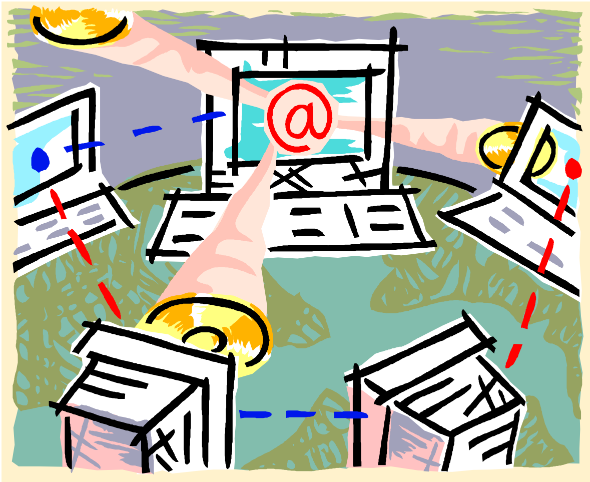
Probability Models of Information Exchange on Networks

Lecture 6

Elchanan Mossel
UC Berkeley



Many Other Models

• There are many models of information exchange on networks.
• Q: Which model to chose?

• My answer – good features of models include:

• “Canonical” models.
• Amenable to analysis.
• Studied intensively before.

• Ok to invent your own models.
• Models are always just models …

Ariel Rubinstein on theoretical economics

- When talking about economics:
- “Everything I say is personal, based upon the entire range of my life experience which also includes the fact that professionally I engage in economics theory. However, to the best of my understanding, economic theory has nothing to do say about the heart of the issue under discussion here. I am not sure I know what an opinion is. I am not attempting to predict the rate of inflation tomorrow …”

Some other natural models

- Growth models: percolation models, DLA etc.

- Competition models: Competing growth.

- Infection models: Contact process, SIR, SIS …

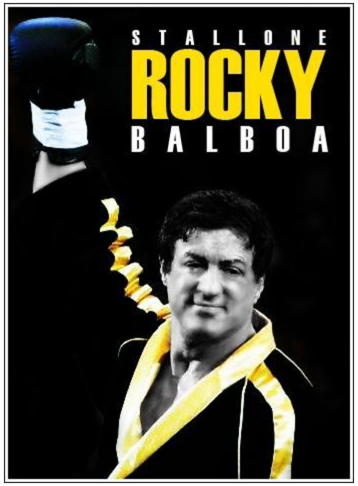
- Aspects of modeling:
- Dynamic networks
- Random networks
- …

- Today: two examples of percolation based processes.

Example1: models of collective behavior

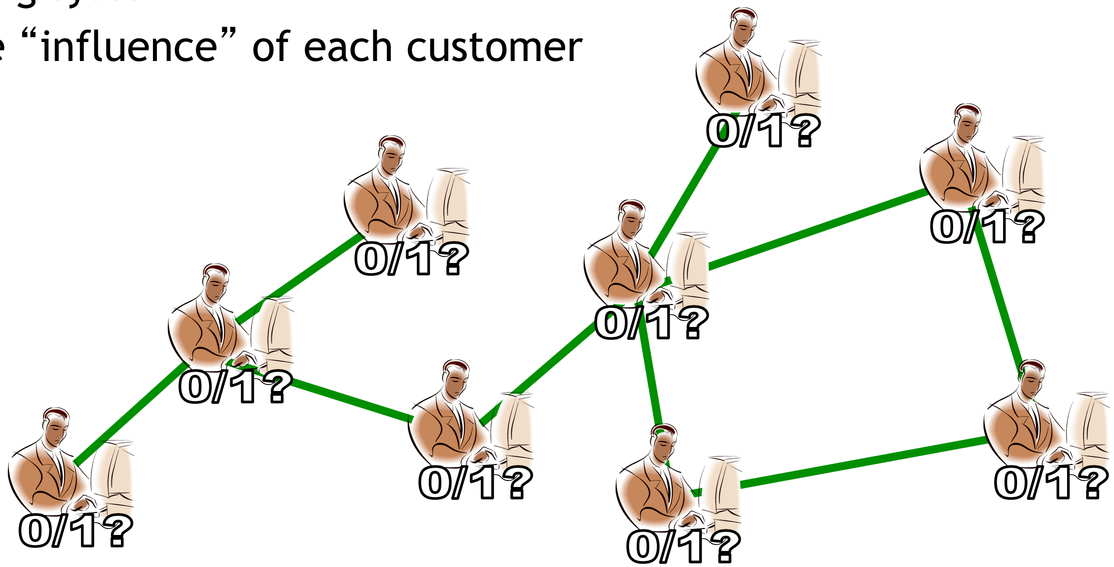
- examples:
  - joining a riot
  - adopting a product
  - going to a movie

- model features:
  - binary decision
  - cascade effect
  - network structure



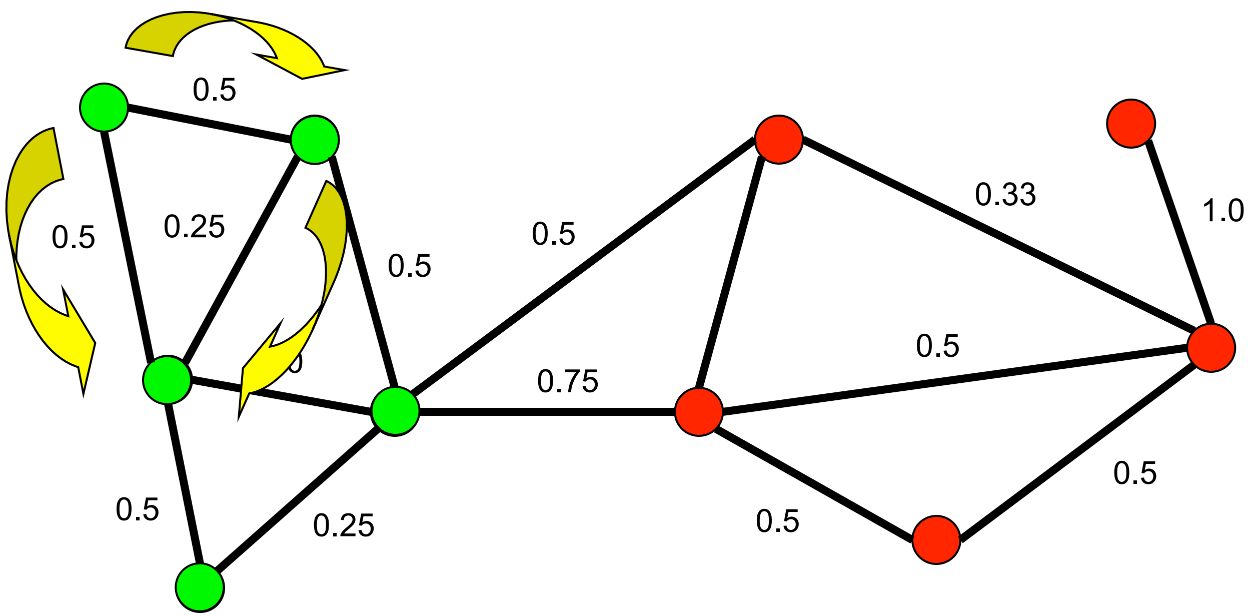
viral marketing

- referrals, word-of-mouth can be very effective
  - ex.: google+
- viral marketing
  - goal: **mining the network value** of potential customers
  - how: **target a small set of trendsetters**, seeds
- example [Domingos-Richardson' 02]
  - collaborative filtering system
  - use MRF to compute "influence" of each customer



independent cascade model

- when a node is activated
  - it gets one chance to activate each neighbour
  - probability of success from u to v is $p_{u,v}$



generalized models

- graph $G=(V,E)$; initial activated set S_0

- generalized threshold model [Kempe-Kleinberg-Tardos' 03,' 05]
  - activation functions: $f_u(S)$ where S is set of activated nodes
  - threshold value: θ_u uniform in $[0,1]$
  - dynamics: at time t,set S_t to S_{t-1} and add all nodes with $f_u(S_{t-1}) \geq \theta_u$
    (note the process stops after (at most) $n-1$ steps)
- generalized cascade model [KKT' 03,' 05]
  - when node u is activated:
    - gets one chance to activate each neighbours
    - probability of success from u to v: $p_u(v,S)$ where S is set of nodes who have already tried (and failed) to activate u
  - assumption: the $p_u(v,.)$' s are "order-independent"

- **theorem** [KKT' 03] - the two models are equivalent

influence maximization

- **definition** - the **influence** $\sigma(S)$ given the initial seed S is the expected size of the infected set at termination

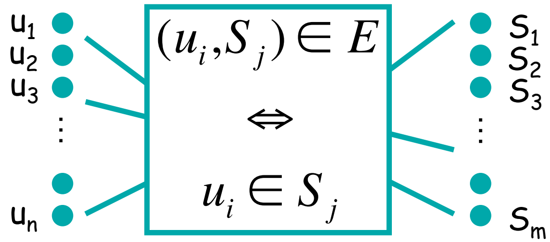
$$\sigma(S) = \mathrm{E}_S\left[\left|S_{n-1}\right|\right]$$

- **definition** - in the **influence maximization problem (IMP)**, we want to find the seed S of fixed size k that maximizes the influence

$$S^* = \arg\max\left\{\sigma(S) : S \subseteq V,\ |S| = k\right\}$$

- **theorem** [KKT' 03] - the IMP is **NP-hard**
  - reduction from *Set Cover*: ground set $U = \{u_1,\ldots,u_n\}$ and collection of cover subsets $S_1,\ldots,S_m$

independent cascade model



$$\exists S,\ |S| = k,\ \sigma(S) \ge n + k\,?$$

submodularity

- **definition** - a set function $f : V \rightarrow R$ is **submodular** if for all A, B in V

$$f(A) + f(B) \geq f(A \cap B) + f(A \cup B)$$

- example: $f(S) = g(|S|)$ where g is concave

- interpretation: “discrete concavity” or “diminishing returns”, indeed submodularity equivalent to

$$\forall S \subseteq T, \forall v \in V, \quad f(T \cup \{v\}) - f(T) \leq f(S \cup \{v\}) - f(S)$$

- threshold models:
  - it is natural to assume that the **activation functions have diminishing returns**
  - supported by observations of [Leskovec-Adamic-Huberman’06] in the context of viral marketing

main result

- **theorem** [M-Roch' 06; first conjectured in KKT' 03] - in the generalized threshold model, if all activation functions are monotone and submodular, then the **influence is also submodular**

- **corollary** [M-Roch' 06] - IMP admits a $(1 - e^{-1} - \varepsilon)$-approximation algorithm (for all $\varepsilon > 0$)
  - this follows from a general result on the approximation of submodular functions [Nemhauser-Wolsey-Fisher' 78]

- known special cases [KKT' 03, ' 05]:
  - linear threshold model, independent cascade model
  - decreasing cascade model, "normalized" submodular threshold model

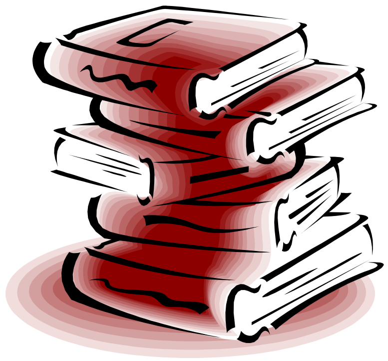
$$\forall S \subseteq T,\ p_u(v,S) \geq p_u(v,T) \text{ or equiv. } \frac{f_u(S \cup \{v\}) - f_u(S)}{1 - f_u(S)} \geq \frac{f_u(T \cup \{v\}) - f_u(T)}{1 - f_u(T)}$$

related work

- sociology
  - threshold models: [Granovetter'78], [Morris'00]
  - cascades: [Watts'02]

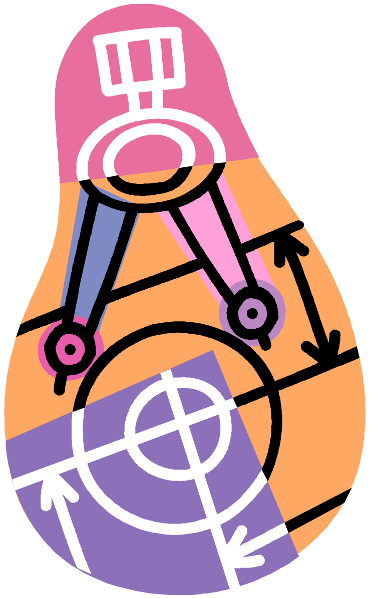
- data mining
  - viral marketing: [KKT'03,'05], [Domingos-Richardson'02]
  - recommendation networks: [Leskovec-Singh-Kleinberg'05], [Leskovec-Adamic-Huberman'06]

- economics
  - game-theoretic point of view: [Ellison'93], [Young'02]



- probability theory
  - Markov random fields, Glauber dynamics
  - percolation
  - interacting particle systems: voter model, contact process

proof sketch



coupling

- we use the generalized threshold model
- arbitrary sets A, B; consider 4 processes:
  - (A_t) started at A
  - (B_t) started at B
  - (C_t) started at $A \cap B$
  - (D_t) started at $A \cup B$
- it suffices to **couple the 4 processes** in such a way that for all t

$$C_t \subseteq A_t \cap B_t$$
$$D_t \subseteq A_t \cup B_t$$

- indeed, at termination

$$|A_{n-1}| + |B_{n-1}| \geq |A_{n-1} \cap B_{n-1}| + |A_{n-1} \cup B_{n-1}| \geq |C_{n-1}| + |D_{n-1}|$$

(note this works with $|.|$ replaced with any w monotone, submodular)

proof ideas

- our goal:

$$C_t \subseteq A_t \cap B_t \quad (1) \qquad D_t \subseteq A_t \cup B_t \quad (2)$$

- antisense coupling
  - obvious way to couple: use same θ_u's for all 4 processes
  - satisfies (1) but not (2)
  - "antisense": using θ_u for (A_t) and $(1-\theta_u)$ for (B_t) "maximizes union"
  - we combine both couplings
- piecemeal growth
  - seed sets can be introduced in stages
  - we add $A \cap B$ then $A \setminus B$ and finally $B \setminus A$
- need-to-know
  - not necessary to pick all θ_u's at beginning
  - can unveil only what we need to know:

$$\theta_v \in \left[f_v(S_{t-2}), f_v(S_{t-1}) \right]?$$

piecemeal growth

- process started at S: (S_t)
- **partition** of S: $S^{(1)},\ldots,S^{(K)}$
- consider the process (T_t):
  - pick θ_u's
  - run the process with seed $S^{(1)}$ until termination
  - add $S^{(2)}$ and continue until termination
  - add $S^{(3)}$ and so on

- **lemma** - the sets S_{n-1} and T_{Kn-1} have the same distribution

antisense coupling

- disjoint sets: S, T
- partition of S: $S^{(1)},\ldots,S^{(K)}$
- piecemeal process with seeds $S^{(1)},\ldots,S^{(K)},T$: (S_t)

- consider the process (T_t):
  - pick θ_u's
  - run piecemeal process with seeds $S^{(1)},\ldots,S^{(K)}$ until termination
  - add T and continue with **threshold values**

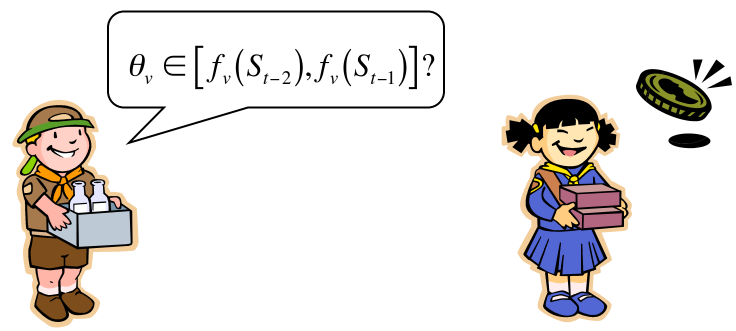
$$\theta_v' = 1 - \theta_v + f_v\left(T_{Kn-1}\right)$$

- **lemma** - the sets $S_{(K+1)n-1}$ and $T_{(K+1)n-1}$ have the same distribution

need-to-know

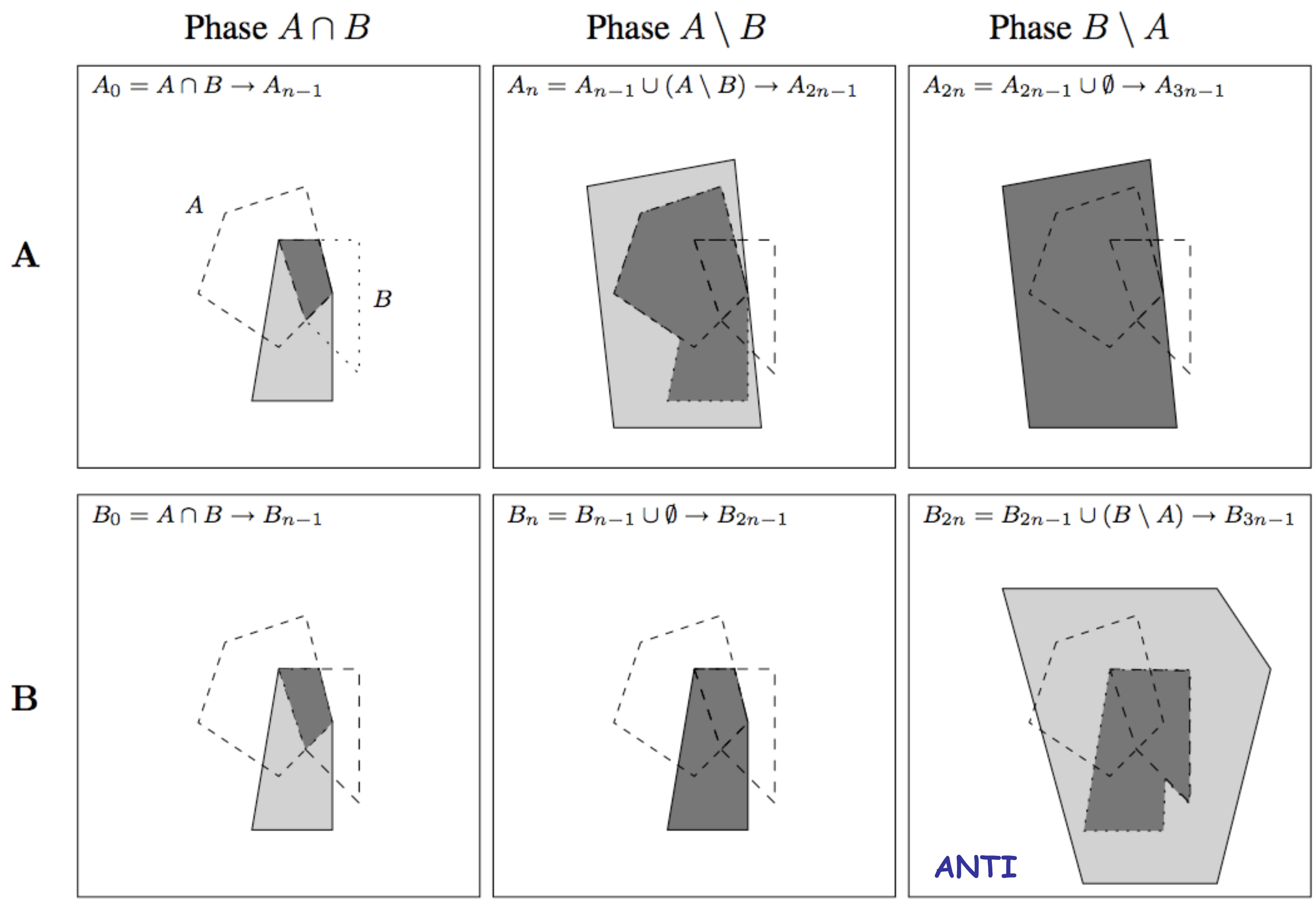
- proof of lemma
  - run the first K stages identically in both processes
  - note that for all v **not in** $S_{Kn-1} = T_{Kn-1}$, θ_v is uniformly distributed in $[f_v(T_{Kn-1}),1]$
  - but $\theta_v' = 1 - \theta_v + f_v(T_{Kn-1})$ has the same distribution

simulation 1



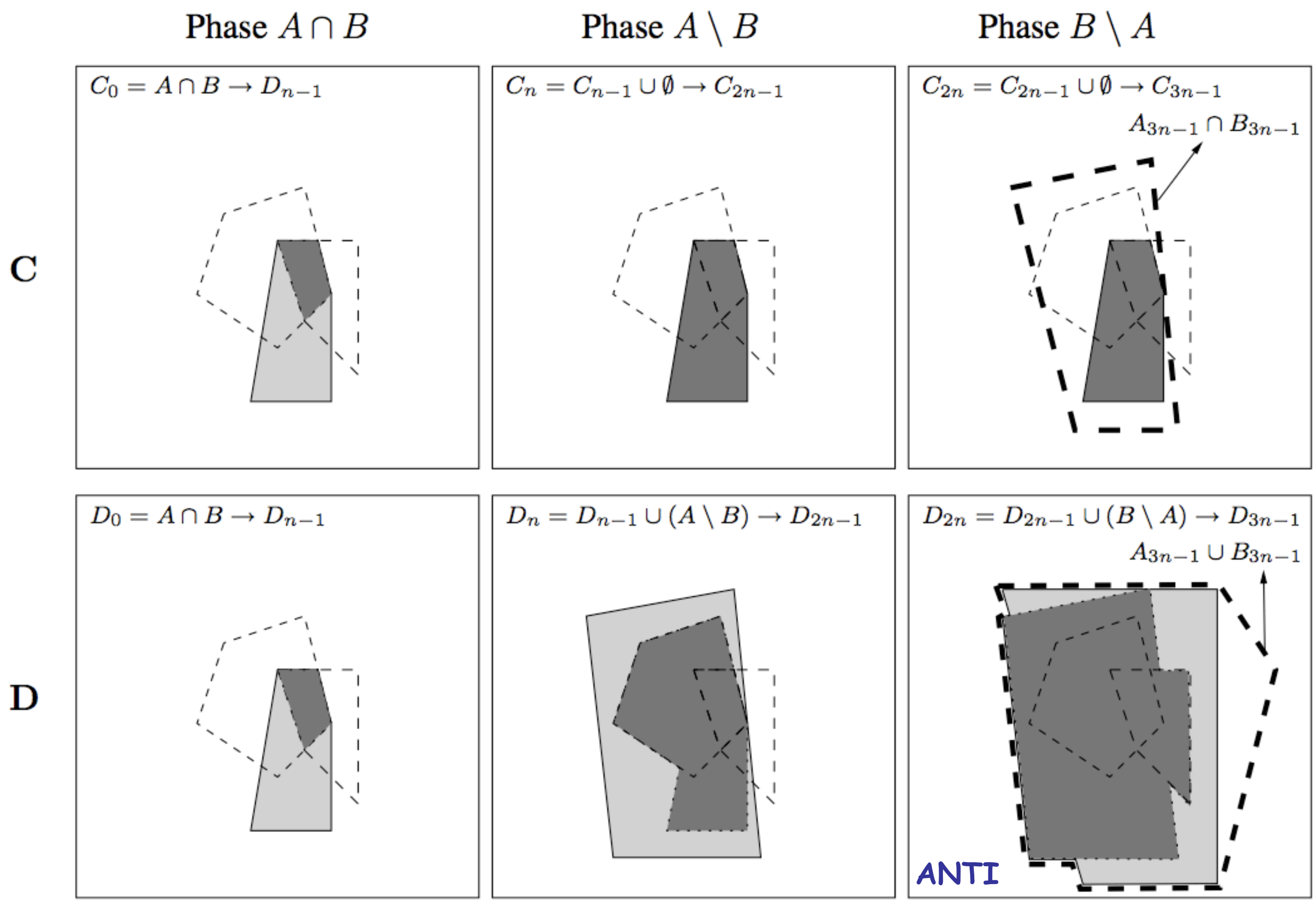
simulation 2

proof I



Phase $A \cap B$
Phase $A \setminus B$
Phase $B \setminus A$
A
B
$A_0 = A \cap B \rightarrow A_{n-1}$
$A_n = A_{n-1} \cup (A \setminus B) \rightarrow A_{2n-1}$
$A_{2n} = A_{2n-1} \cup \emptyset \rightarrow A_{3n-1}$
$B_0 = A \cap B \rightarrow B_{n-1}$
$B_n = B_{n-1} \cup \emptyset \rightarrow B_{2n-1}$
$B_{2n} = B_{2n-1} \cup (B \setminus A) \rightarrow B_{3n-1}$
A
B
ANTI

proof II


Phase $A \cap B$
Phase $A \setminus B$
Phase $B \setminus A$
C
D
$C_0 = A \cap B \to D_{n-1}$
$C_n = C_{n-1} \cup \emptyset \to C_{2n-1}$
$C_{2n} = C_{2n-1} \cup \emptyset \to C_{3n-1}$
$A_{3n-1} \cap B_{3n-1}$
$D_0 = A \cap B \to D_{n-1}$
$D_n = D_{n-1} \cup (A \setminus B) \to D_{2n-1}$
$D_{2n} = D_{2n-1} \cup (B \setminus A) \to D_{3n-1}$
$A_{3n-1} \cup B_{3n-1}$
ANTI

proof III

- new processes have correct final distribution
- up to time $2n-1$, $B_t = C_t$ and $A_t = D_t$ so that

$$C_t \subseteq A_t \cap B_t \qquad D_t \subseteq A_t \cup B_t$$

- for time $2n$, note that

$$B_{2n-1} \subseteq D_{2n-1}$$
$$B_{2n} = B_{2n-1} \cup (T \setminus S) \qquad D_{2n} = D_{2n-1} \cup (T \setminus S)$$

- so by **monotonicity** and **submodularity**

$$f_v(B_{2n}) - f_v(B_{2n-1}) \geq f_v(D_{2n}) - f_v(D_{2n-1})$$

- then proceed by induction

general result

- we have proved:

  theorem [Mossel-R' 06] - in the generalized threshold model, if all activation functions are submodular, then for any monotone, submodular function w, the generalized influence

  $$\sigma_w(S) = \mathrm{E}_S[w(S_{n-1})]$$

  is submodular
- Note: A closure property for sub-modular functions!

Competing first passage percolation on random regular graphs

Elchanan Mossel

University of California, Berkeley

July 24, 2013

Based on a joint work with Tonći Antunović Yael Dekel, Elchanan Mossel and Yuval Peres

First passage percolation

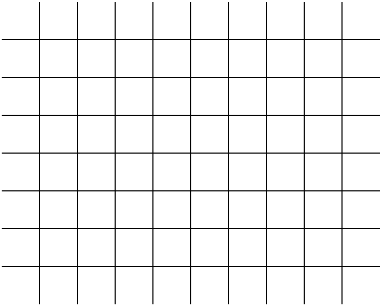
First passage percolation:
Fix a graph $G = (V, E)$, consider iid edge lengths $(\ell_e)_{e \in E}$.
Define the random metric on V

$$d(x, y) = \inf_{\Gamma} \ell(\Gamma),$$

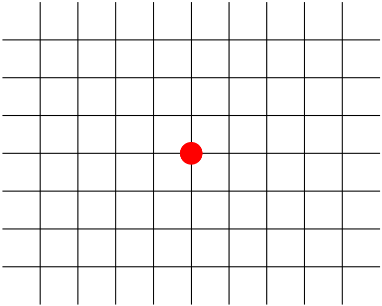
where the infimum is taken over all paths Γ connecting x and y and $\ell(\Gamma)$ is the sum of lengths of the edges on Γ.

An important case: $\ell_e \sim \exp(\lambda)$.
Process $r \mapsto B(0, r)$ evolves as a Markov process, new vertices are added at the rate $\lambda \times$ the number of neighbors in $B(0, r)$.

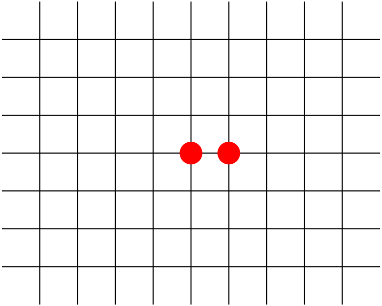
First passage percolation



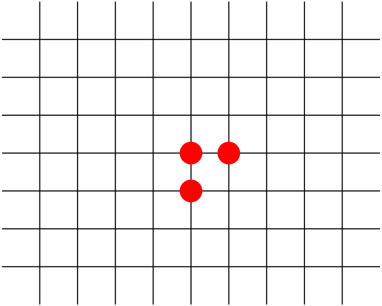
First passage percolation



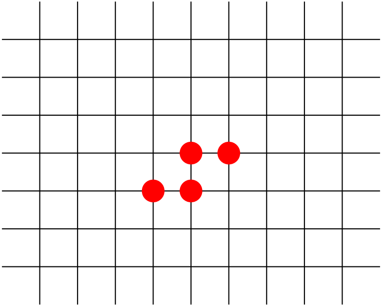
First passage percolation



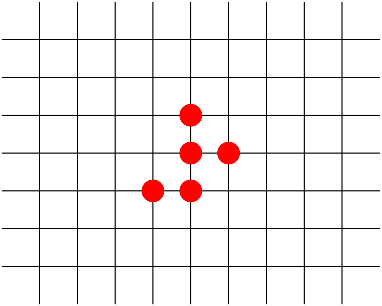
First passage percolation



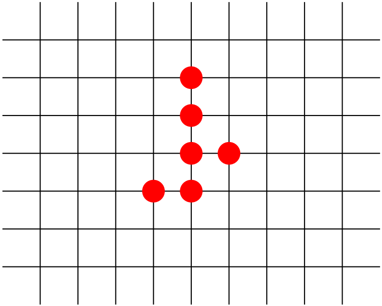
First passage percolation



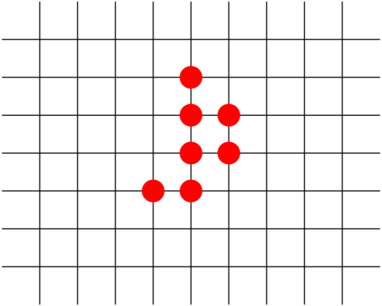
First passage percolation



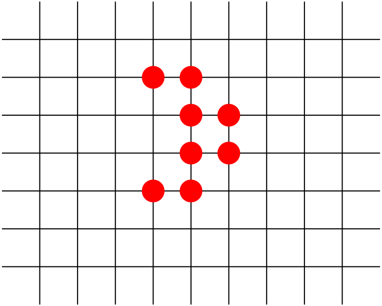
First passage percolation



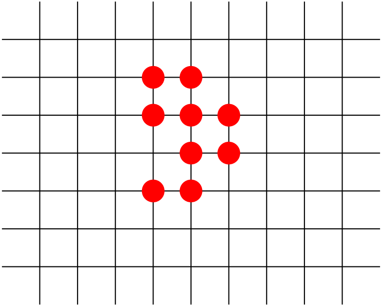
First passage percolation



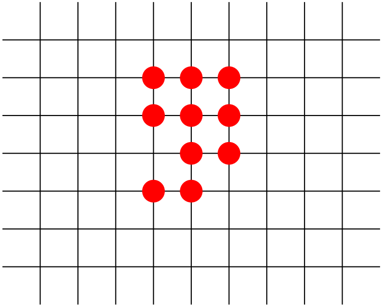
First passage percolation



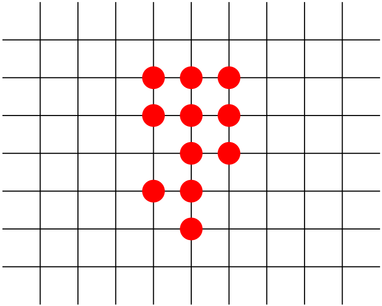
First passage percolation



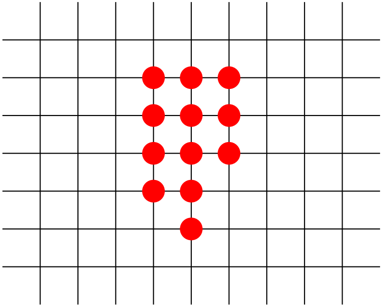
First passage percolation



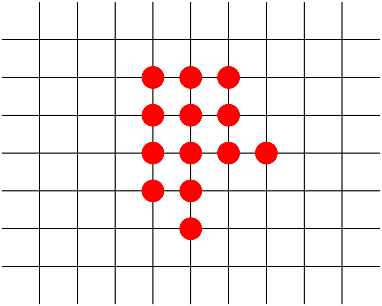
First passage percolation



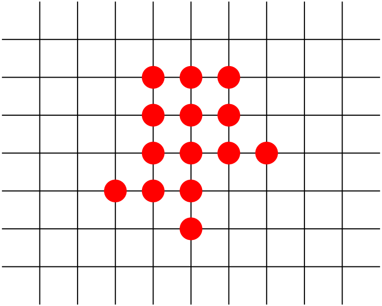
First passage percolation



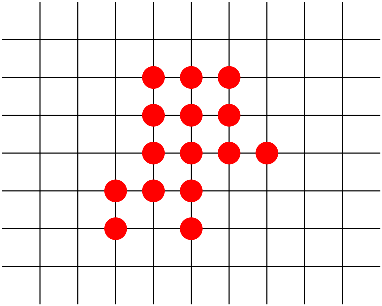
First passage percolation



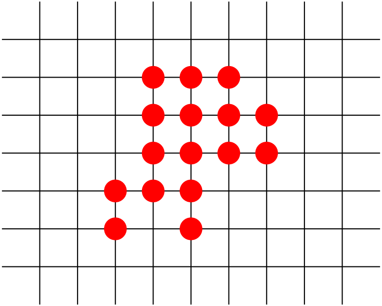
First passage percolation



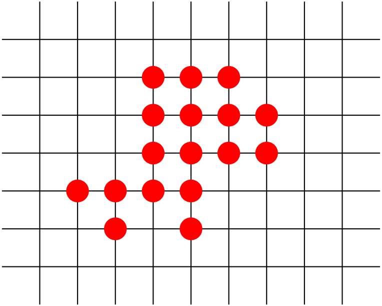
First passage percolation



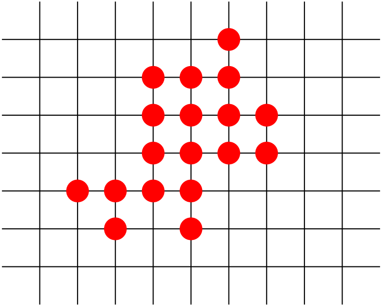
First passage percolation



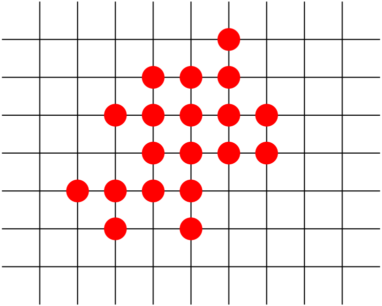
First passage percolation



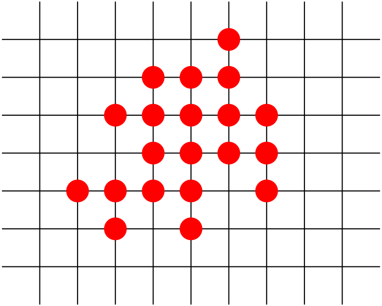
First passage percolation



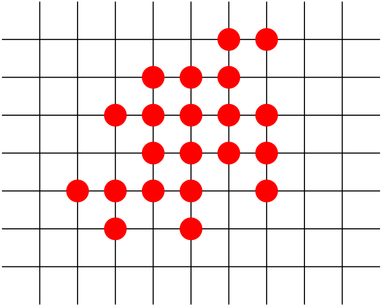
First passage percolation



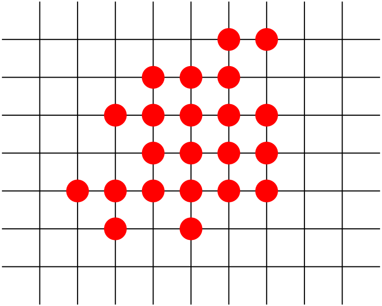
First passage percolation



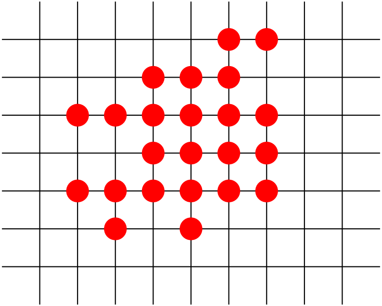
First passage percolation



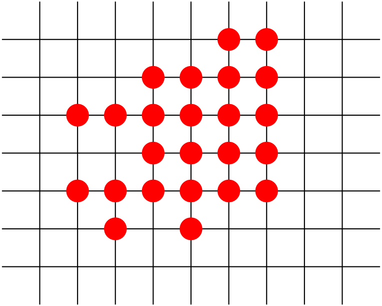
First passage percolation



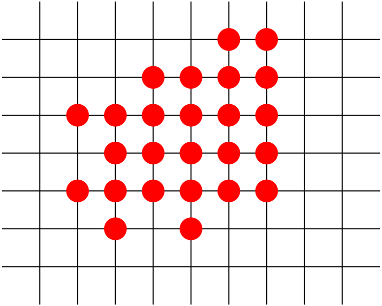
First passage percolation



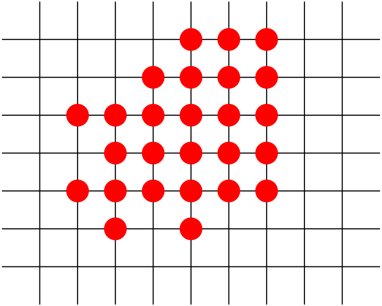
First passage percolation



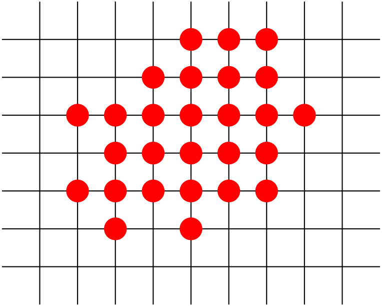
First passage percolation



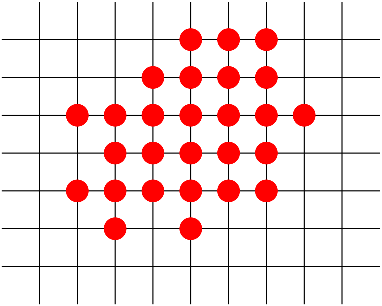
First passage percolation



First passage percolation



First passage percolation



Theorem (Cox-Durrett shape theorem)

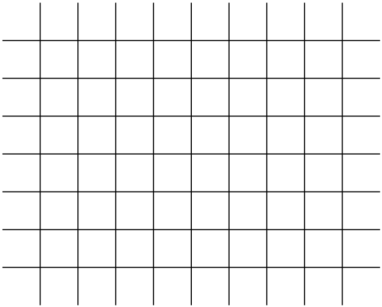
There exists a compact convex set A such that for any $\delta > 0$

$$\lim_{r \to \infty} \mathbb{P}\big((1-\delta) rA \subset B(0,r) \subset (1+\delta) rA\big) = 1.$$

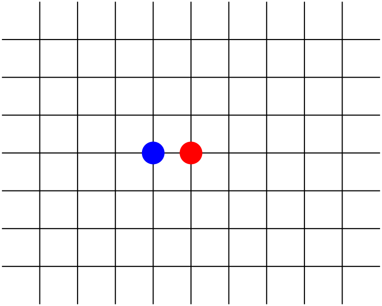
Competing first passage percolation (also called Two-type Richardson Model by Häggström, Pemantle):

- Start with one red vertex and one blue vertex, other uncolored.
- Uncolored vertices become red at the rate ($\lambda_R \times$ the number of red neighbors) and blue at the rate ($\lambda_B \times$ the number of blue neighbors).
- Once colored, vertices never change the color.

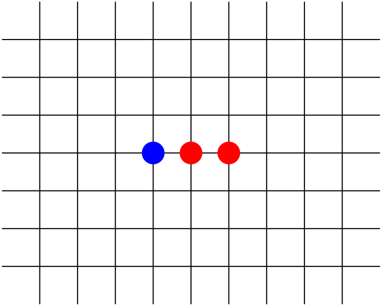
Competing first passage percolation



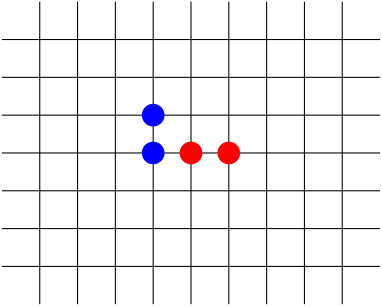
Competing first passage percolation



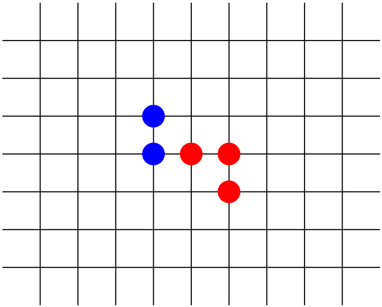
Competing first passage percolation



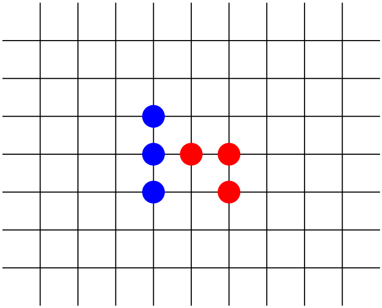
Competing first passage percolation



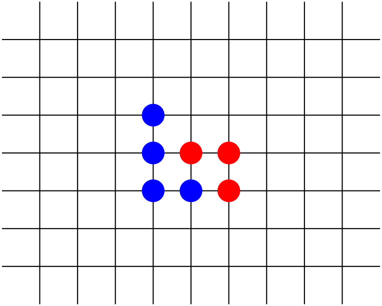
Competing first passage percolation



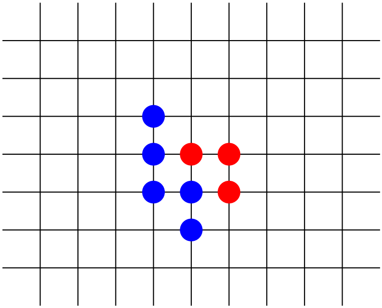
Competing first passage percolation



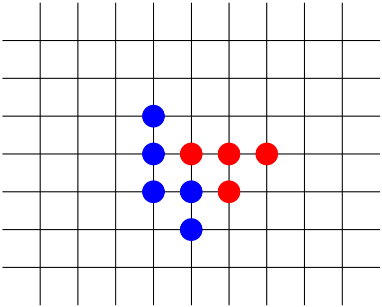
Competing first passage percolation



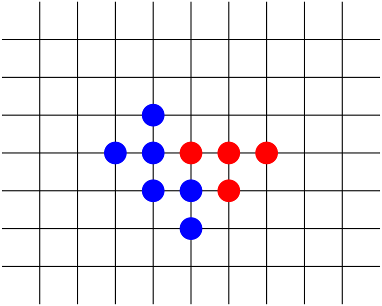
Competing first passage percolation



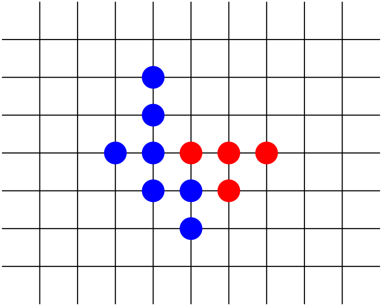
Competing first passage percolation



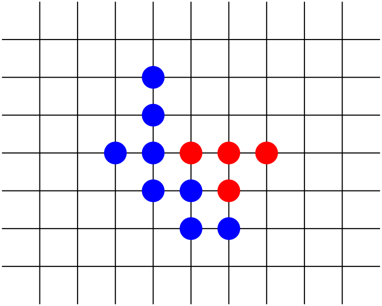
Competing first passage percolation



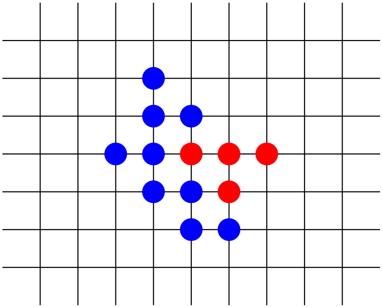
Competing first passage percolation



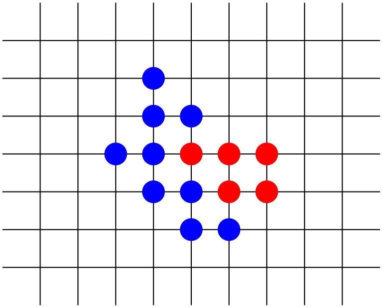
Competing first passage percolation



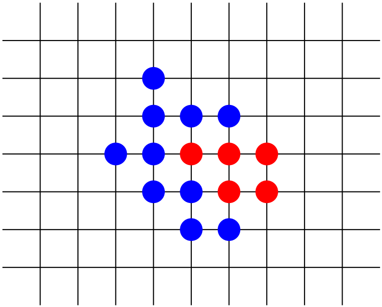
Competing first passage percolation



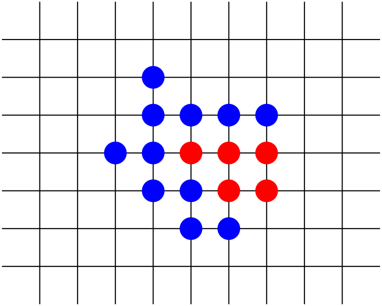
Competing first passage percolation



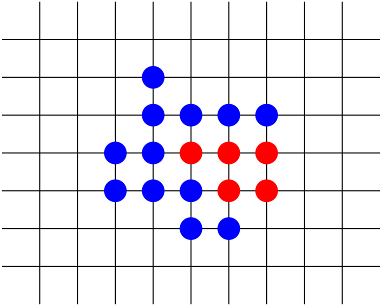
Competing first passage percolation



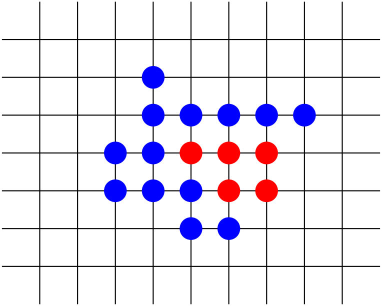
Competing first passage percolation



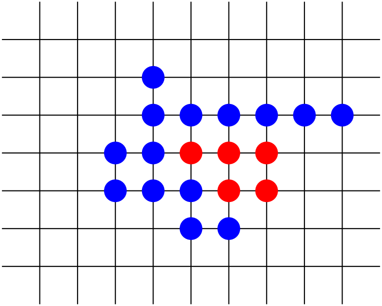
Competing first passage percolation



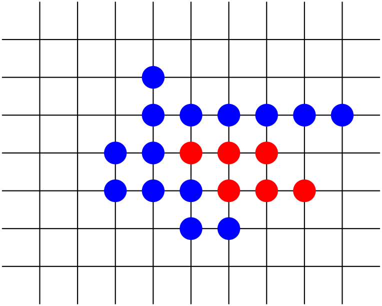
Competing first passage percolation



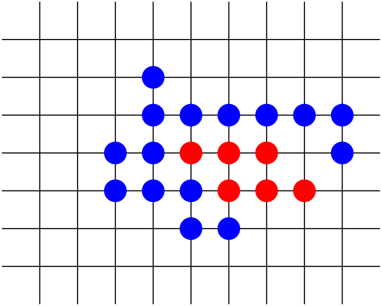
Competing first passage percolation



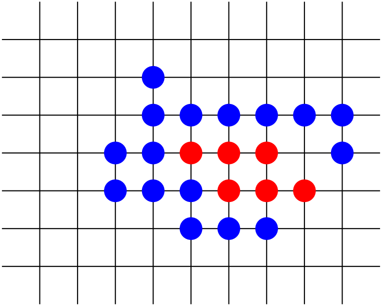
Competing first passage percolation



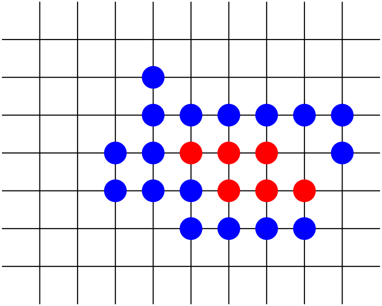
Competing first passage percolation



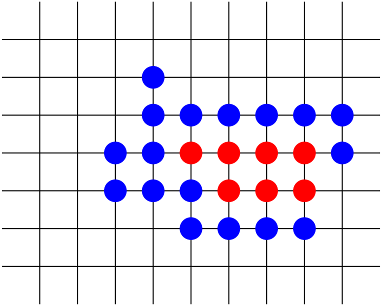
Competing first passage percolation



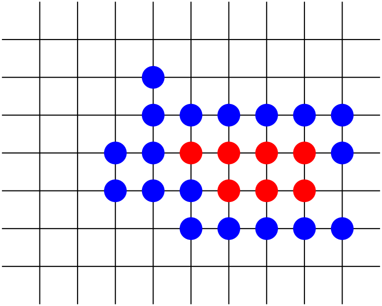
Competing first passage percolation



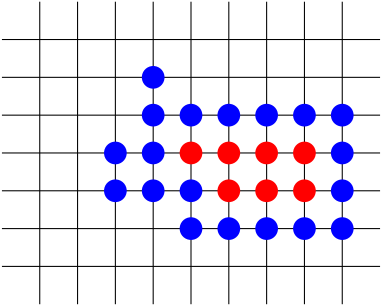
Competing first passage percolation



Competing first passage percolation



Competing first passage percolation



Theorem (Häggström, Pemantle)

On 2D lattice, for $\lambda_R = \lambda_B$

$$\mathbb{P}(\textit{both red and blue} \to \infty) > 0;$$

for at most countable set S

$$\frac{\lambda_R}{\lambda_B} \notin S \Rightarrow \mathbb{P}(\textit{both red and blue} \to \infty) = 0.$$

Random regular graphs

- Have only bounded number of short cycles.
- Neighborhoods or typical vertices are trees.
- Expander properties.
- **Configuration model**

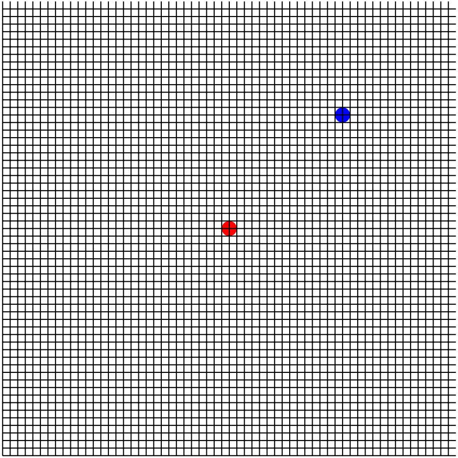
Competing process on random graphs

Let G_n be random d-regular graph on n vertices.

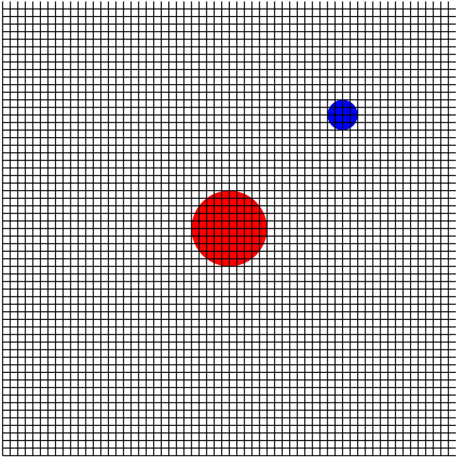
- Uniformly choose $\mathbf{r}(n)$ vertices of G_n and color it red and $\mathbf{b}(n)$ vertices and color it blue ($\mathbf{r}(n)$ and $\mathbf{b}(n)$ are given functions).
- Run the same dynamics as in the competing first passage percolation model with rates λ_R and λ_B.
- Consider the number of red and blue vertices R_n^{final} and B_n^{final} when the graphs is exhausted.

Question: Can we estimate R_n^{final} and B_n^{final}?

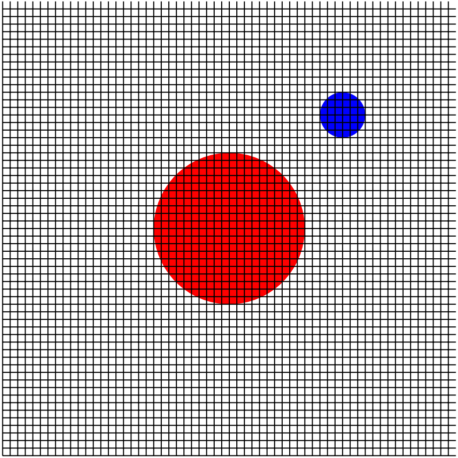
Compare with the Torus



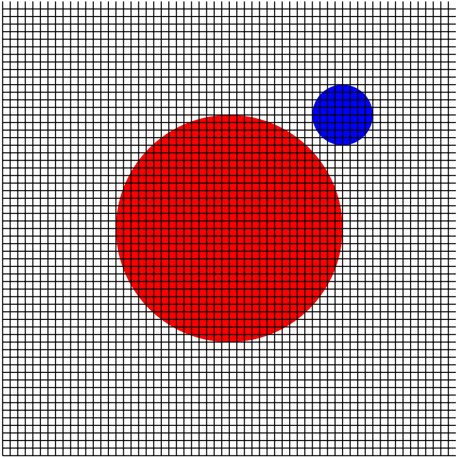
Compare with the Torus



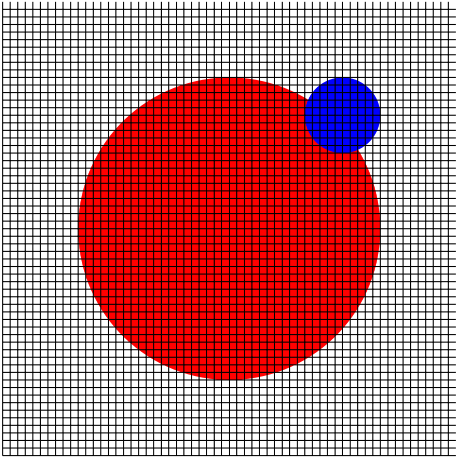
Compare with the Torus



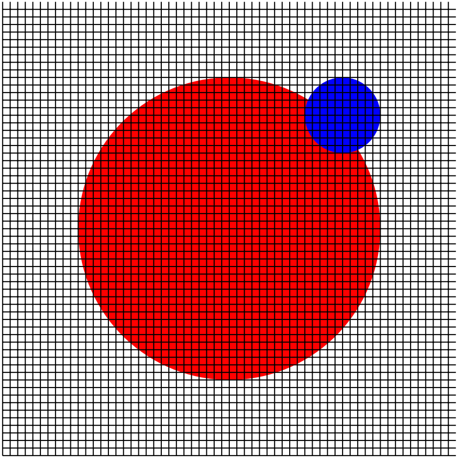
Compare with the Torus



Compare with the Torus



Compare with the Torus



Both processes occupy $\Theta(n) = \Theta(k^2)$ vertices.

Results - asymptotics

Theorem (Antunović, Dekel, M, Peres)

Up to a constant factor with high probability

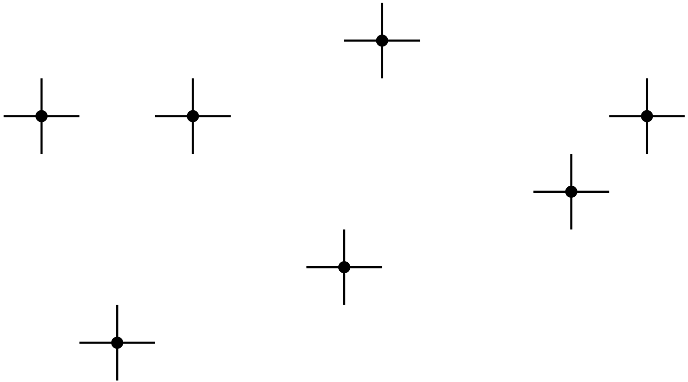
$$R_n^{total} \sim \mathbf{r}(n)\left(\frac{n}{\mathbf{b}(n)}\right)^{\lambda_R/\lambda_B} \wedge n.$$

In particular if $\mathbf{r}(n) = n^{\rho}$ *and* $\mathbf{b}(n) = n^{\beta}$ *then*

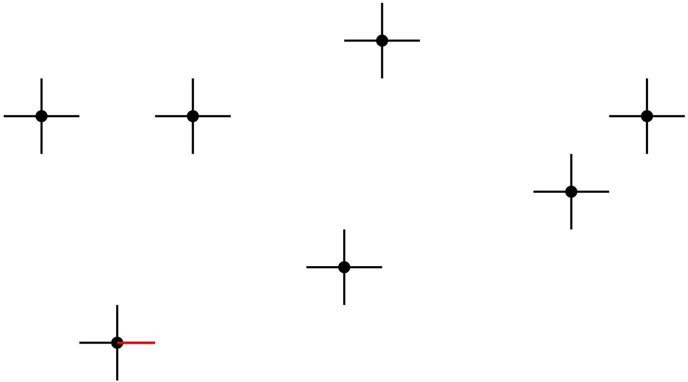
$$R_n^{total} \sim \begin{cases} n^{\rho+(1-\beta)\lambda_R/\lambda_B}, & \text{for } \rho < 1-(1-\beta)\lambda_R/\lambda_B, \\ n, & \text{for } \rho \geq 1-(1-\beta)\lambda_R/\lambda_B. \end{cases}$$

"Balance" occurs at $(1-\rho)\lambda_B = (1-\beta)\lambda_R$.

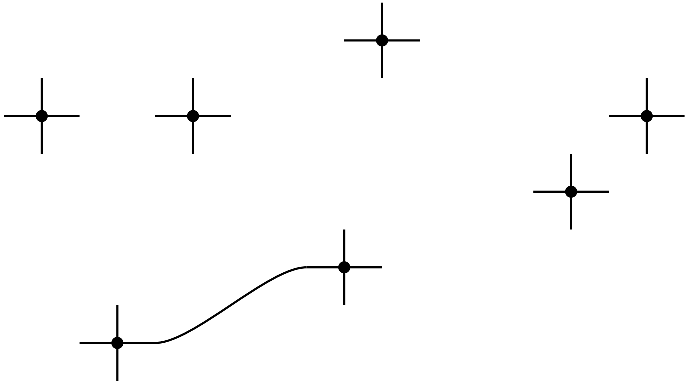
Configuration model



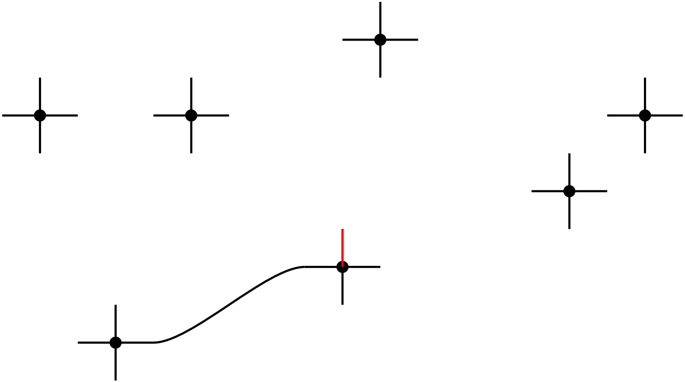
Configuration model



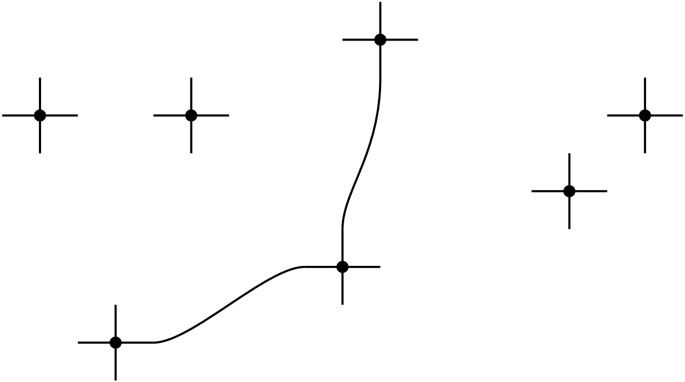
Configuration model



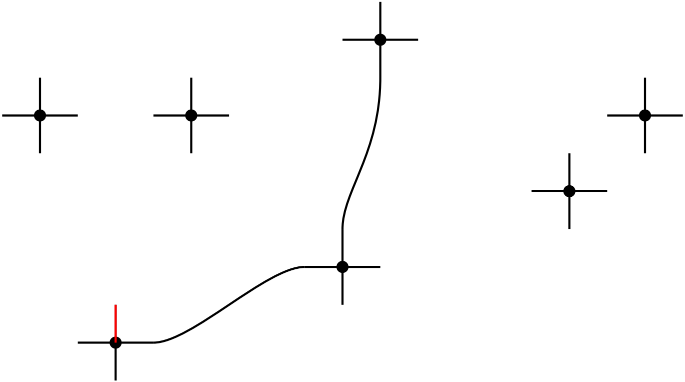
Configuration model



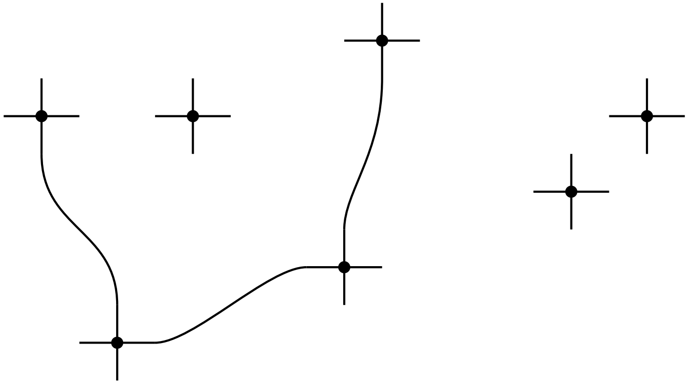
Configuration model



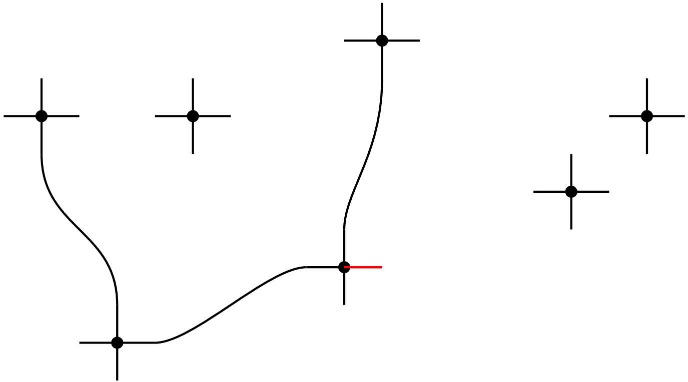
Configuration model



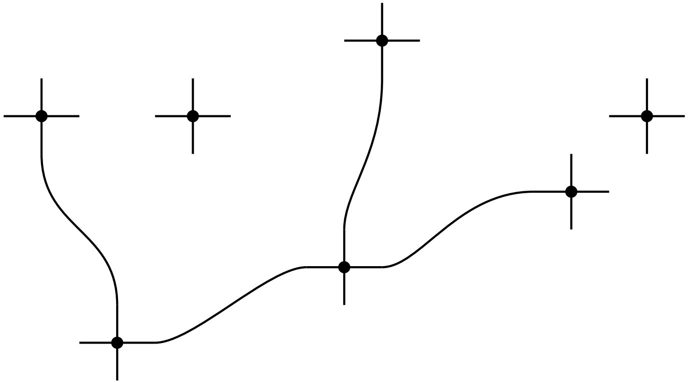
Configuration model



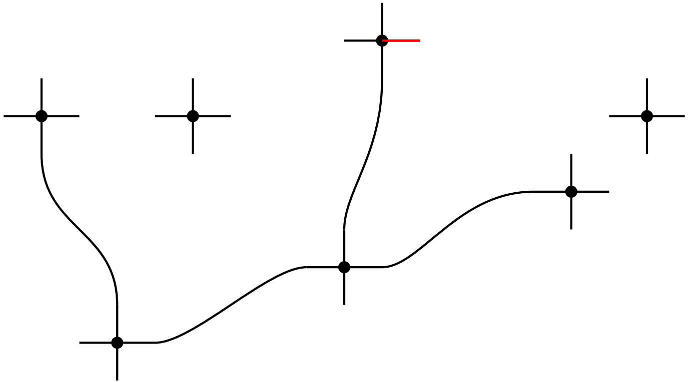
Configuration model



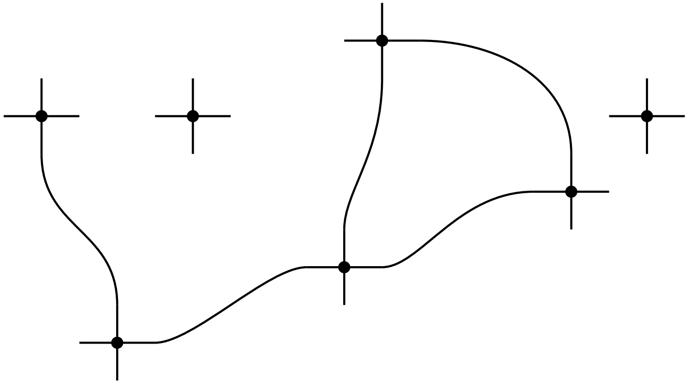
Configuration model



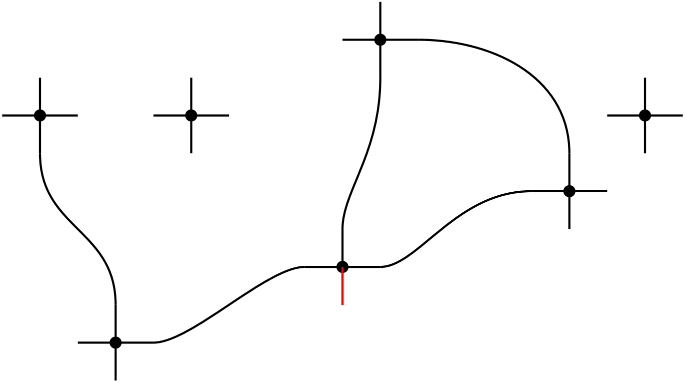
Configuration model



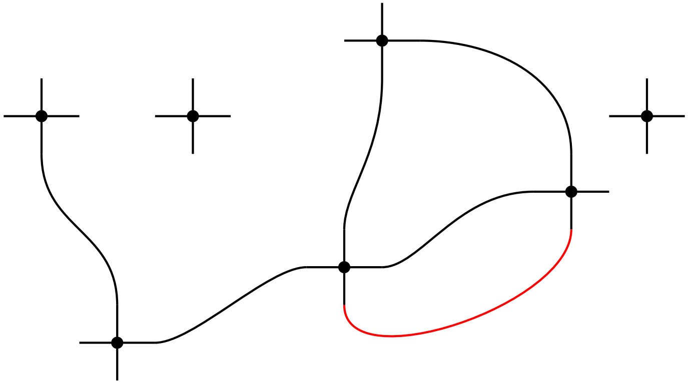
Configuration model



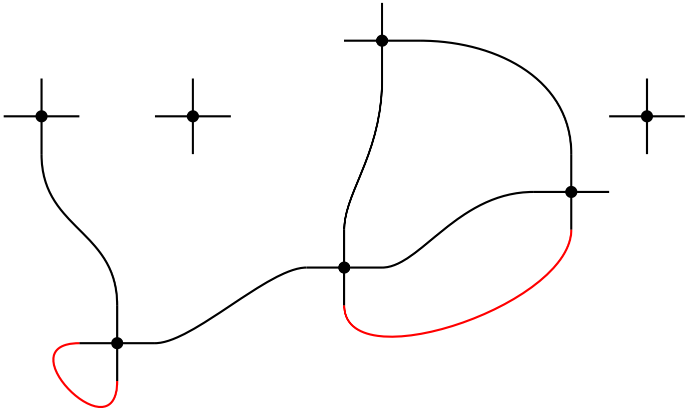
Configuration model



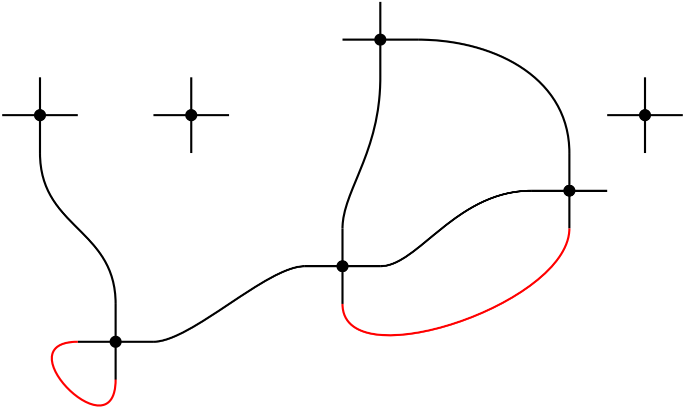
Configuration model



Configuration model

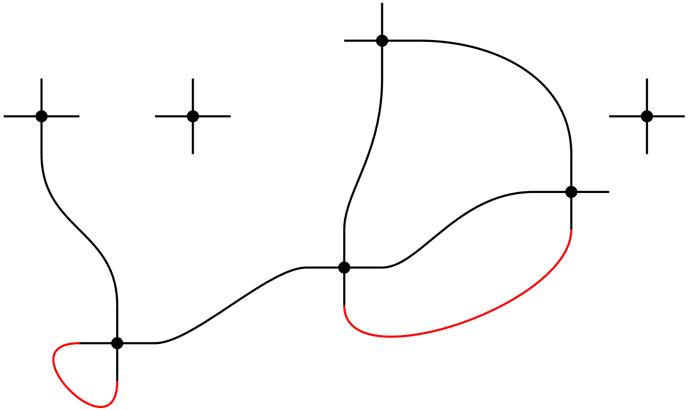


Configuration model



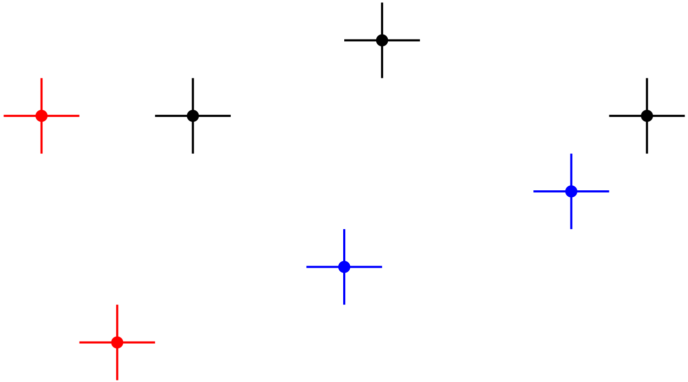
$\mathbb{P}$(Bad configuration) is bounded away from 1.

Configuration model

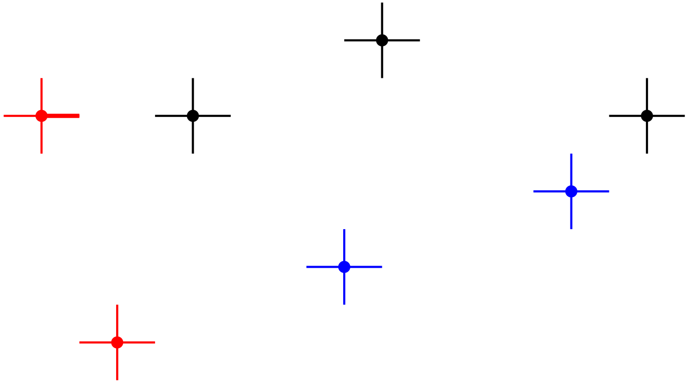


$\mathbb{P}$(Bad configuration) is bounded away from 1.
Conditioned that there are no Bad configuration the algorithm generates a random regular graph.

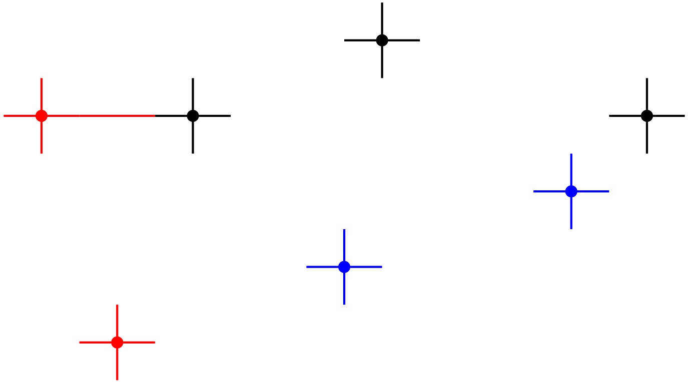
Couple CFPP and the configuration model

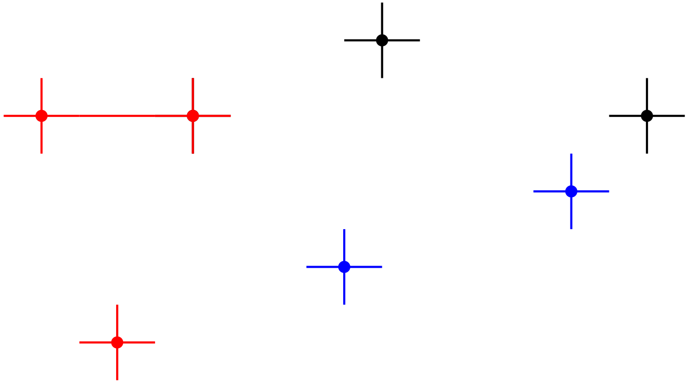


Couple CFPP and the configuration model

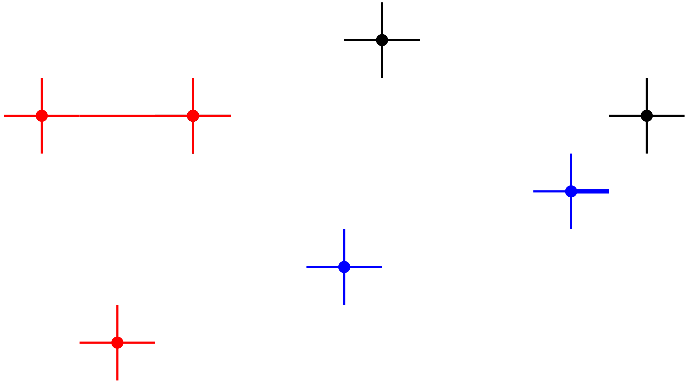


Couple CFPP and the configuration model

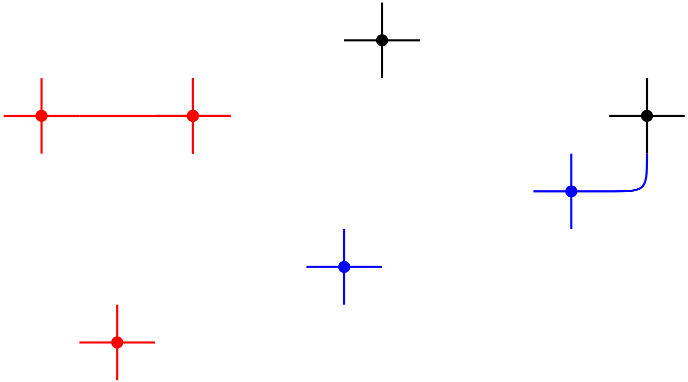




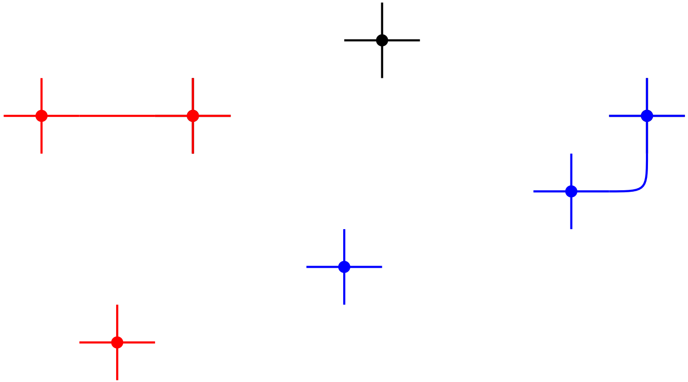
Couple CFPP and the configuration model



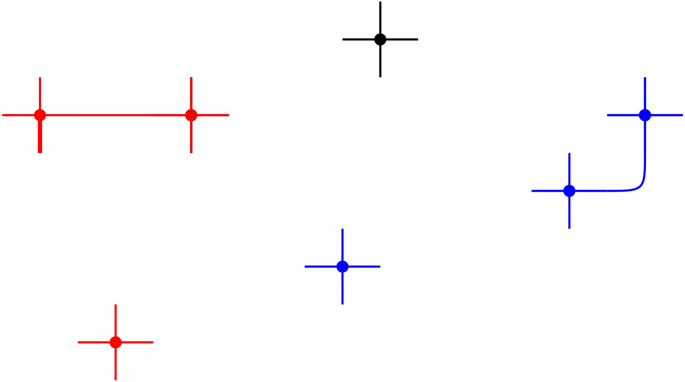
Couple CFPP and the configuration model

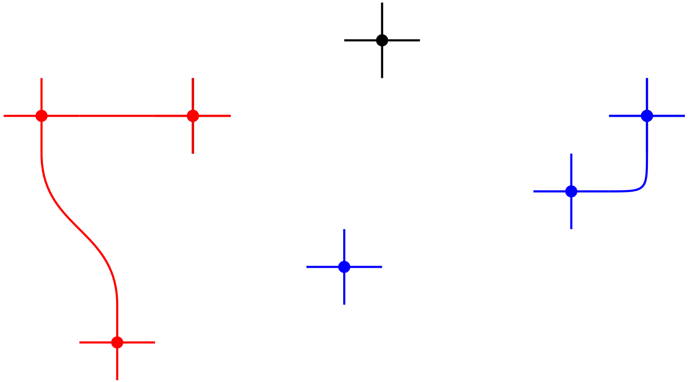


Couple CFPP and the configuration model

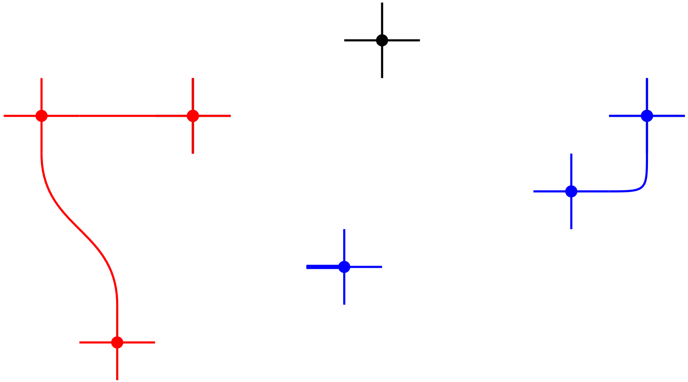


Couple CFPP and the configuration model

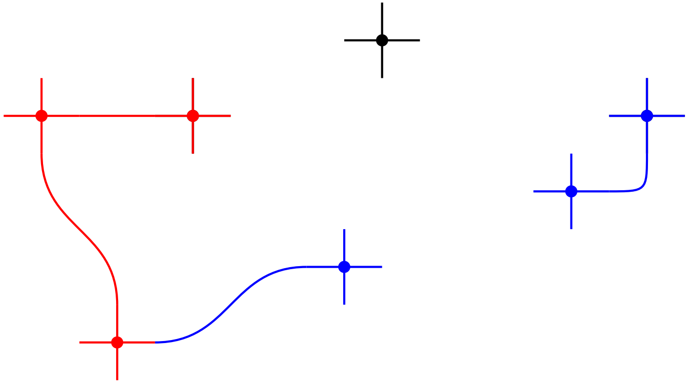




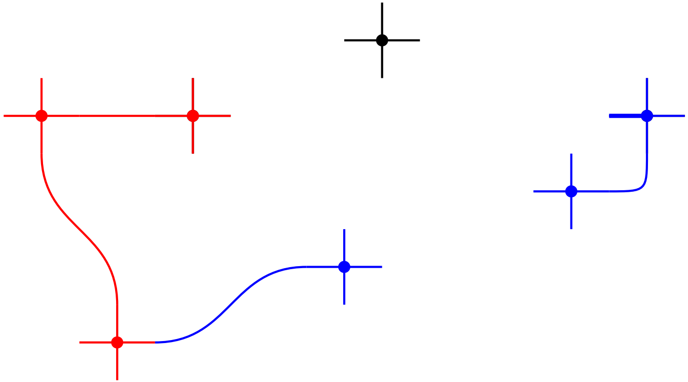
Couple CFPP and the configuration model



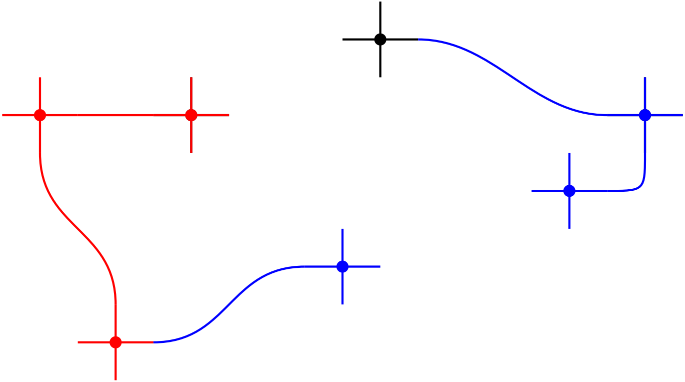
Couple CFPP and the configuration model



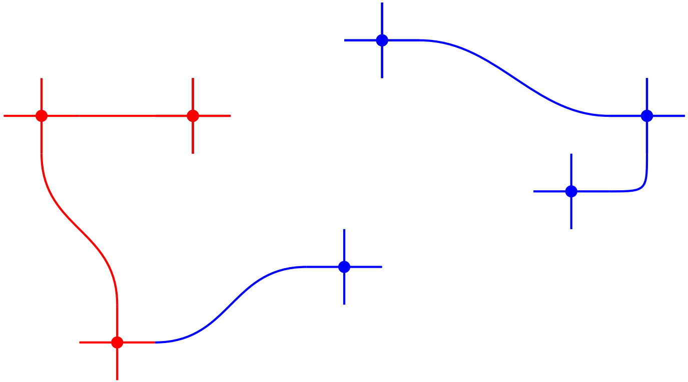
Couple CFPP and the configuration model



Couple CFPP and the configuration model



Couple CFPP and the configuration model



Couple CFPP and the configuration model

We will keep track of:
$\mathcal{R}_t$ and $\mathcal{B}_t$ ... number of red and blue "half-edges" at step t.

Couple CFPP and the configuration model

We will keep track of:
$\mathcal{R}_t$ and $\mathcal{B}_t$ ... number of red and blue "half-edges" at step t.
Transition probabilities:

$$(\mathcal{R}_{t+1},\mathcal{B}_{t+1}) = \begin{cases} (\mathcal{R}_t + d - 2, \mathcal{B}_t), & w.p.\ \frac{\lambda_R \mathcal{R}_t}{\lambda_R\mathcal{R}_t+\lambda_B\mathcal{B}_t}\frac{dn-2t-\mathcal{R}_t-\mathcal{B}_t}{dn-2t-1} \\ (\mathcal{R}_t, \mathcal{B}_t + d - 2), & w.p.\ \frac{\lambda_B \mathcal{B}_t}{\lambda_R\mathcal{R}_t+\lambda_B\mathcal{B}_t}\frac{dn-2t-\mathcal{R}_t-\mathcal{B}_t}{dn-2t-1} \\ (\mathcal{R}_t - 2, \mathcal{B}_t), & w.p.\ \frac{\lambda_R \mathcal{R}_t}{\lambda_R\mathcal{R}_t+\lambda_B\mathcal{B}_t}\frac{\mathcal{R}_t-1}{dn-2t-1} \\ (\mathcal{R}_t, \mathcal{B}_t - 2), & w.p.\ \frac{\lambda_B \mathcal{B}_t}{\lambda_R\mathcal{R}_t+\lambda_B\mathcal{B}_t}\frac{\mathcal{B}_t-1}{dn-2t-1} \\ (\mathcal{R}_t - 1, \mathcal{B}_t - 1), & w.p.\ \frac{(\lambda_R+\lambda_B)\mathcal{B}_t\mathcal{R}_t}{(\lambda_R\mathcal{R}_t+\lambda_B\mathcal{B}_t)(dn-2t-1)} \end{cases}$$

Couple CFPP and the configuration model

To control $\mathcal{R}_t$ and $\mathcal{B}_t$ we use martingale techniques to control

$$X_t = \mathcal{R}_t + \mathcal{B}_t$$

and

$$Y_t = \frac{\mathcal{R}_t}{\mathcal{B}_t^{\lambda_R/\lambda_B}}$$

Couple CFPP and the configuration model

To control $\mathcal{R}_t$ and $\mathcal{B}_t$ we use martingale techniques to control

$$X_t = \mathcal{R}_t + \mathcal{B}_t$$

and

$$Y_t = \frac{\mathcal{R}_t}{\mathcal{B}_t^{\lambda_R/\lambda_B}}$$

X_t ... number of active half-edges in the configuration model

Couple CFPP and the configuration model

To control $\mathcal{R}_t$ and $\mathcal{B}_t$ we use martingale techniques to control

$$X_t = \mathcal{R}_t + \mathcal{B}_t$$

and

$$Y_t = \frac{\mathcal{R}_t}{\mathcal{B}_t^{\lambda_R/\lambda_B}}$$

X_t ... number of active half-edges in the configuration model
Y_t ... is the continuous time martingale when both processes evolve without any interactions (and self-interactions)

Couple CFPP and the configuration model

$$X_t = \mathcal{R}_t + \mathcal{B}_t$$

Process $(X_t, dn - 2t - X_t)$ evolves as an urn model

$$\begin{pmatrix} -2 & 0 \\ d-2 & -d \end{pmatrix}$$

Couple CFPP and the configuration model

$$X_t = \mathcal{R}_t + \mathcal{B}_t$$

Process $(X_t, dn - 2t - X_t)$ evolves as an urn model

$$\begin{pmatrix} -2 & 0 \\ d-2 & -d \end{pmatrix}$$

As $n \to \infty$

$$X_t = (1 \pm o(1)) \left(dn - 2t - (dn - X_0) \left(1 - \frac{2t}{dn} \right)^{d/2} \right),$$

for all $0 \leq t < dn/2$.

Couple CFPP and the configuration model

$$X_t = \mathcal{R}_t + \mathcal{B}_t$$

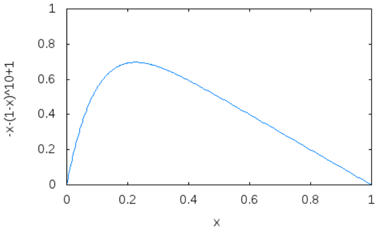
Process $(X_t, dn - 2t - X_t)$ evolves as an urn model

$$\begin{pmatrix} -2 & 0 \\ d-2 & -d \end{pmatrix}$$

As $n \to \infty$

$$X_t = (1 \pm o(1)) \left(dn - 2t - (dn - X_0)\left(1 - \frac{2t}{dn}\right)^{d/2} \right),$$

for all $0 \leq t < dn/2$.



$$Y_t = \frac{\mathcal{R}_t}{\mathcal{B}_t^{\lambda_R/\lambda_B}}$$

As long as X_t is large

$$Y_t = (1 \pm o(1)) Y_0 \left(1 - \frac{2t}{dn}\right)^{\frac{\lambda_B - \lambda_R}{2\lambda_B}}.$$

Couple CFPP and the configuration model

$$Y_t = \frac{\mathcal{R}_t}{\mathcal{B}_t^{\lambda_R/\lambda_B}}$$

As long as X_t is large

$$Y_t = (1 \pm o(1)) Y_0 \Big(1 - \frac{2t}{dn} \Big)^{\frac{\lambda_B - \lambda_R}{2\lambda_B}}.$$

Observe the "extra advantage" of the faster process.

THANKS!!!

Thank you!

- Speakers: …
- Organizers:
- Laurent, Lionel, Tasia
- Heather Peterson
- You!