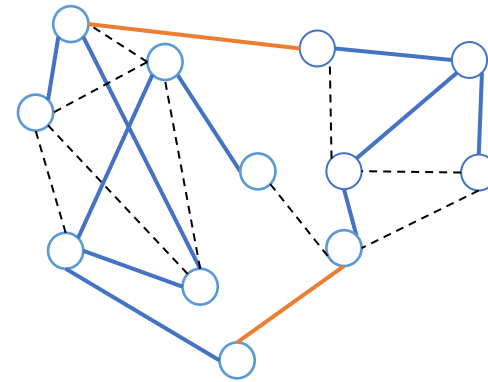
**Step 0:** For some  $\epsilon > 0$  let  $p_1 = p_c(G) + \epsilon$  be such that  $1 - p = (1 - p_1)(1 - \epsilon)$ . Consider two copies of percolation  $G_n(p_1)$  and  $G_n(\epsilon)$ . The union of them gives an instance of  $G_n(p)$ .



The original graph  $G_n$



$G_n(p_1)$



$G_n(\epsilon)$



# Proof: Super Critical Case $p > p_c(G)$

**Step 0:** For some  $\epsilon > 0$  let  $p_1 = p_c(G) + \epsilon$  be such that  $1 - p = (1 - p_1)(1 - \epsilon)$ . Consider two copies of percolation  $G_n(p_1)$  and  $G_n(\epsilon)$ . The union of them gives an instance of  $G_n(p)$ .

**Step 1:** There exists some  $\delta > 0$  such that for all  $K > 0$ , whp there are  $\delta n$  nodes with component larger than  $K$  in  $G_n(p_1)$ , i.e., let  $Z_K = \{\text{nodes with component larger than } K\}$  for all  $n \geq n_0$

$$\mathbb{P}_{G_n(p_1)} (|Z_K| \leq \delta n) \leq \exp\left(-\frac{\delta^2 n}{2d^{2k}}\right).$$

**Step 2 (Sprinkling):** There is a path in  $G_n(\epsilon)$  between any two large partition of components in  $Z_K$ :

$$\begin{aligned} & \mathbb{P}_{G_n(\epsilon)} \left( \exists A, B \subseteq Z_K: A, B \text{ disconnected in } G_n(\epsilon) \text{ and } G_n(p_1), |A|, |B| \geq \frac{\delta n}{3} \mid G_n(p_1) \right) \\ & \leq \exp(-nc_{\{\alpha, \delta, d, \epsilon\}}) \end{aligned}$$

**Step 3:**  $\mathbb{P}_{G_n(p)} \left( \text{contains a component of size } \frac{\delta n}{3} \right) \rightarrow 1, \text{ as } n \rightarrow \infty.$

# Step 1: Existence of relatively large components

**Step 1:** There exists some  $\delta > 0$  such that for all  $K > 0$ , whp there are  $\delta n$  nodes with component larger than  $K$  in  $G_n(p_1)$ , i.e., let  $Z_K = \{\text{nodes with component larger than } K\}$  for all  $n \geq n_0$

$$\mathbb{P}_{G_n(p_1)} (|Z_K| \leq \delta n) \leq \exp\left(-\frac{\delta^2 n}{2d^{2k}}\right).$$

- There exists  $\delta$  such that for all  $K > 0$ :

$$\mathbb{P}_{G(p)} (o \text{ connects to } K \text{ boundary}) \geq 4\delta.$$

- For any  $K$ , there exists  $n_0$  such that for all  $n \geq n_0$

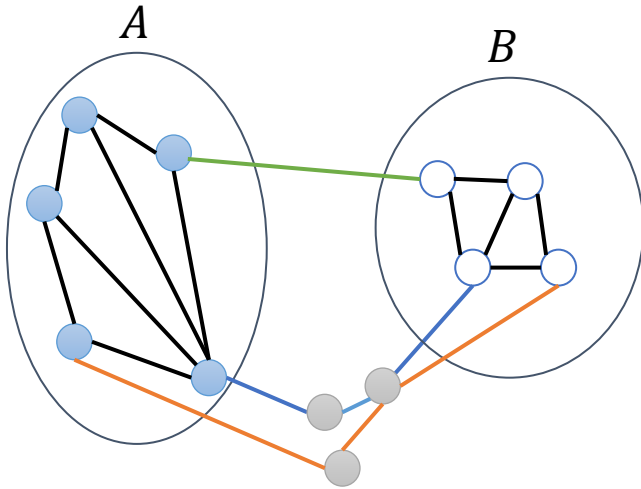
$$\mathbb{P}_{G_n(p)} (\text{a uniform random node is in } Z_K) \geq 2\delta.$$

- $\mathbb{E}[Z_K] \geq 2\delta n$
- Changing the status of an edge changes the membership of at most  $d^K$  nodes in  $Z_K$ .

# Step 2: Sprinkling

**Step 2 (Sprinkling):** There is a path in  $G_n(\epsilon)$  between any two large partition of components in  $Z_K$ :

$$\mathbb{P}_{G_n(\epsilon)} \left( \exists A, B \subseteq 2^{Z_K}: A, B \text{ disconnected in } G_n(\epsilon) \text{ and } G_n(p_1), |A|, |B| \geq \frac{\delta n}{3} \mid G_n(p_1) \right) \leq \exp(-nc_{\{\alpha, \delta, d, \epsilon\}})$$



**Menger's Theorem.** Let  $G$  be a finite undirected graph and  $A$  and  $B$  two disjoint set of vertices. Then the minimum edge-cut between  $A$  and  $B$  is equal to the number of pairwise edge-independent paths from  $A$  to  $B$ .

There are  $\frac{\delta \alpha n}{3}$  edge-disjoint paths in  $G_n$  between  $A$  and  $B$  (expansion). Since the average degree is bounded by  $d$ , the length of half of these paths is bounded by  $\ell = \frac{6d}{\delta \alpha}$ . (# paths =  $\frac{\delta \alpha n}{6}$ ) Each path appear in  $G_n(\epsilon)$  with probability  $\epsilon^\ell$ .

The probability that non of the paths appear in  $G_n(\epsilon)$  :  $(1 - \epsilon^\ell)^{\#paths}$

Number of  $A, B$  partitions in  $G_n(p_1)$  :  $2^{\frac{n}{K}}$

$$\text{Finally: } 2^{\frac{n}{K}} \left(1 - \epsilon^{\frac{6d}{\delta \alpha}}\right)^{\frac{\delta \alpha n}{6}} \leq \exp\left(n\left(\frac{1}{K} - \frac{\delta \alpha}{6} \epsilon^{\frac{6d}{\delta \alpha}}\right)\right)$$

# Brief History of Sprinkling

[Erdős, Rényi'60]

[Posa'76][Ajtai, Kolmós, Szemerédi '82]

[Bollobás, Riordan '01] [Alon, Benjamini, Stacey '02]

[Borgs, Chayes, van der Hofstad, Slade, Spencer '07]

**[Benjamini, Nachmias, Peres '09]**

[Janson, Rucinski'10] [van der Hofstad, Nachmias '17]

[Krivelevich, Sudakov '17]

[Dudek, C. Reiher, A. Ruci'nski, and M. Schacht '20]

[Nenadov, Trujic '21][Easo, Hutchcroft '21]

# Relative Size of the Giant in Expanders

## Theorem 2. [A., Borgs, Saberi '21]

Let  $\{G_n\}_{n \in \mathbb{N}}$  be a sequence of (possibly random) large-set expanders with bounded average degree converging locally in probability to  $(G, o) \in \mathfrak{G}_*$  with non-random distribution  $\mu$ . Let  $C_i$  be the  $i^{\text{th}}$  largest component. If  $p \neq p_c(\mu)$ ,

$$\frac{|C_1|}{n} \xrightarrow{\mathbb{P}} \zeta(p),$$

Also for all  $p \in [0, 1]$ ,  $\frac{|C_2|}{n} \xrightarrow{\mathbb{P}} 0$ .

$\xrightarrow{\mathbb{P}}$ : convergence in probability in percolation and  $\mu$ .

$\zeta(p) := \mathbb{E}_{(G, o) \sim \mu} [\mathbb{P}_{G(p)}(|\text{connected component of } o| = \infty)]$ .

**Takeaway:** Giant in convergent expanders is unique, and its size converges to its limit.

# Large-set Expanders

## Definition. (Expander)

$G$  is  $\alpha$ -expander if  $\phi(G) \geq \alpha$ , where

$$\phi(G) = \min_{S \subseteq V(G)} \frac{E(S, \bar{S})}{\min(|S|, |\bar{S}|)}$$

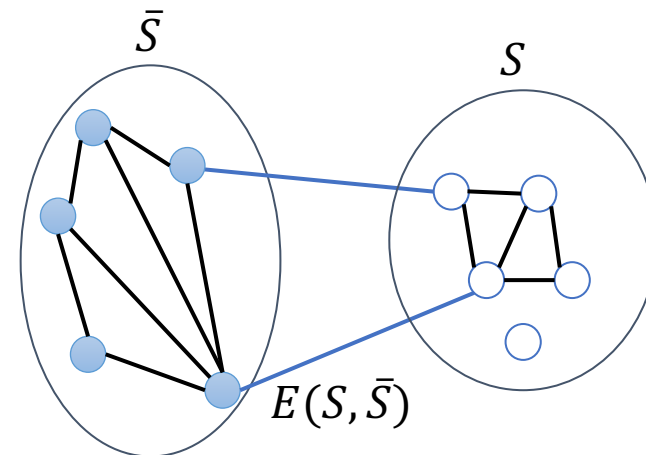
## Definition. (Large-set Expander)

$G$  with average degree bounded by  $d$  is  $(\alpha, \epsilon, d)$  large-set expander if  $\phi_\epsilon(G) \geq \alpha$ , where

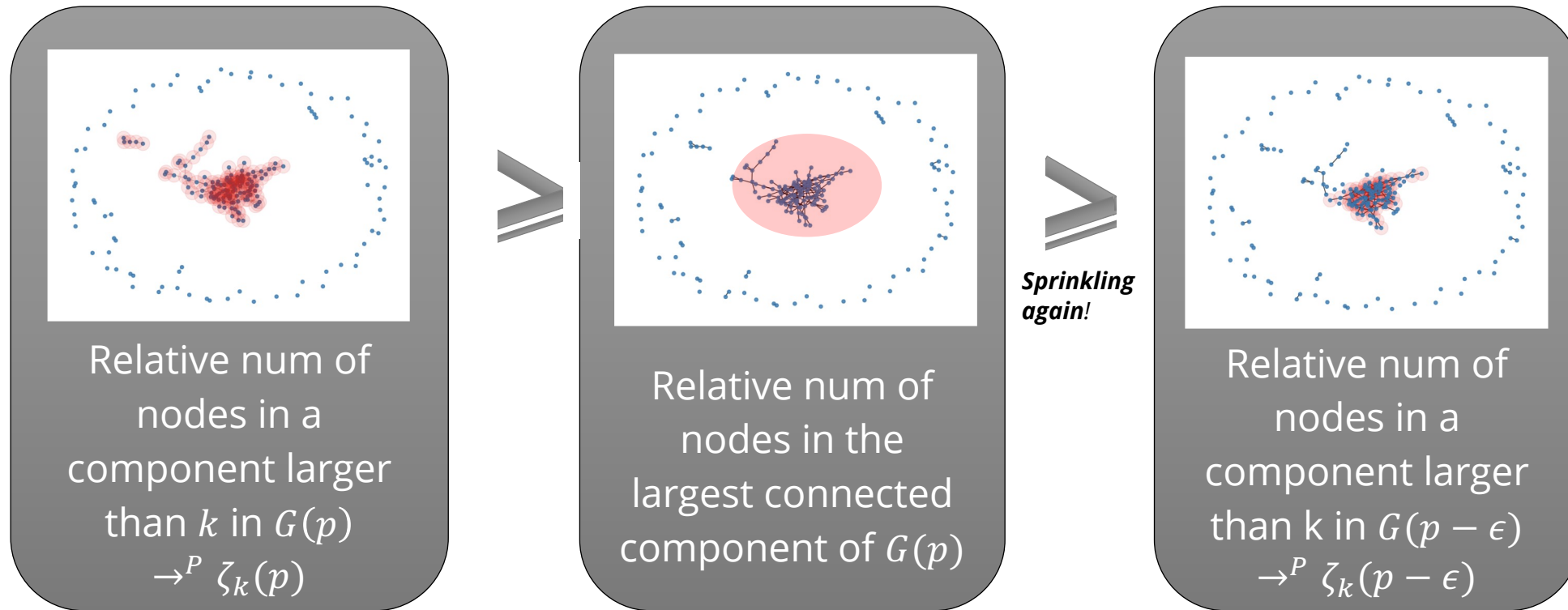
$$\phi_\epsilon(G) = \min_{\substack{S \subseteq V(G) \\ |S| \geq \epsilon n}} \frac{E(S, \bar{S})}{\min(|S|, |\bar{S}|)}$$

## Definition. (Sequence of Large-set Expander)

A sequence of possibly random graphs  $\{G_n\}_{n \in \mathbb{N}}$  is called a large-set expander sequence with bounded average degree, if there exists  $\bar{d} < \infty$  and  $\alpha > 0$  such that for all  $\epsilon \in (0, .5)$ , the probability that  $G_n$  is an  $(\alpha, \epsilon, d)$  large-set expander goes to 1 as  $n \rightarrow \infty$ .



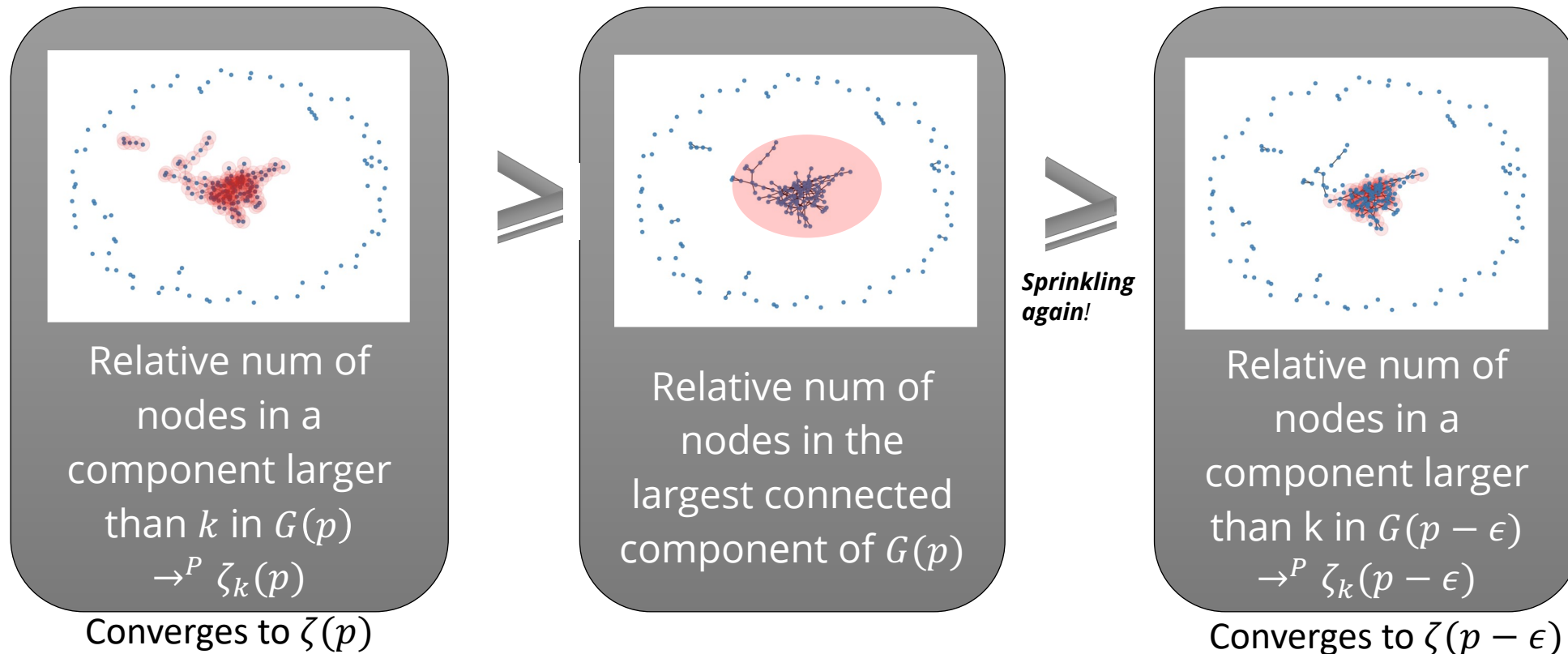
# Proof Sketch: Size of the Giant Converges



$$\zeta_k(p) := \mathbb{E}_{(G,o) \sim \mu} [\mathbb{P}_{G(p)}(|\text{connected component of } o| \geq k)].$$

$$\lim_{k \rightarrow \infty} \zeta_k(p) = \zeta(p).$$

# Proof Sketch: Size of the Giant Converges



**Lemma.** For a sequence of graphs satisfying the assumptions of Theorem 2,  $\zeta(p)$  is continuous for all  $p \neq p_c(\mu)$ . Equivalently, the limit  $\mu$  is ergodic.

(Sourav Sarkar proved this lemma for deterministic sequence of convergent expanders in 2018.)



# Directed Percolation Creates a Bow-Tie

## Theorem 3. [A., Borgs, Saberi]

For a sequence of large-set expanders as in Theorem 2,  $\frac{|SCC_2|}{n} \xrightarrow{\mathbb{P}} 0$ .

Also, if  $p > p_c(G)$ :

- $\liminf_{n \rightarrow \infty} \frac{|SCC_1|}{n} \geq \zeta^2(p)$  and  $\frac{|SCC_1|}{\mathbb{E}|SCC_1|} \xrightarrow{\mathbb{P}} 1$ .

Critical probability is the same as undirected.

- $\frac{1}{n} |SCC_1^+| \xrightarrow{\mathbb{P}} \zeta(p)$  and  $\frac{1}{n} |SCC_1^-| \xrightarrow{\mathbb{P}} \zeta(p)$

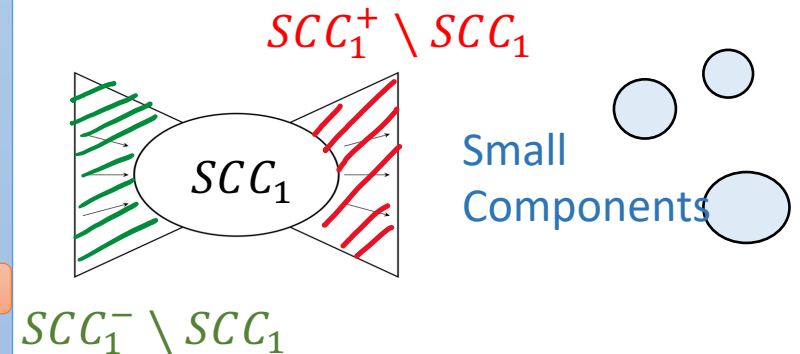
The number of super-spreaders and infected is local and always the same.

- For a uniform random node  $v$  whp  $|out(v) \setminus SCC_1^+| = o(n)$ , and  $|in(v) \setminus SCC_1^-| = o(n)$ .

An outbreak is inevitable when  $\epsilon n$  people are infected.

If  $p < p_c(G)$ , for a uniform random node  $v$  whp

$$\frac{|out(v)|}{n} \xrightarrow{\mathbb{P}} 0, \frac{|in(v)|}{n} \xrightarrow{\mathbb{P}} 0, \text{ and } \frac{|SCC_1|}{n} \xrightarrow{\mathbb{P}} 0.$$



## Recall. Strongly Connected Component (SCC)

A directed graph is SCC if there exists a directed path between any pairs of nodes.

# Proof Idea 1: Coupling to Undirected

## Lemma. (informal)

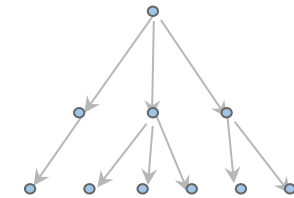
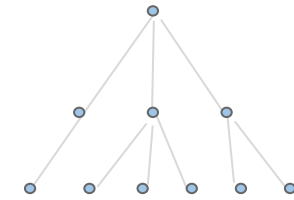
For a fixed, or random vertex  $v$

$$\mathbb{P}_{D_n(p)}(|Out(v)| \geq k) = \mathbb{P}_{G_n(p)}(|C(v)| \geq k)$$

Coupling two trees and using uniqueness of the giant in  $G_n(p)$  one can derive

$$|SCC_2| = o(n)$$

for convergent sequence of large-set expanders.



# Proof Idea: Strongly Connected Component

## Step1. (Lower bound on $\mathbb{E}[|SCC_1|^2]$ )

Couplings, plus FKG gives lower bound on expectation of

$$\sum_i |SCC_i|^2 = \sum_x |SCC(x)|$$

Use  $|SCC_2| = o(n)$  to get a lower bound on  $\mathbb{E}|SCC_1|^2$ .

## Step2. (Upper bound on $\text{var}(|SCC_1|)$ )

Uses the sharpening of the Efron-Stein bounds by [Falik and Samorodnitsky '07]

Bounding the influence of an edge on the size of  $|SCC_1|$

# Takeaways

On converging expanders:

Critical probability, and the size of the giant converges to its limit.

We give an algorithm to estimate the limit.

Directed percolation in convergent expanders has a bow-tie structure.

Graph limits enables us to connect the discrete world to the continuous world. Can we find more applications?

# References

Alon, Benjamini, Stacey, *“Percolation on finite graphs and isoperimetric inequalities”* (2002)

Benajmini, Nachmias, Peres, *“Is the critical percolation probability local?”* (2009)

Alimohammadi, Borgs, Saberi, *“Algorithms Using Local Graph Features to Predict Epidemics”* (2022)

Alimohammadi, Borgs, Saberi, *“Locality of Random Diagraphs on Expanders”* (2021)

Thank You

yeganeh@stanford.edu