Impact of Clustering on Epidemics in Random Networks

Joint work with Marc Lelarge

TREC

20 December 2011

Outline

Outline

1. Introduction : Social Networks and Epidemics
2. Random Graph Model
3. Epidemic Model (from Game Theory)

Social networks

Examples :

- Population
- Internet
- World-Wide Web
- Collaborations
- ...

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

- V = set of vertices
- E = set of edges

Social networks

Examples :

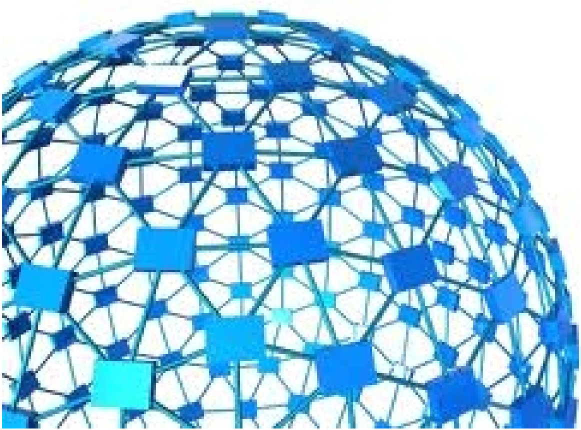
- Population *(individuals)*
- Internet *(autonomous systems)*
- World-Wide Web *(sites)*
- Collaborations *(individuals)*
- ...

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

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

Large networks

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

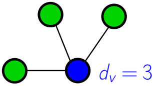


Large networks

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 $\boldsymbol{p} = (p_k)_{k \geq 0}$

Define a model of *random* graphs

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

Define a model of *random* graphs

- having (asymptotically) the observed properties :
  - *scale-free* networks $\Leftrightarrow$ power law degree distribution
    i.e. there exists $\tau > 0$ such that, for all $k \geq 0$, $p_k \propto k^{-\tau}$
    (small number of nodes having a large number of edges)
  - networks with *clustering*
    ("The friends of my friends are my friends", Newman, '03)
- tractable

Epidemics on random graphs

Spread of :

- Human diseases
- Computer viruses
- Information
- New ideas and practices (diffusion of innovations)
- ...

Two types of epidemic models :

- Diffusion (classical SI model)
- Contagion (from Game Theory)

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)

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

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 $\not\rightarrow$

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

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 $\not\rightarrow$

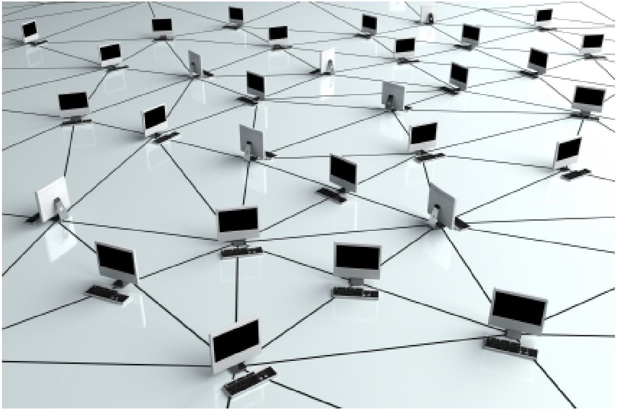
Situation	Payoff (for both users)
	q
	$1 - q > q$
	0

Total payoff
= sum of payoffs from
all your neighbors

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

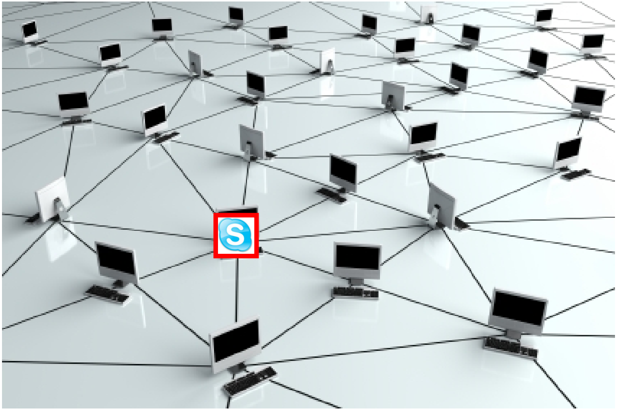
Switch from (MSN logo) to (Skype logo) $\Leftrightarrow \frac{|\text{Neighbors using Skype}|}{|\text{Neighbors}|} > q$

Example : $q = 0.2$



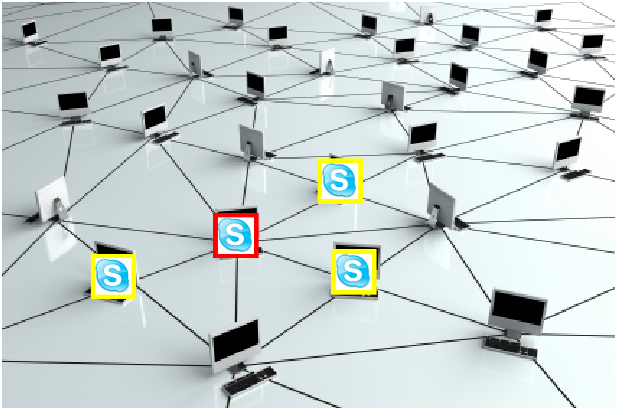
Switch from to $\Leftrightarrow \frac{|\text{Neighbors using Skype}|}{|\text{Neighbors}|} > q$

Example : $q = 0.2$



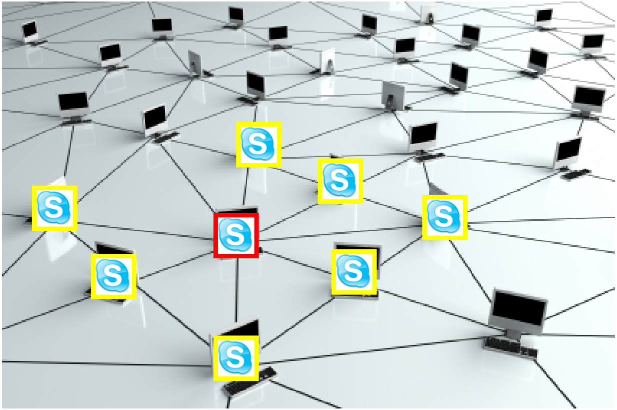
Switch from to $\Leftrightarrow \frac{|\text{Neighbors using Skype}|}{|\text{Neighbors}|} > q$

Example : $q = 0.2$



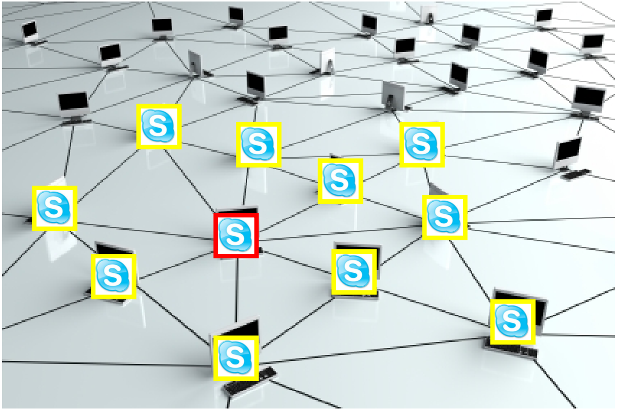
Switch from  to  $\Leftrightarrow \frac{|\text{Neighbors using Skype}|}{|\text{Neighbors}|} > q$

Example : $q = 0.2$



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

Example : $q = 0.2$



Population size $\to \infty$:
Final nb of infected nodes negligeable or not / population size ?

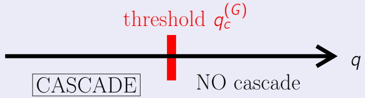
Given graph $G = (V, E)$, parameter q varies

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

More precisely :

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

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

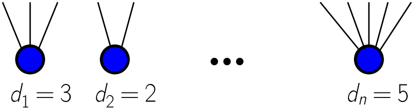


Outline

(i) Start from a uniform graph with given vertex degrees
(ii) Add clustering

(i) Original graph (with given vertex degrees) :

$V = \{1, ..., n\}$
d_i = degree of node i



$G(n, \boldsymbol{d})$ = uniform random graph with vertex degrees $\boldsymbol{d} = (d_i)_{i=1}^n$
Asymptotic degree distribution $\boldsymbol{p} = (p_k)_{k \geq 0}$

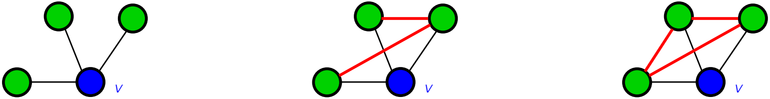
Ref. : (Lelarge,'11) for the study of contagion and diffusion models on graphs with given vertex degrees

(ii) Clustering Coefficient of $G = (V, E)$:

$$C^{(G)} := \frac{3 \times \text{nb of triangles}}{\text{nb of connected triples}} = \frac{\sum_v P_v}{\sum_v N_v}$$

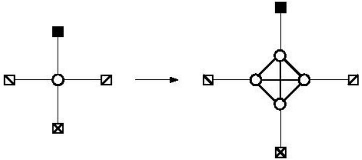
$P_v :=$ nb of pairs of neighbors of v sharing an edge together,
$N_v :=$ nb of pairs of neighbors of v : $N_v = d_v(d_v - 1)/2$.

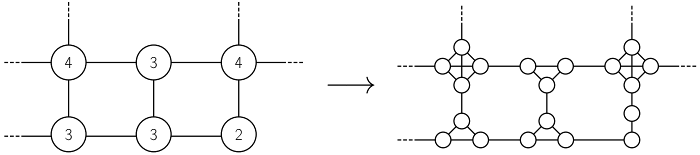
Example : $N_v = 3$



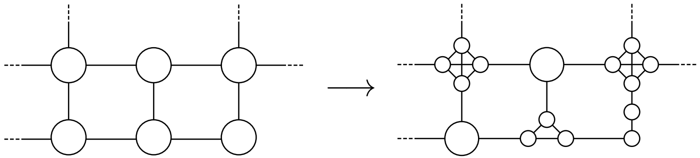
$P_v = 0$ $P_v = 2$ $P_v = 3$

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





- 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).



- $\tilde{G}(n, \boldsymbol{d}, \gamma)$ = resulting random graph (with additional cliques)
  Similar model : (Trapman,'07), (Gleeson,'09)
- **Particular cases :**
  - $\gamma = 0 \Rightarrow \tilde{G}(n, \boldsymbol{d}, \gamma) = G(n, \boldsymbol{d})$,
  - $\gamma = 1 \Rightarrow$ all vertices in $G(n, \boldsymbol{d})$ have been replaced by cliques.

We can compute :

- New asymptotic degree distribution $\tilde{\boldsymbol{p}} = (\tilde{p}_k)_{k\geq 0}$
- Asymptotic clustering coefficient $C > 0$

What we are interested in... :

Theorems : epidemic threshold and cascade size

Comparing epidemics in two graphs with

- SAME asymptotic degree distribution $\tilde{\boldsymbol{p}} = (\tilde{p}_k)_{k\geq 0}$
- DIFFERENT clustering coefficients

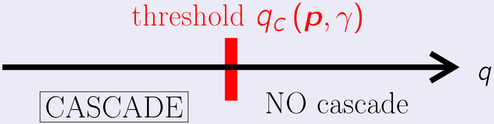
Outline

Contagion model with parameter q, on the (random) graph $\tilde{G}(n, \boldsymbol{d}, \gamma)$:

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

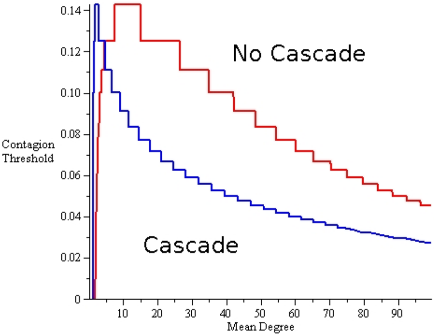
proportion of its infected neighbors $> q$

Theorem



Contagion Threshold (q_c) vs Mean Degree

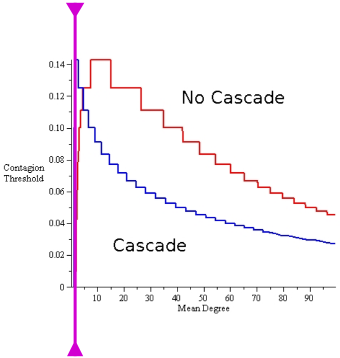
Asymptotic degree distribution : $\tilde{p}_k \propto k^{-\tau} e^{-k/50}$

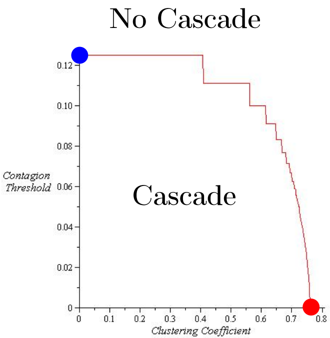


— Graph with clustering (cliques)
— Graph with no clustering

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

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

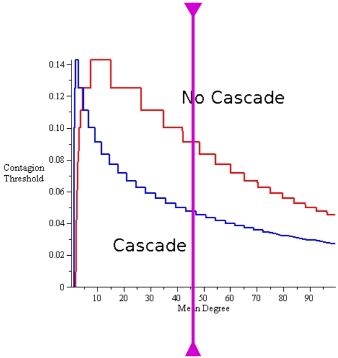


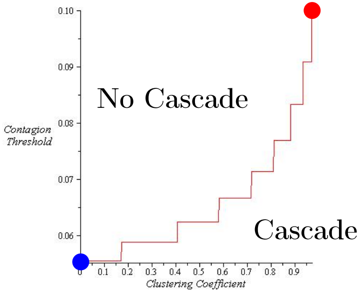


— Graph with maximal clustering coefficient
— Graph with no clustering

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

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

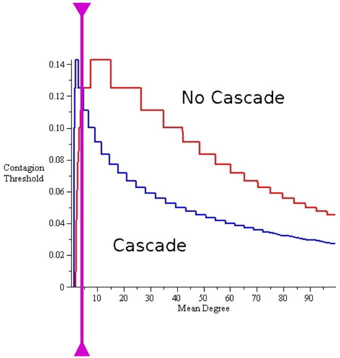


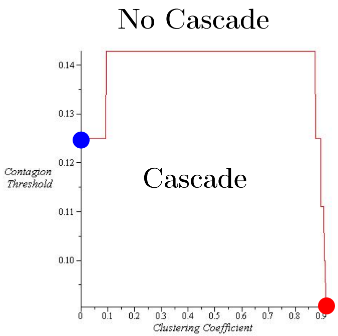


— Graph with maximal clustering coefficient
— Graph with no clustering

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

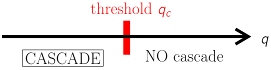
Mean degree $\tilde{\lambda} \approx 3.22$





— Graph with maximal clustering coefficient
— Graph with no clustering

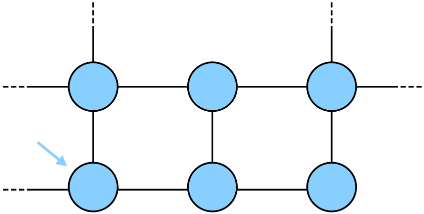
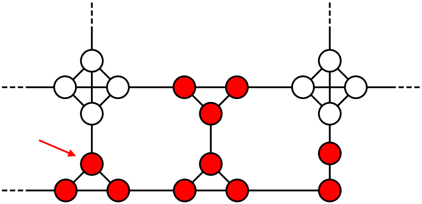
Idea of the proof for



Contagion model with parameter q, on the (random) graph $\tilde{G}(n, \boldsymbol{d}, \gamma)$:

- 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
	

Conclusion

- Clustering decreases the contagion threshold for low values of the mean degree, while the opposite happens in the high values regime
- Clustering increases the diffusion threshold in random regular graphs

Conclusion

- Clustering decreases the contagion threshold for low values of the mean degree, while the opposite happens in the high values regime
- Clustering increases the diffusion threshold in random regular graphs

Thanks for your attention !

References

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.

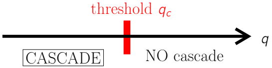
M. Lelarge.
Diffusion and cascading behavior in random networks.
under revision for Games and Economic Behavior, arxiv/0805.3155, 2009.

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.

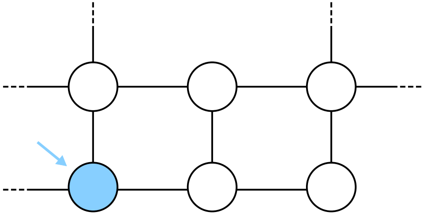
Appendix I : Idea of the proof for



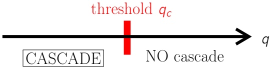
Contagion model with parameter q, on the (random) graph $\tilde{G}\,(n, \boldsymbol{d}, \gamma)$:

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

proportion of its infected neighbors $> q = \frac{1}{3}$

INITIAL GRAPH	GRAPH WITH CLIQUES
	

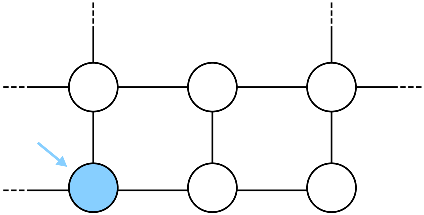
Appendix I : Idea of the proof for



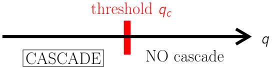
Contagion model with parameter q, on the (random) graph $\tilde{G}(n, \boldsymbol{d}, \gamma)$:

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

proportion of its infected neighbors $> q = \frac{1}{3}$

INITIAL GRAPH	GRAPH WITH CLIQUES
	

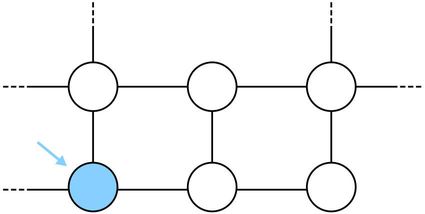
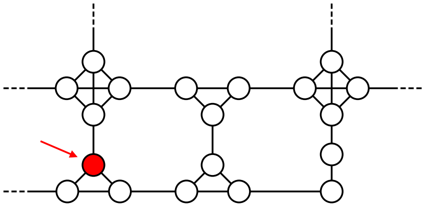
Appendix I : Idea of the proof for



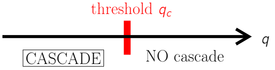
Contagion model with parameter q, on the (random) graph $\tilde{G}(n, \boldsymbol{d}, \gamma)$:

- 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
	

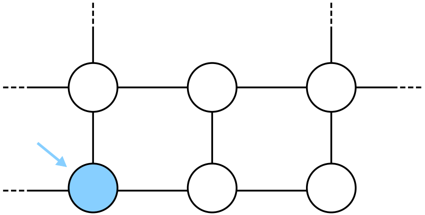
Appendix I : Idea of the proof for



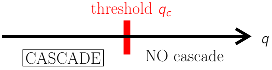
Contagion model with parameter q, on the (random) graph $\tilde{G}(n, \boldsymbol{d}, \gamma)$:

- 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
	

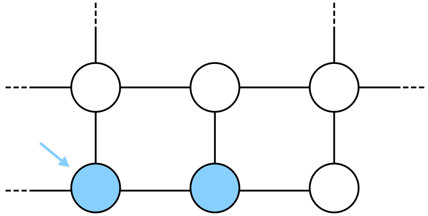
Appendix I : Idea of the proof for



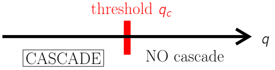
Contagion model with parameter q, on the (random) graph $\tilde{G}(n, \boldsymbol{d}, \gamma)$:

- 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
	

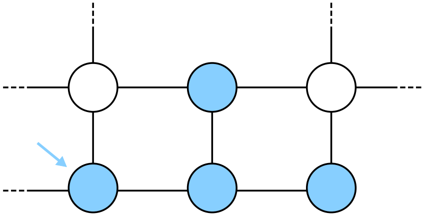
Appendix I : Idea of the proof for



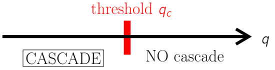
Contagion model with parameter q, on the (random) graph $\tilde{G}(n, \boldsymbol{d}, \gamma)$:

- 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
	

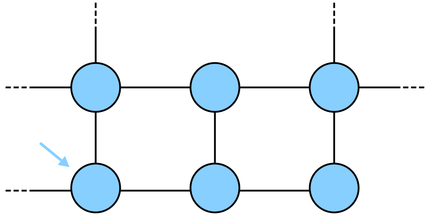
Appendix I : Idea of the proof for



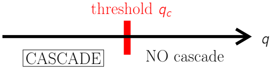
Contagion model with parameter q, on the (random) graph $\tilde{G}(n, \boldsymbol{d}, \gamma)$:

- 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
	

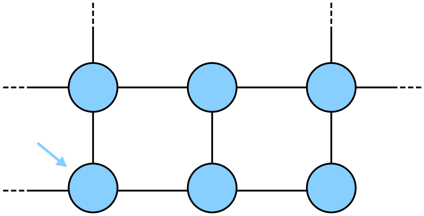
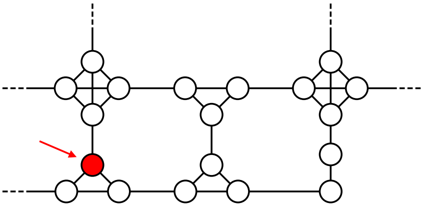
Appendix I : Idea of the proof for



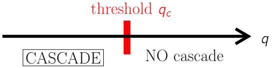
Contagion model with parameter q, on the (random) graph $\tilde{G}(n, \boldsymbol{d}, \gamma)$:

- 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
	

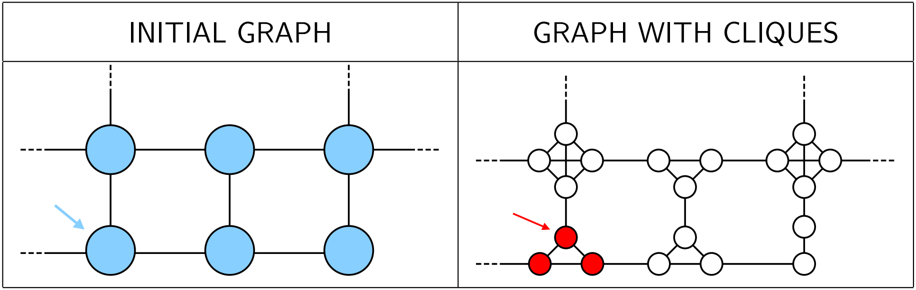
Appendix I : Idea of the proof for



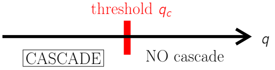
Contagion model with parameter q, on the (random) graph $\tilde{G}(n, \boldsymbol{d}, \gamma)$:

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

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



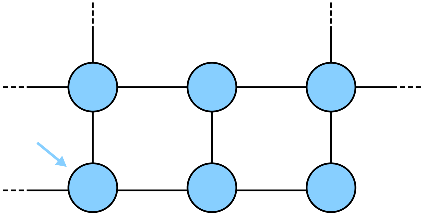
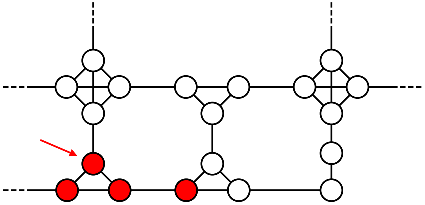
Appendix I : Idea of the proof for



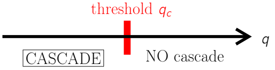
Contagion model with parameter q, on the (random) graph $\tilde{G}(n, \boldsymbol{d}, \gamma)$:

- 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
	

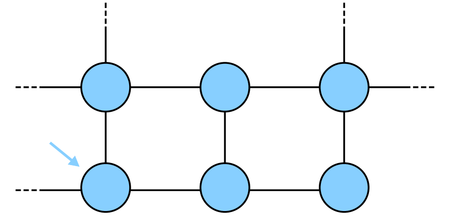
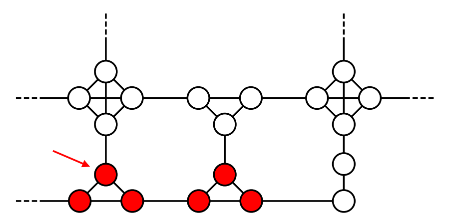
Appendix I : Idea of the proof for



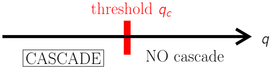
Contagion model with parameter q, on the (random) graph $\tilde{G}(n, \boldsymbol{d}, \gamma)$:

- 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
	

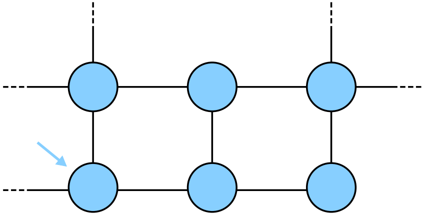
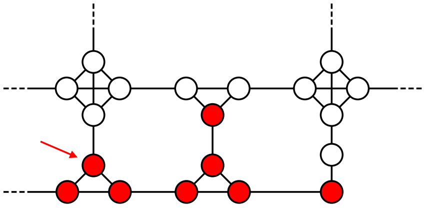
Appendix I : Idea of the proof for



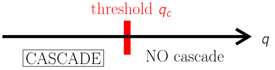
Contagion model with parameter q, on the (random) graph $\tilde{G}(n, \boldsymbol{d}, \gamma)$:

- 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
	

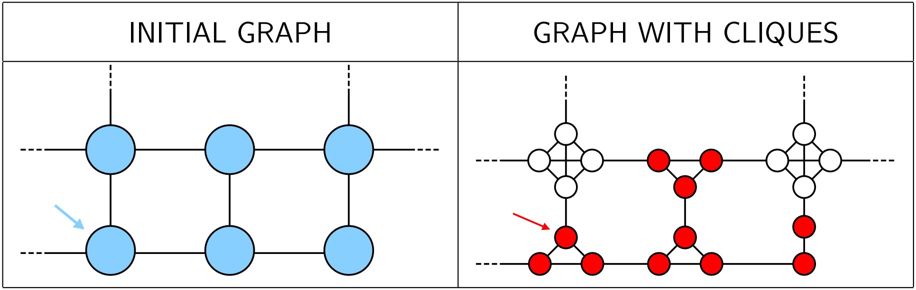
Appendix I : Idea of the proof for



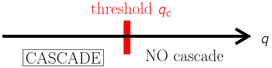
Contagion model with parameter q, on the (random) graph $\tilde{G}(n, \boldsymbol{d}, \gamma)$:

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

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



Appendix I : Idea of the proof for



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

Cascade $\Leftrightarrow \mathcal{P}^{(n)}$ large

Appendix II

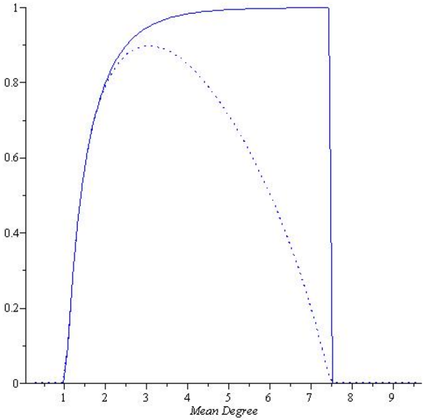
Fraction of infected neighbors needed to become infected :

$q = 0.15$ (fixed)

$\boldsymbol{p} \sim \mathrm{Po}(\lambda)$ and $\gamma = 0$

$\Rightarrow \tilde{p}_r = p_r = \frac{e^{-\lambda}\lambda^r}{r!}$

and $C = 0$



$\cdots$ Pivotal players in the graph with no clustering

— Cascade size in the graph with no clustering

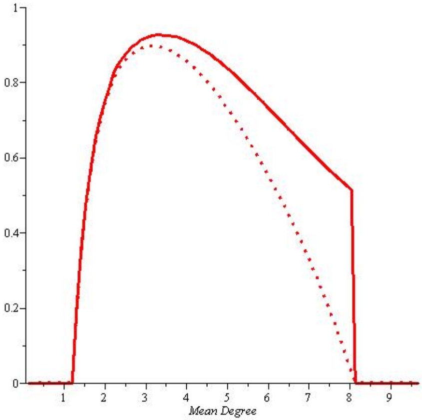
Appendix II

Fraction of infected neighbors needed to become infected :

$q = 0.15$ (fixed)

$\boldsymbol{p} \sim \mathrm{Po}(\lambda)$ and $\gamma = 0.2$

$\Rightarrow \tilde{p}_r = \frac{0.2r+0.8}{0.2\lambda+0.8} \frac{e^{-\lambda}\lambda^r}{r!}$
and $C = \frac{0.2\lambda}{0.2\lambda+1.2} > 0$



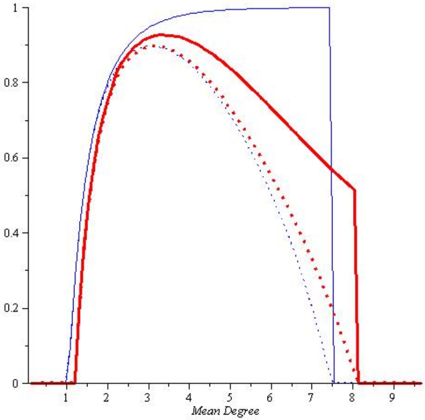
··· Pivotal players in the graph with positive clustering

— Cascade size in the graph with positive clustering

Appendix II

Fraction of infected neighbors needed to become infected :

$$q = 0.15 \text{ (fixed)}$$



⋯ Pivotal players in the graph with no clustering

— Cascade size in the graph with no clustering

⋯ Pivotal players in the graph with positive clustering

— Cascade size in the graph with positive clustering