Analysis of Large Random Graphs

Emilie Coupechoux - Marc Lelarge

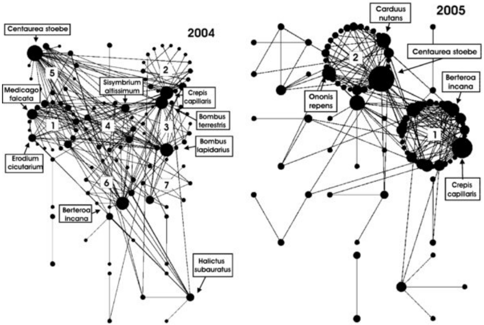
INRIA-ENS

5 December 2012

Outline

Outline

Complex networks



Plants and bees : co-occurrence network

Complex networks

Examples :

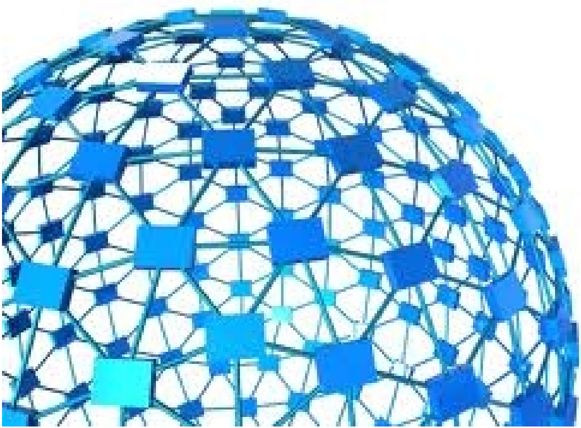
- Population *(individuals)*
- Internet *(autonomous systems)*
- World-Wide Web *(sites)*
- Collaborations *(individuals)*
- ...

Model : finite graph $G = (V, E)$

- V = set of vertices / *nodes*
- E = set of edges / *connections between nodes*

Large networks

Empirical data = *local* observations $\Rightarrow$ statistics on the network

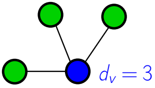


Large networks

Empirical data = *local* observations ⇒ statistics on the network

Example :

Degree = number of acquantainces for each node



- Observation : degree of each node
- Computation : for all $k \geq 0$, probability p_k that a node (chosen uniformly at random) has degree k :

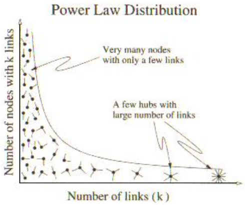
$$p_k = \frac{\text{nb of nodes with degree } k}{\text{total nb of nodes}}$$

- Deduction : (empirical) distribution of degrees $\boldsymbol{p} = (p_k)_{k \geq 0}$

Definition 1 : *Scale-free* networks

Power law degree distribution :
there exists $\tau > 0$ such that, for all $k \geq 0$,

$$p_k = \frac{\text{nb of nodes with degree } k}{\text{total nb of nodes}} \quad \propto \quad k^{-\tau}$$



(small number of nodes having a large number of edges)

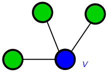
Definition 2 : *Clustering* coefficient of a graph

("The friends of my friends are my friends", Newman, '03)

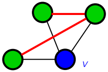
$$C := \frac{3 \times \text{nb of triangles}}{\text{nb of connected triples}} = \frac{\sum_v P_v}{\sum_v N_v} > 0$$

$P_v :=$ nb of pairs of neighbors of v sharing an edge together,
$N_v :=$ nb of pairs of neighbors of v : $N_v = d_v(d_v - 1)/2$.

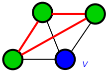
Example :
$N_v = 3$



$P_v = 0$



$P_v = 2$



$P_v = 3$

Define a model of *random* graphs

- having (asymptotically) the observed properties :
  - *scale-free* networks
  - networks with *clustering*
- tractable

Epidemic models

DIFFUSION
(Bond percolation)

CONTAGION
(Watts threshold model)

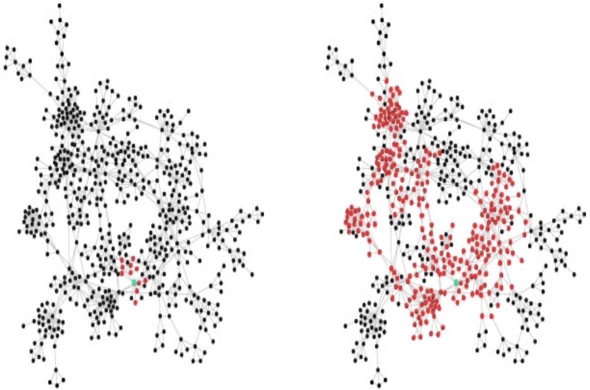
Examples :

- Spread of a disease
- Spread of e-mail viruses
- Diffusion of innovations
- Adoption of a new technology

Influence of the neighbors :

- Independence
- Probabilistic
- Joint influence
- Deterministic

Cascade phenomenon



What we are interested in...

When the epidemic starts from only one individual

- Phase transition
- Cascade size
- Effect of clustering on both phase transition and cascade size

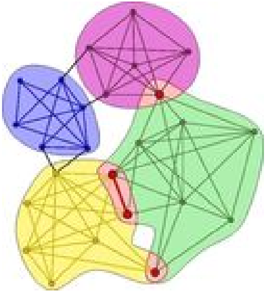
When the epidemic starts from a positive proportion of the population

- Epidemic size

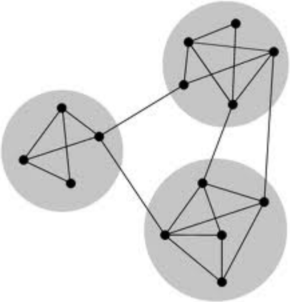
For which kind of random graph models ?

Organization in communities

Overlapping communities



Separate communities



Random graphs with separate communities

DIFFUSION	CONTAGION

Phase transition and cascade size

DIFFUSION	CONTAGION
Ball *et al.* (same community sizes + heuristics) Trapman, Gleeson *et al.* (heuristics)	Lelarge (without clustering)

Effect of clustering

Ball *et al.* (Poisson distribution)
Gleeson *et al.*

Positive proportion of infected individuals at the start

Random graphs with separate communities

DIFFUSION	CONTAGION
Phase transition and cascade size	
Ball *et al.* (same community sizes + heuristics) Trapman, Gleeson *et al.* (heuristics) $\surd$	Lelarge (without clustering) $\surd$
Effect of clustering	
Ball *et al.* (Poisson distribution) Gleeson *et al.* $\surd$	$\surd$
Positive proportion of infected individuals at the start	
$\surd$	$\surd$

Random graphs with overlapping communities

DIFFUSION

CONTAGION

Phase transition and cascade size

Newman (BP)

Britton *et al.* (Poisson distribution)

Effect of clustering

Britton *et al.* (Poisson distribution)

Newman (effect on cascade size)

Positive proportion of infected individuals at the start

Random graphs with overlapping communities

DIFFUSION	CONTAGION
Phase transition and cascade size	
Newman (BP) Britton *et al.* (Poisson distribution)	$\surd$ (BP)
Effect of clustering	
Britton *et al.* (Poisson distribution) Newman (effect on cascade size) $\surd$ (Poisson distribution)	
Positive proportion of infected individuals at the start	
$\surd$	$\surd$ (Variant of the model)

Outline

Random graph model with separate communities

- Configuration model $G(n, \boldsymbol{d})$:

Degree sequence $\boldsymbol{d} = (d_i)_{1 \leq i \leq n}$

$$\frac{\sharp\{i : d_i = r\}}{n} \underset{n \to \infty}{\longrightarrow} p_r = \text{probability that a vertex has degree } r$$

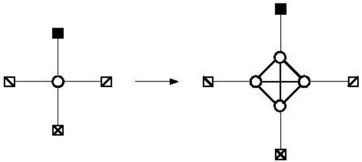
But : converges locally to a tree

Random graph model with separate communities

- Configuration model $G(n, \boldsymbol{d})$:

$$\boldsymbol{d} = (d_i)_{1 \leq i \leq n}$$

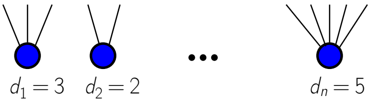
- Idea : Replace a vertex of degree r in $G(n, \boldsymbol{d})$ by a clique of size r :



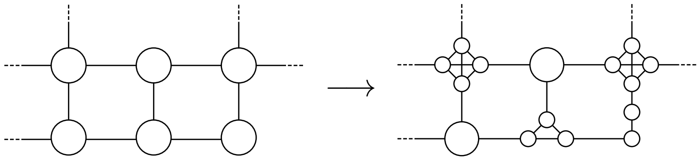
Random graph model with separate communities

- Configuration model $G(n, \boldsymbol{d})$:

$$\boldsymbol{d} = (d_i)_{1 \leq i \leq n}$$



- Idea : Replace a vertex of degree r in $G(n, \boldsymbol{d})$ by a clique of size r.
- Adding cliques *randomly* : Let $\gamma \in [0, 1]$.
  Each vertex is replaced by a clique with probability γ (independently for all vertices).

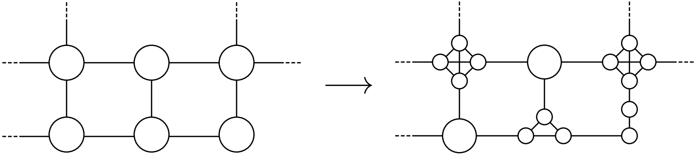


Random graph model with separate communities

- Configuration model $G(n, \boldsymbol{d})$:

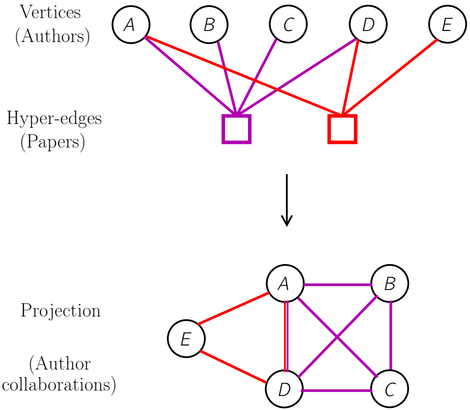
$$\boldsymbol{d} = (d_i)_{1 \le i \le n}$$

- Idea : Replace a vertex of degree r in $G(n, \boldsymbol{d})$ by a clique of size r.
- Adding cliques *randomly* : Let $\gamma \in [0, 1]$.
  Each vertex is replaced by a clique with probability γ (independently for all vertices).



- Advantage : clustering coefficient can be easily tuned, while keeping the *whole* degree distribution fixed

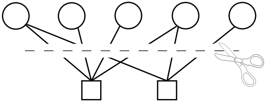
Community structure of real-world networks



Random graph model with overlapping communities

Given degrees for both top and bottom vertices

Uniform matching of the half-edges



Projection of this random hypergraph

Advantage : any distribution for

- the community sizes
- the number of communities an individual belongs to

Drawbacks :

- Strong dependence between the edges
- Clustering coefficient and degree distributions cannot be tuned independently

$\Longrightarrow$ Equivalent branching process for the random hypergraph

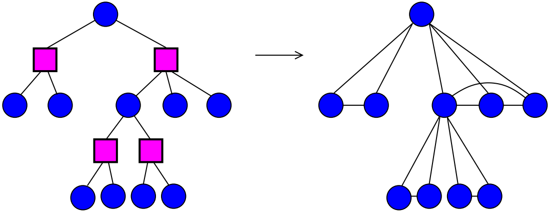
Random graph model with overlapping communities

In the hypergraph :

- p_d = probability that a vertex has degree d
- q_w = probability that a hyper-edge has weight w

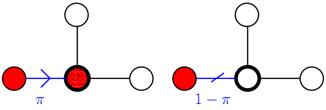
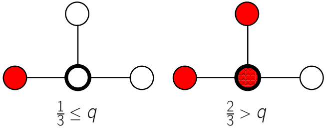
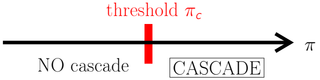
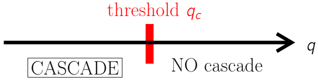
Alternating branching process :

- dp_d/λ = probability that a vertex has $d-1$ children
- wq_w/μ = probability that a hyper-edge has $w-1$ children



Epidemic models on these random graphs

Final nb of infected nodes negligeable or not / population size ?

	DIFFUSION MODEL	CONTAGION MODEL
Ref.	Bond percolation	Morris, Watts
Para- -meter	π = probability that an edge transmits the epidemic 	A vertex is infected $\Leftrightarrow$ fraction of infected neighbors $> q$ 
Thm		

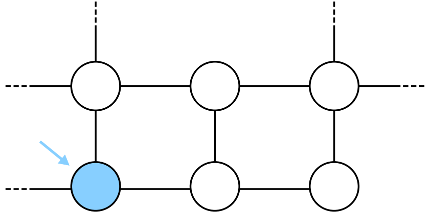
Effect of clustering on these thresholds and on the cascade size

Outline

Random graph with separate communities

Diffusion model with a given probability π of transmission

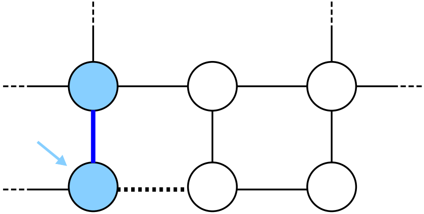
- At the beginning, activate a given vertex (= the seed of the epidemic)
- Transmit the epidemic through any edge with probability π

INITIAL GRAPH	GRAPH WITH CLIQUES
	

Random graph with separate communities

Diffusion model with a given probability π of transmission

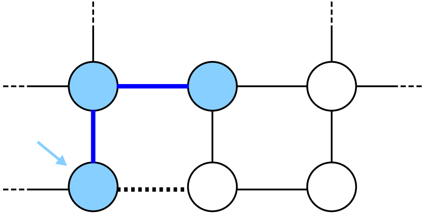
- At the beginning, activate a given vertex (= the seed of the epidemic)
- Transmit the epidemic through any edge with probability π

INITIAL GRAPH	GRAPH WITH CLIQUES
	

Random graph with separate communities

Diffusion model with a given probability π of transmission

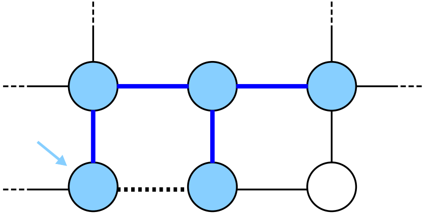
- At the beginning, activate a given vertex (= the seed of the epidemic)
- Transmit the epidemic through any edge with probability π

INITIAL GRAPH	GRAPH WITH CLIQUES
	

Random graph with separate communities

Diffusion model with a given probability π of transmission

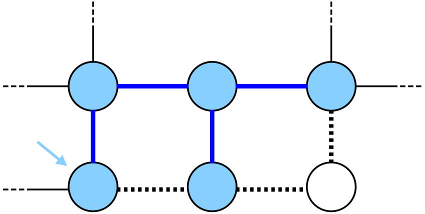
- At the beginning, activate a given vertex (= the seed of the epidemic)
- Transmit the epidemic through any edge with probability π

INITIAL GRAPH	GRAPH WITH CLIQUES
	

Random graph with separate communities

Diffusion model with a given probability π of transmission

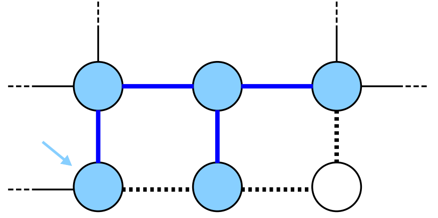
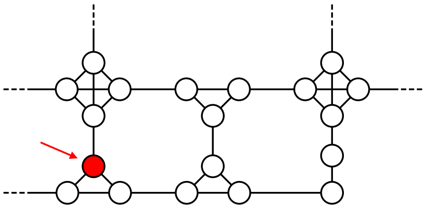
- At the beginning, activate a given vertex (= the seed of the epidemic)
- Transmit the epidemic through any edge with probability π

INITIAL GRAPH	GRAPH WITH CLIQUES
	

Random graph with separate communities

Diffusion model with a given probability π of transmission

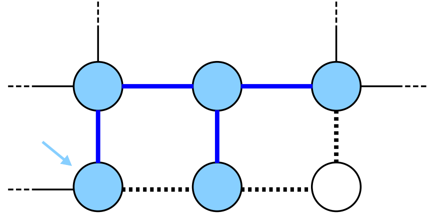
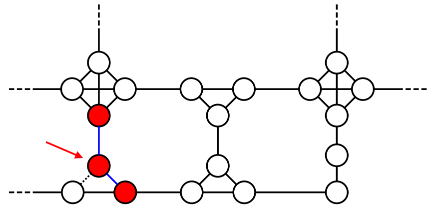
- At the beginning, activate a given vertex (= the seed of the epidemic)
- Transmit the epidemic through any edge with probability π

INITIAL GRAPH	GRAPH WITH CLIQUES
	

Random graph with separate communities

Diffusion model with a given probability π of transmission

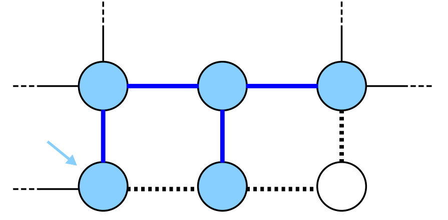
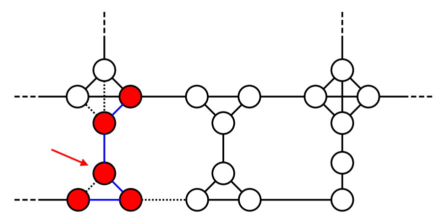
- At the beginning, activate a given vertex (= the seed of the epidemic)
- Transmit the epidemic through any edge with probability π

INITIAL GRAPH	GRAPH WITH CLIQUES
	

Random graph with separate communities

Diffusion model with a given probability π of transmission

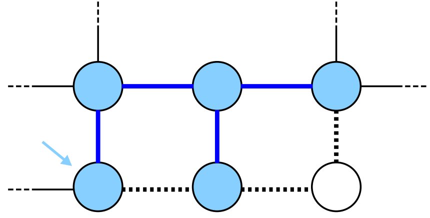
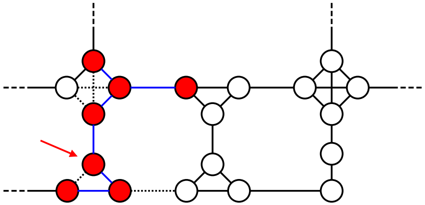
- At the beginning, activate a given vertex (= the seed of the epidemic)
- Transmit the epidemic through any edge with probability π

INITIAL GRAPH	GRAPH WITH CLIQUES
	

Random graph with separate communities

Diffusion model with a given probability π of transmission

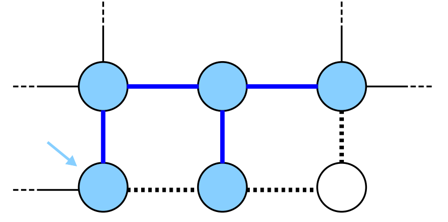
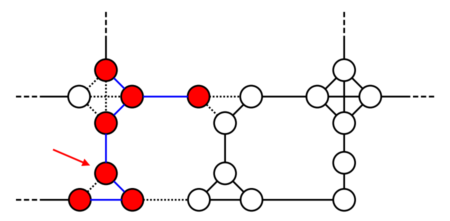
- At the beginning, activate a given vertex (= the seed of the epidemic)
- Transmit the epidemic through any edge with probability π

INITIAL GRAPH	GRAPH WITH CLIQUES
	

Random graph with separate communities

Diffusion model with a given probability π of transmission

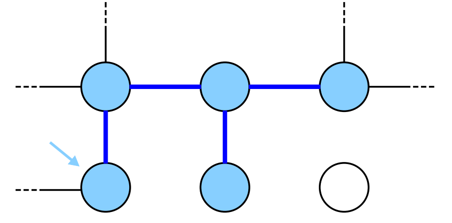
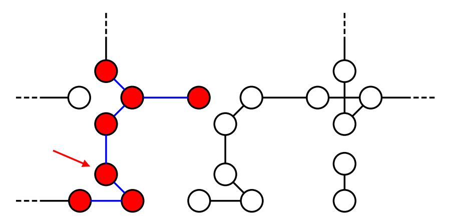
- At the beginning, activate a given vertex (= the seed of the epidemic)
- Transmit the epidemic through any edge with probability π

INITIAL GRAPH	GRAPH WITH CLIQUES
	

Random graph with separate communities

Diffusion model with a given probability π of transmission

- At the beginning, activate a given vertex (= the seed of the epidemic)
- Transmit the epidemic through any edge with probability π

INITIAL GRAPH	GRAPH WITH CLIQUES
	

Connected component of the seed
in the bond percolated graph

Random graph with separate communities

Theorem (DIFFUSION THRESHOLD)

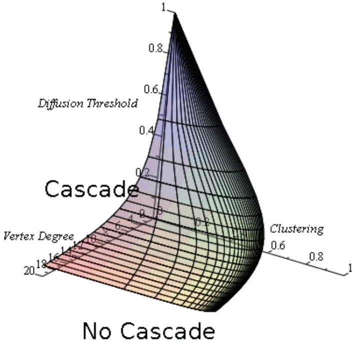
Let π_c be the solution of the equation : $\boxed{\pi' = \frac{\mathbb{E}[D_{\pi'}]}{\mathbb{E}[D_{\pi'}(D_{\pi'}-1)]}}$,
where $D_{\pi'}$ is a random variable with a given distribution that depends on $\boldsymbol{p}$, γ and π'.

- $\pi > \pi_c$: There exists *whp* a giant component in the percolated graph, *i.e.* a single node can trigger a global cascade.
- $\pi < \pi_c$: The size of the epidemic generated by a vertex u (chosen uniformly at random) is negligeable : $o_p(n)$.

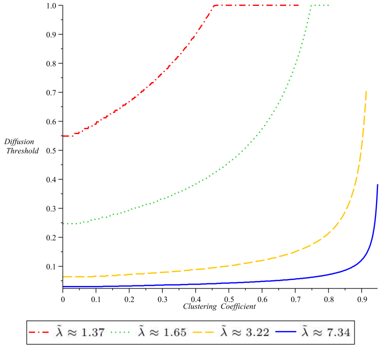
Random graph with separate communities

Diffusion Threshold π_c vs Clustering

Random d-regular graphs

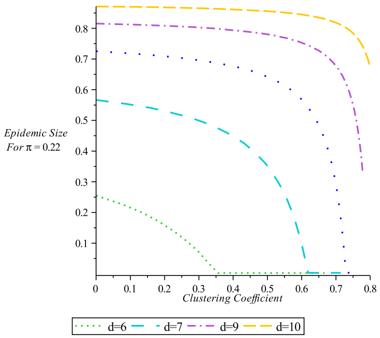


Power law graphs



Random graph with separate communities

Cascade size vs Clustering
(Infection probability : $\pi = 0.22$)



Comparison with the diffusion on the random graph model with overlapping communities

Effect of clustering on the diffusion threshold :

- In graphs with separate communities, clustering 'inhibits' the diffusion process (cf. also Ball *et al.*) ;
- In graphs with overlapping communities, clustering 'helps' the diffusion to spread (cf. also Britton *et al.*).

Effect of clustering on the cascade size :

- In both cases, clustering reduces the cascade size (cf. Newman for graphs with overlapping communities).

Outline

Game-theoretic contagion model on a given graph $G = (V, E)$, with parameter $q \in (0, 1/2)$:

Two possible choices : ($\leftrightarrow$ susceptible) or ($\leftrightarrow$ infected)

Initially : all use , except one who uses

Possible switch $\rightarrow$, but no switch $\not\rightarrow$

Situation	Payoff (for both users)
	q
	$1 - q > q$
	0

Total payoff
= sum of payoffs from all your neighbors

Switch from to $\Leftrightarrow \frac{|\text{Neighbors using Skype}|}{|\text{Neighbors}|} > q.$

Infinite deterministic graph $G = (V, E)$

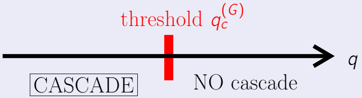
Parameter q varies :

$$q \text{ small} \Rightarrow \boxed{\text{CASCADE}}$$
$$q \text{ higher} \Rightarrow \text{NO cascade}$$

More precisely :

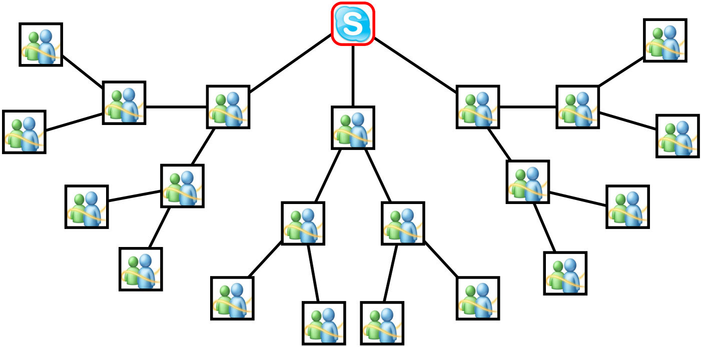
$q_1 \geq q_2$, cascade for $q_1 \Rightarrow$ cascade for q_2

Contagion threshold $q_c^{(G)} := \sup \left\{ q \,\middle|\, \text{CASCADE in } G \text{ for parameter } q \right\}$



Switch from  to  $\Leftrightarrow \frac{|\text{Neighbors using Skype}|}{|\text{Neighbors}|} > q$

Example : $G = d$-regular tree



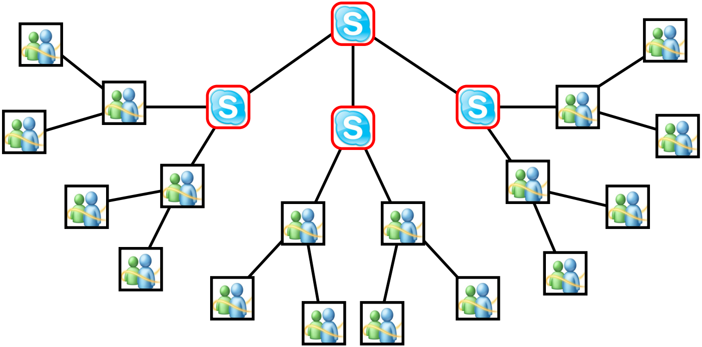
$q \geq 1/d \quad \Rightarrow \quad$ NO cascade

$q < 1/d \quad \Rightarrow \quad$ CASCADE

$\Rightarrow \quad q_c^{(G)} = 1/d$

Switch from to $\Leftrightarrow \dfrac{|\text{Neighbors using Skype}|}{|\text{Neighbors}|} > q$

Example : $G = d$-regular tree



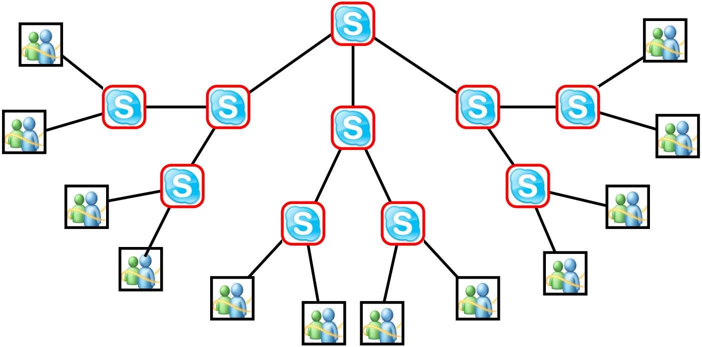
$q \geq 1/d \quad \Rightarrow \quad$ NO cascade

$q < 1/d \quad \Rightarrow \quad$ CASCADE

$\Rightarrow \quad q_c^{(G)} = 1/d$

Switch from [MSN] to [Skype] $\Leftrightarrow \frac{|\text{Neighbors using Skype}|}{|\text{Neighbors}|} > q$

Example : G = d-regular tree



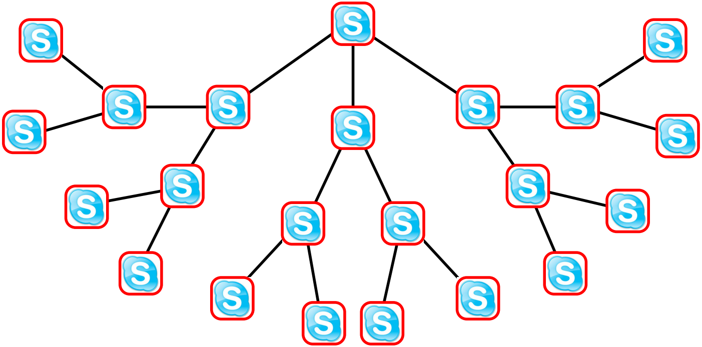
$q \geq 1/d \quad \Rightarrow \quad$ NO cascade

$q < 1/d \quad \Rightarrow \quad$ CASCADE

$\Rightarrow \quad q_c^{(G)} = 1/d$

Switch from to $\Leftrightarrow \frac{|\text{Neighbors using Skype}|}{|\text{Neighbors}|} > q$

Example : $G = d$-regular tree



$q \geq 1/d \Rightarrow$ NO cascade

$q < 1/d \Rightarrow$ CASCADE

$$\Rightarrow q_c^{(G)} = 1/d$$

Random graph with separate communities

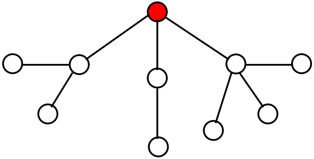
- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

proportion of its infected neighbors $> q$

Heuristically...

The random graph $G(n, \boldsymbol{d})$ converges locally to a random tree such that :

$$\mathbb{P}\left(r-1 \text{ children}\right) = rp_r/\lambda$$



$q = \frac{1}{4}$

Infected nodes = those with degree $< 1/q$

Infinite tree (of infected nodes)

$$\iff \sum_{r<1/q} (r-1)\frac{rp_r}{\lambda} > 1$$

Random graph with separate communities

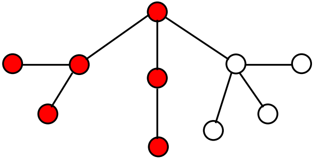
- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

proportion of its infected neighbors $> q$

Heuristically...

The random graph $G(n, \boldsymbol{d})$ converges locally to a random tree such that :

$$\mathbb{P}(r-1 \text{ children}) = rp_r/\lambda$$



$q = \frac{1}{4}$

Infected nodes = those with degree $< 1/q$

Infinite tree (of infected nodes)

$$\iff \sum_{r<1/q} (r-1)\frac{rp_r}{\lambda} > 1$$

Random graph with separate communities

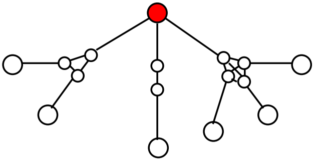
- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

proportion of its infected neighbors $> q$

Heuristically...

The random graph $G(n, \boldsymbol{d})$ converges locally to a random tree such that :

$$\mathbb{P}(r-1 \text{ children}) = rp_r/\lambda$$



$q = \frac{1}{4}$

Infected nodes = those with degree $< 1/q$

Infinite tree (of infected nodes)

$$\iff \sum_{r<1/q} (r-1)\frac{rp_r}{\lambda} > 1$$

Random graph with separate communities

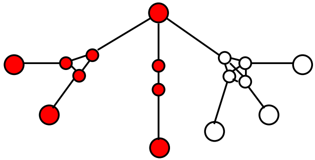
- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

proportion of its infected neighbors $> q$

Heuristically...

The random graph $G(n, \boldsymbol{d})$ converges locally to a random tree such that :

$$\mathbb{P}(r-1 \text{ children}) = rp_r/\lambda$$



$q = \frac{1}{4}$

Infected nodes = those with degree $< 1/q$

Infinite tree (of infected nodes)

$$\Longleftrightarrow \sum_{r<1/q} (r-1)\frac{rp_r}{\lambda} > 1$$

Random graph with separate communities

$$q_c := q_c(\boldsymbol{p}) = \sup\left\{q' : \sum_{r<1/q'}(r-1)\frac{rp_r}{\lambda} > 1\right\}$$

Fixed q, $\mathcal{P}^{(n)}$ = set of pivotal players in $\tilde{G}(n, \boldsymbol{d}, \gamma)$:

- G_0 = induced subgraph with vertices of degree $< 1/q$
- Pivotal players = vertices in the largest connected component of G_0

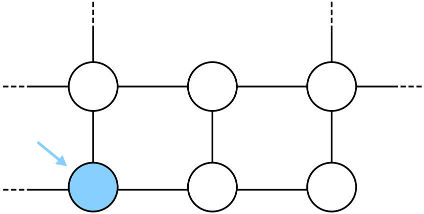
Theorem (CONTAGION THRESHOLD)

- $q < q_c$: $|\mathcal{P}^{(n)}| = \Theta_p(n)$
  Each pivotal player can trigger a global cascade.
- $q > q_c$: the size of the epidemic generated by a vertex u (chosen uniformly at random) is negligeable : $o_p(n)$.

Random graph with separate communities

- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

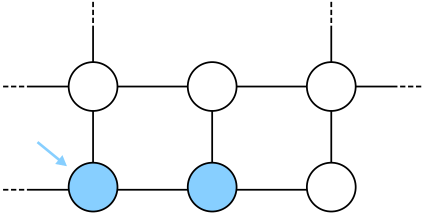
proportion of its infected neighbors $> q = \frac{1}{4}$

INITIAL GRAPH	GRAPH WITH CLIQUES
	

Random graph with separate communities

- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

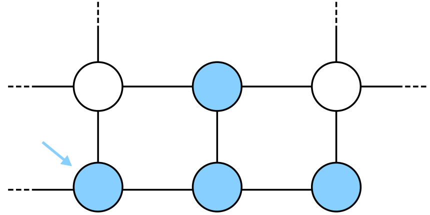
proportion of its infected neighbors $> q = \frac{1}{4}$

INITIAL GRAPH	GRAPH WITH CLIQUES
	

Random graph with separate communities

- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

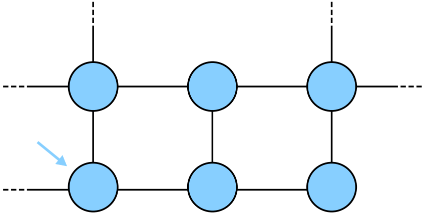
proportion of its infected neighbors $> q = \frac{1}{4}$

INITIAL GRAPH	GRAPH WITH CLIQUES
	

Random graph with separate communities

- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

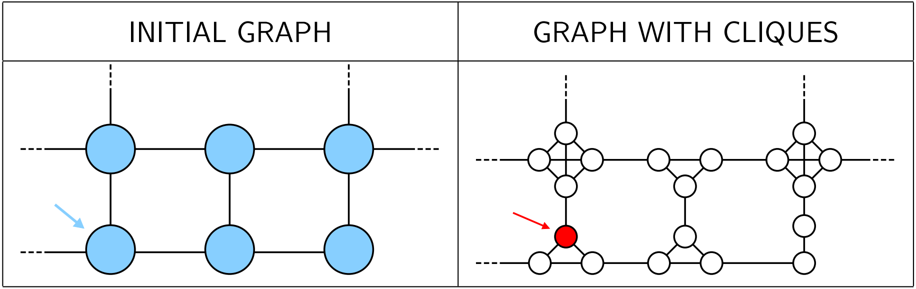
proportion of its infected neighbors $> q = \frac{1}{4}$

INITIAL GRAPH	GRAPH WITH CLIQUES
	

Random graph with separate communities

- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

proportion of its infected neighbors $> q = \frac{1}{4}$

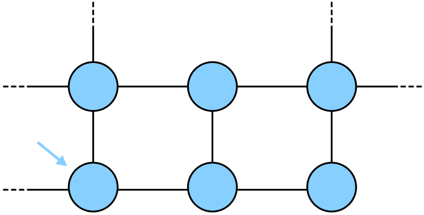


Random graph with separate communities

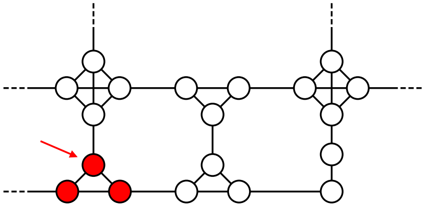
- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

proportion of its infected neighbors $> q = \frac{1}{4}$

INITIAL GRAPH



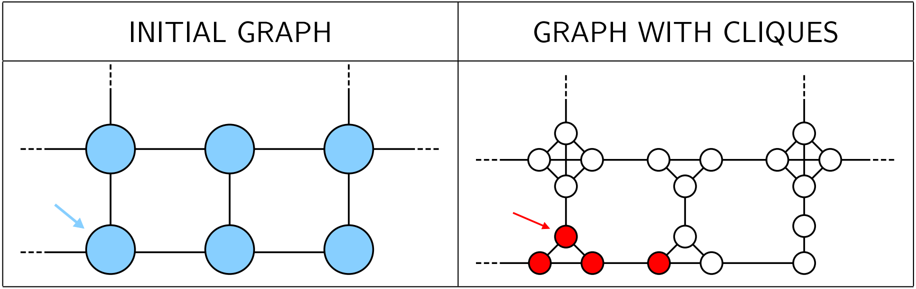
GRAPH WITH CLIQUES



Random graph with separate communities

- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

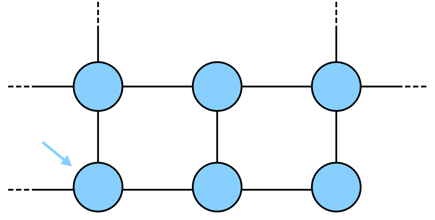
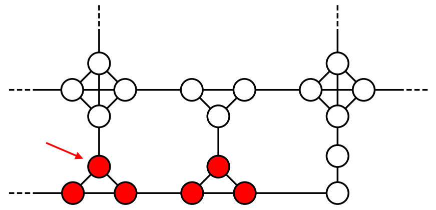
proportion of its infected neighbors $> q = \frac{1}{4}$



Random graph with separate communities

- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

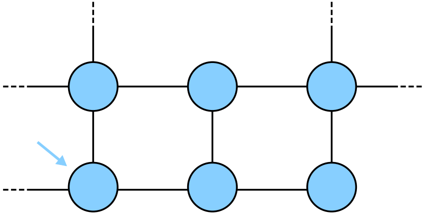
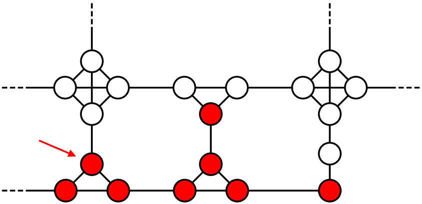
proportion of its infected neighbors $> q = \frac{1}{4}$

INITIAL GRAPH	GRAPH WITH CLIQUES
	

Random graph with separate communities

- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

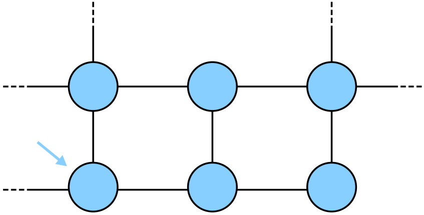
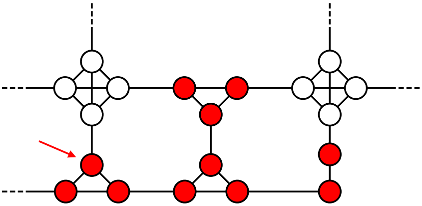
proportion of its infected neighbors $> q = \frac{1}{4}$

INITIAL GRAPH	GRAPH WITH CLIQUES
	

Random graph with separate communities

- At the beginning, one infected vertex (= the seed of the epidemic)
- At each step, each vertex becomes infected if :

proportion of its infected neighbors $> q = \frac{1}{4}$

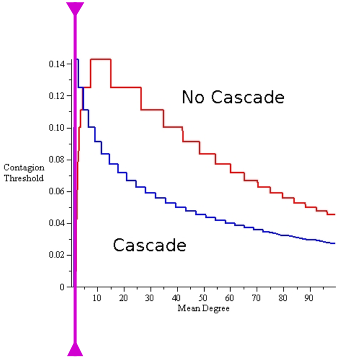
INITIAL GRAPH	GRAPH WITH CLIQUES
	

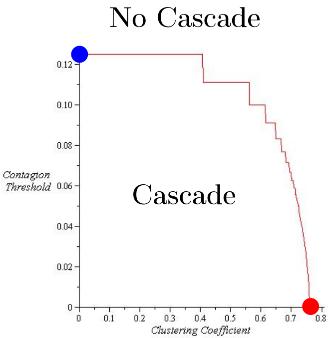
$\Longrightarrow$ Clustering decreases the cascade size.

Effect of Clustering on the Contagion Threshold

Asymptotic degree distribution :
$\tilde{p}_k \propto k^{-\tau} e^{-k/50}$

Mean degree $\tilde{\lambda} \approx 1.65$

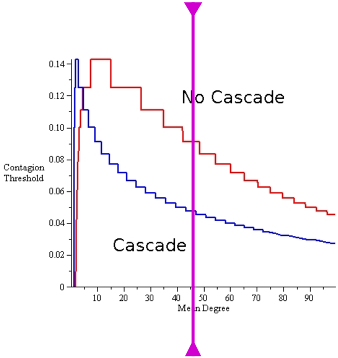




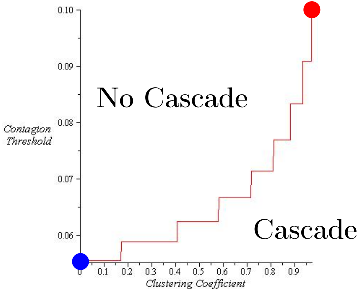
— Graph with maximal clustering coefficient
— Graph with no clustering

Effect of Clustering on the Contagion Threshold

Asymptotic degree distribution :
$\tilde{p}_k \propto k^{-\tau} e^{-k/50}$



Mean degree $\tilde{\lambda} \approx 46$

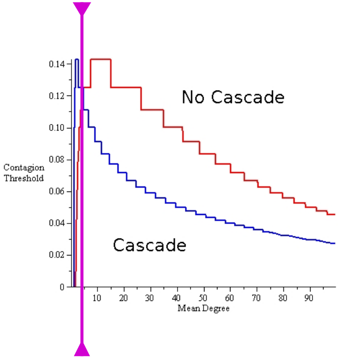


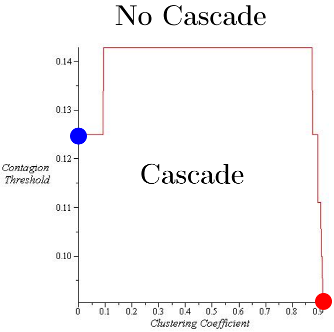
— Graph with maximal clustering coefficient
— Graph with no clustering

Effect of Clustering on the Contagion Threshold

Asymptotic degree distribution :
$\tilde{p}_k \propto k^{-\tau} e^{-k/50}$

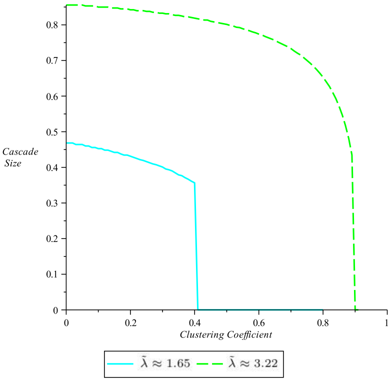
Mean degree $\tilde{\lambda} \approx 3.22$

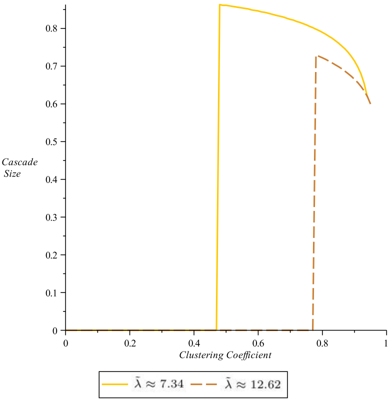




— Graph with maximal clustering coefficient
— Graph with no clustering

Effect of Clustering on the Cascade Size



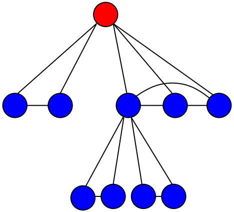


Asymptotic degree distribution : $\tilde{p}_k \propto k^{-\tau} e^{-k/50}$

Random graph with overlapping communities

- At the beginning, the root is infected
- Each vertex becomes infected if :

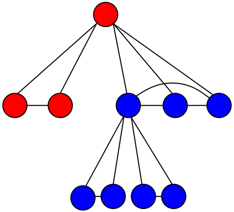
proportion of its infected neighbors $> q = \frac{1}{5}$



Random graph with overlapping communities

- At the beginning, the root is infected
- Each vertex becomes infected if :

proportion of its infected neighbors $> q = \frac{1}{5}$

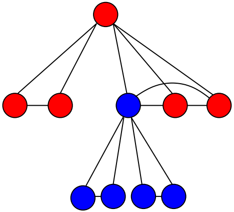


- Independence among the different branches

Random graph with overlapping communities

- At the beginning, the root is infected
- Each vertex becomes infected if :

proportion of its infected neighbors $> q = \frac{1}{5}$

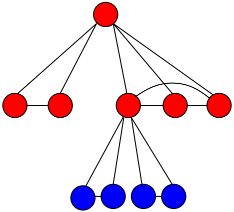


- Independence among the different branches
- Inside a clique : first ones = those with smaller degree

Random graph with overlapping communities

- At the beginning, the root is infected
- Each vertex becomes infected if :

proportion of its infected neighbors $> q = \frac{1}{5}$

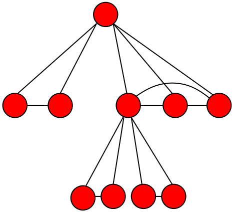


- Independence among the different branches
- Inside a clique : first ones = those with smaller degree
- Then : possibly other vertices infected inside the clique

Random graph with overlapping communities

- At the beginning, the root is infected
- Each vertex becomes infected if :

proportion of its infected neighbors $> q = \frac{1}{5}$



- Independence among the different branches
- Inside a clique : first ones = those with smaller degree
- Then : possibly other vertices infected inside the clique

Random graph with overlapping communities

- Independence among the different branches
  $\Longrightarrow$ Study independently each branches

Random graph with overlapping communities

- Independence among the different branches
  $\Longrightarrow$ Study independently each branches
- Inside a clique : first ones = those with smaller degree
  $\Longrightarrow$ Order the vertices according to their degree

Random graph with overlapping communities

- Independence among the different branches
  $\Longrightarrow$ Study independently each branches
- Inside a clique : first ones = those with smaller degree
  $\Longrightarrow$ Order the vertices according to their degree
- Then : possibly other vertices infected inside the clique
  $\Longrightarrow$ i-th vertex infected if needs less than i infected neighbors

Random graph with overlapping communities

- Independence among the different branches
  $\Longrightarrow$ Study independently each branches
- Inside a clique : first ones = those with smaller degree
  $\Longrightarrow$ Order the vertices according to their degree
- Then : possibly other vertices infected inside the clique
  $\Longrightarrow$ i-th vertex infected if needs less than i infected neighbors
- Need to know the number of children of each vertex
  $\Longrightarrow$ Use of a multi-type branching process

Random graph with overlapping communities

Definition of a multi-type (and alternating) branching process such that :

- nb of vertices at a given generation
  = nb of infected vertices (in the original process)
- type of a vertex
  = its number of grandchildren (in the original process)
- type of a hyper-edge
  = its number of children (in the original process)

Random graph with overlapping communities

Definition of a multi-type (and alternating) branching process such that :

- nb of vertices at a given generation
  = nb of infected vertices (in the original process)
- type of a vertex
  = its number of grandchildren (in the original process)
- type of a hyper-edge
  = its number of children (in the original process)

Mean number of grandchildren of a type y of a vertex of type x
$\Longrightarrow$ Matrix with largest eigenvalue ρ.

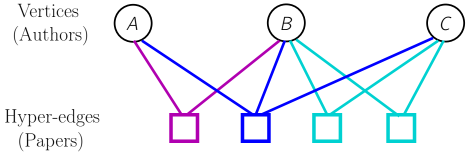
Theorem (CONTAGION THRESHOLD)

- If $q \geq 1/2$, then there is no cascade.
- If $q < 1/2$, then there is a cascade if and only if either $\rho > 1$ or $p_2 = q_2 = 1$.

Variant of the epidemic model on random hypergraphs

With the previous model : individuals equally influenced by each neighbor

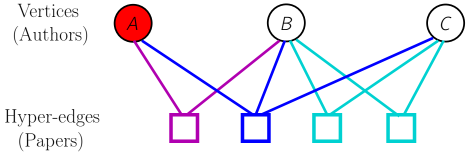
Epidemic model that takes into account the number of papers in common



Variant of the epidemic model on random hypergraphs

With the previous model : individuals equally influenced by each neighbor

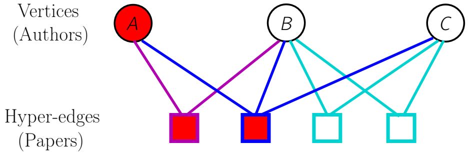
Epidemic model that takes into account the number of papers in common



Variant of the epidemic model on random hypergraphs

With the previous model : individuals equally influenced by each neighbor

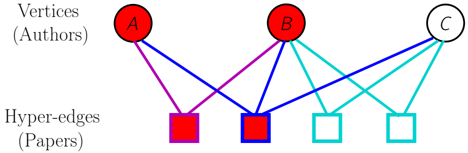
Epidemic model that takes into account the number of papers in common



Variant of the epidemic model on random hypergraphs

With the previous model : individuals equally influenced by each neighbor

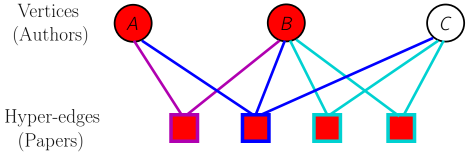
Epidemic model that takes into account the number of papers in common



Variant of the epidemic model on random hypergraphs

With the previous model : individuals equally influenced by each neighbor

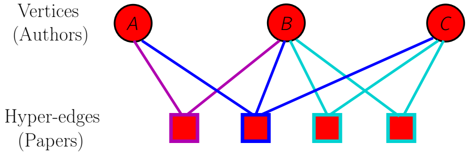
Epidemic model that takes into account the number of papers in common



Variant of the epidemic model on random hypergraphs

With the previous model : individuals equally influenced by each neighbor

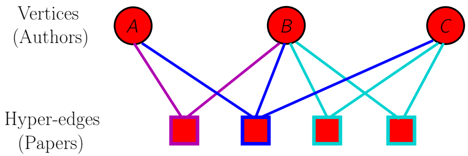
Epidemic model that takes into account the number of papers in common



Variant of the epidemic model on random hypergraphs

With the previous model : individuals equally influenced by each neighbor

Epidemic model that takes into account the number of papers in common



- Initially : positive fraction of infected individuals
- Definition of a Markov chain
- Differential equation approximations for Markov chains
- Application : upper bound for the largest component size

Outline

Conclusion and perspectives

Random graphs with separate communities :

- Model of random graphs with a given degree distribution, and a tunable clustering coefficient
- Effect of clustering on the diffusion model :
  - Clustering increases the diffusion threshold
  - Clustering decreases the cascade size
- Effect of clustering on the contagion model :
  - Clustering decreases the contagion threshold for low values of the mean degree, while the opposite happens in the high values regime
  - Clustering decreases the cascade size (when a cascade is possible)
- Non-negligeable proportion of infected individuals at the start

Conclusion and perspectives

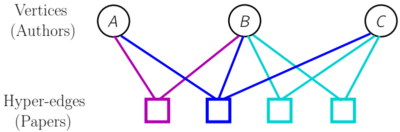
Random graphs with overlapping communities :

- On the equivalent branching process :
  - Clustering decreases the diffusion threshold
  - Use of a multi-type branching process to study the contagion
- Variant of the epidemic model that takes into account the number of groups in common

Conclusion and perspectives

Random graphs with overlapping communities :

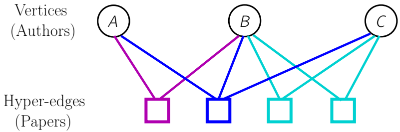
- On the equivalent branching process :
  - Clustering decreases the diffusion threshold
  - Use of a multi-type branching process to study the contagion
- Variant of the epidemic model that takes into account the number of groups in common
- No redundancy in this model :



Conclusion and perspectives

Random graphs with overlapping communities :

- On the equivalent branching process :
  - Clustering decreases the diffusion threshold
  - Use of a multi-type branching process to study the contagion
- Variant of the epidemic model that takes into account the number of groups in common
- No redundancy in this model :



Thanks for your attention !

References

F. Ball, D. Sirl, and P. Trapman.
Analysis of a stochastic sir epidemic on a random network incorporating household structure.
Mathematical Biosciences, 224(2) :53–73, 2010.

T. Britton, M. Deijfen, A. N. Lagerås, and M. Lindholm.
Epidemics on random graphs with tunable clustering.
J. Appl. Probab., 45(3) :743–756, 2008.

A. Hackett, S. Melnik, and J. P. Gleeson.
Cascades on a class of clustered random networks.
Physical Review E, 83, 2011.

M. Lelarge.
Diffusion and cascading behavior in random networks.
Games and Economic Behavior, 75(2) :752–775, 2012.

M. E. J. Newman.
Properties of highly clustered networks.
Phys. Rev. E, 68(2) :026121, Aug 2003.

P. Trapman.
On analytical approaches to epidemics on networks.
Theoretical Population Biology, 71(2) :160–173, 2007.

D. J. Watts and S. H. Strogatz.
Collective dynamics of 'small-world' networks.
Nature, 393(6684) :440–442, June 1998.