Fuzzy Conventions

Marcin Pęski

University of Toronto

March 1, 2022

Introduction

- Social interactions, positive externalities.
  - maintaining neat front yard,
  - engaging in criminal activity,
  - technology adoption.
- A typical result: emergence of a (homogeneous) convention.
- But, in reality, conventions are often fuzzy:
  - there are countries where multiple languages are used,
  - married couples that use both IPhone and Android.
- People care not only about their neighbors, but they differ wrt. tastes, preferences.

Introduction

- Binary coordination games on networks with random utility,
- (Statistical) heterogeneous preferences: i.i.d payoff shocks,
- I am interested in the set of average (i.e., aggregate) behavior $x \in [0, 1]$
  - in static,
  - complete information equilibria,
  - when each agent number of connection is large.
- **Q**: What can we say about equilibrium sets? How do they depend on the network?

Introduction

Model

- agents i, j live on a network with weights $g_{ij} = g_{ji} \geq 0$,
  - $g_i = \sum_j g_{ij}$ is degree of agent i,
- payoffs: $\sum_{j \neq i} g_{ij} u(a_i, a_j, \varepsilon_i)$,
  - each i chooses $a_i \in \{0, 1\}$,
  - i.i.d. payoff shocks $\varepsilon_i \sim F$,
  - positive externalities: $u(., ., \varepsilon_i)$ has increasing differences for each ε,
- average behavior $\mathrm{Av}(a) = \frac{1}{\sum_i g_i} \sum_i g_i a_i$,
- equilibrium set

$$\mathrm{Eq}(g, \varepsilon) = \{\mathrm{Av}(a) : a \text{ is a Nash equilibrium in game } G(g, \varepsilon)\},$$

Introduction

Model

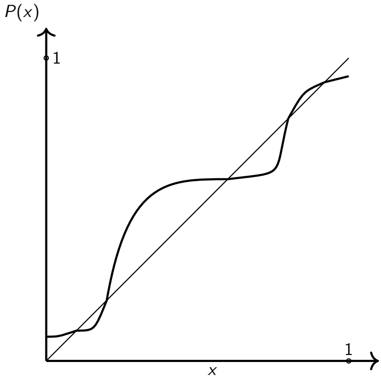
- Object of interest: $\lim \mathrm{Eq}\left(g, .\right)$ as
  - $d\left(g\right) = \max_{i,j} \frac{g_{ij}}{g_i} \to 0$ - large degrees,
  - $w\left(g\right) = \max_{i,j} \frac{g_i}{g_j} < w_{\max} < \infty$ is bounded - not too much inequality.

- 4 theorems that characterize the largest and the smallest possible limit of equilibrium sets across all networks.

Introduction

Results

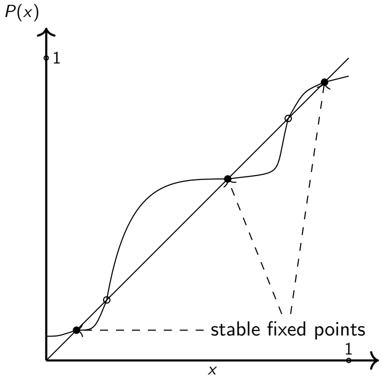
- Let $P(x) = F\{\varepsilon : u(1, x, \varepsilon) \geq u(0, x, \varepsilon)\}$,
- fraction of agents for whom 1 is a best response if x agents play 1 in a continuum toy version.



Introduction

Results

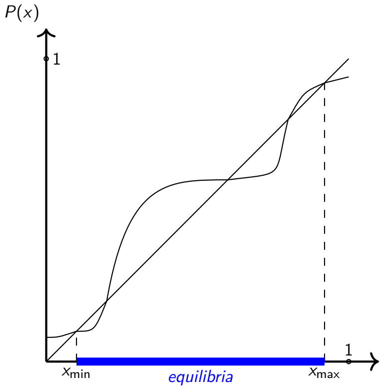
- Let $P(x) = F\{\varepsilon : u(1, x, \varepsilon) \geq u(0, x, \varepsilon)\}$,
- fraction of agents for whom 1 is a best response if x agents play 1 in a continuum toy version.



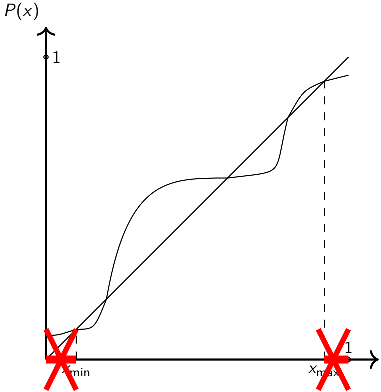
Introduction

Results

- **Theorem 1**: There exists a sequence of networks such that the limit equilibrium set is $[x_{\min}, x_{\max}]$.



- **Theorem 2**: All limit equilibrium sets are contained in $[x_{\min}, x_{\max}]$.

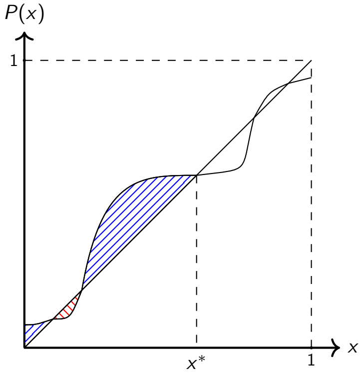


Introduction

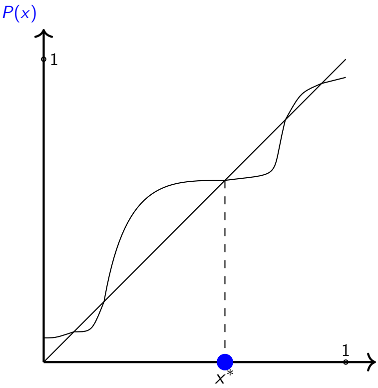
Results

- Define *random utility (RU-) dominant* outcome

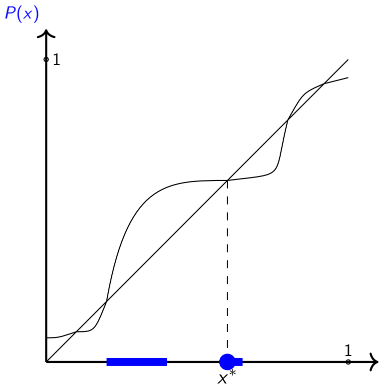
$$x^* \in \arg\max_x \int_0^x \left(y - P^{-1}(y)\right) dy.$$



- **Theorem 3**: There exists a sequence of networks such that the limit equilibrium set is $\{x^*\}$.



- **Theorem 4**: All limit equilibrium sets contain x^*.



Introduction

Results

- 4 theorems that characterize the largest and the smallest possible limit of equilibrium sets across all networks.
- The largest set: partial identification theory.
- The smallest set: equilibrium selection theory.

Introduction

Literature

- Emergence of conventions: evolutionary approach
  - risk-dominance (Harsanyi-Selten 88),
  - complete networks (Kandori, Mailath Rob 93), (Young 93),
  - line (Ellison 93) and some other networks (Ellison 00),
  - all networks: (Peski 10).
- Contagion (Morris 00):
  - some networks (lattices) admit contagion: a finite group of agents can spread risk-dominant behavior to the rest of the network,
  - contagion only works for risk-dominant actions.
- Here,
  - random utility instead of noise (or a perturbation),
  - static solution concept.

Theorem 1

Notation

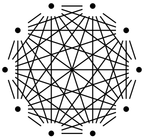
- Define a profile of neighborhood fractions β^a: for each i

$$\beta_i^a = \frac{1}{g_i} \sum_{i \neq j} g_{ij} a_j,$$

- $A \subseteq_\eta B$ if for each $a \in A$, there is $b \in B$ st. $|a - b| \leq \eta$,
- $A =_\eta B$ if $A \subseteq_\eta B$ and $B \subseteq_\eta A$.

Theorem 1

- Let g^n_{complete} be the complete graph with n nodes



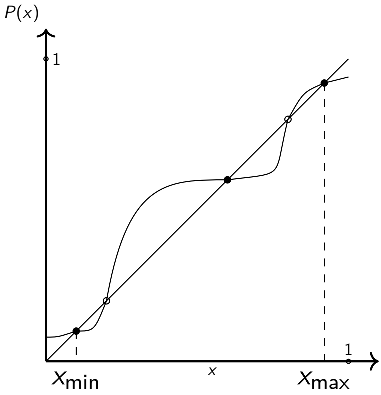
Theorem

If x is a stable fixed point of P, then, for each $\eta > 0$,

$$\lim_{n\to\infty} Prob\left(\{x\} \subseteq_\eta Eq\left(g^n_{complete}, \varepsilon\right)\right) \geq 1 - \eta.$$

- very simple proof,

Theorem 1



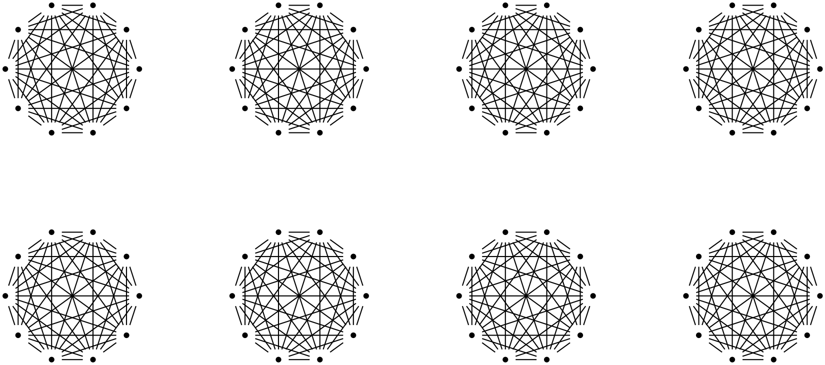
- Generically, $x_{\min}$ and $x_{\max}$ - the smallest and the largest fixed points - are stable.

Theorem 1

Corollary

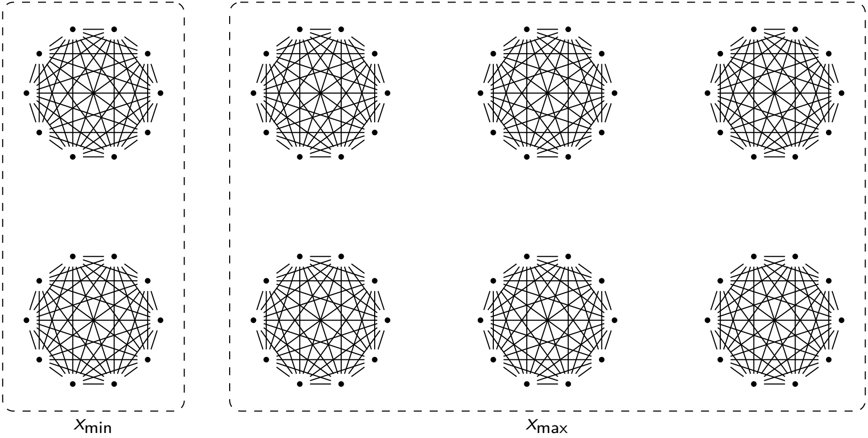
There exists a sequence of graphs g^n such that

$$\lim_{n\to\infty} Prob\left([x_{\min}, x_{\max}] \subseteq_\eta Eq\left(g^n, \varepsilon\right)\right) \geq 1-\eta.$$



Theorem 1

- Here, $x = \frac{2}{8}x_{\text{min}} + \frac{6}{8}x_{\text{max}}$.



Theorem 2

- So far, we showed existence of networks g such that with a large probability,

$$[x_{\text{min}}, x_{\text{max}}] \subseteq_\eta \text{Eq}\left(g, \varepsilon\right).$$

- Next, we show that, for any g st. $d\left(g\right) = \max_{i,j} \frac{g_{ij}}{g_i}$ is sufficiently small,

$$\text{Eq}\left(g, \varepsilon\right) \subseteq_\eta [x_{\text{min}}, x_{\text{max}}].$$

Theorem 2

Theorem

For any $w_{\max} < \infty$, *any sequence of graphs* g_n, *if* $d(g_n) \to 0$ *and* $w(g_n) \le w_{\max}$, *then*

$$\lim_{n\to\infty} Prob\left(Eq\left(g^n, \varepsilon\right) \subseteq_\eta \left[x_{\min}, x_{\max}\right]\right) = 1.$$

Theorem 2

- Proof: surprisingly complicated.
- W.l.o.g., we want to show that, with a large probability, there is no profile a st $\mathrm{Av}(a) > x_{\max}$ and a is an equilibrium.

Theorem 2

- Proof: surprisingly complicated.
- W.l.o.g., we want to show that, with a large probability, there is no profile a st $\mathrm{Av}(a) > x_{\max}$ and a is an equilibrium.
- Bound

$$\begin{aligned} & \mathrm{Prob}\left(\text{there exists } a \text{ st.} \mathrm{Av}(a) \geq x \text{ and } a \text{ is equilibrium}\right) \\ \leq & \# \left\{a : \mathrm{Av}(a) > x\right\} \cdot \mathrm{Prob}\left(a \text{ is equilibrium}\right). \end{aligned}$$

Theorem 2

- Proof: surprisingly complicated.
- W.l.o.g., we want to show that, with a large probability, there is no profile a st $\mathrm{Av}(a) > x_{\max}$ and a is an equilibrium.
- It is easy to show that a is unlikely to be an equilibrium: there exists $\delta > 0$ st. for each a,

$$\mathrm{Prob}\,(a \text{ is equilibrium}) \leq \exp(-\delta N).$$

- But, there are many profiles a:

$$\#\{a : \mathrm{Av}(a) > x\} \sim \exp\left(\left(x \log x + (1-x)\log(1-x)\right) N\right).$$

Theorem 2

- Proof: surprisingly complicated.
- W.l.o.g., we want to show that, with a large probability, there is no profile a st $\mathrm{Av}(a) > x_{\max}$ and a is an equilibrium.
- Problem: there are too many candidate profiles a.
- Observation I: the above proof treats events "a is equilibrium" for all as as disjoint, whereas they are often correlated.
- Observation II: events "a is equilibrium" and "a' is equilibrium" are correlated more if β^a and $\beta^{a'}$ are similar.
  - $\beta_i^a = \frac{1}{g_i} \sum_j g_{ij} a_j$.
- Idea: divide all profiles a into "groups" with similar β^a.

Theorem 2

- Proof: surprisingly complicated.
- W.l.o.g., we want to show that, with a large probability, there is no profile a st $\mathrm{Av}(a) > x_{\max}$ and a is an equilibrium.
- Problem: there are too many candidate profiles a.
- Observation I: the above proof treats events "a is equilibrium" for all as as disjoint, whereas they are often correlated.
- Observation II: events "a is equilibrium" and "a' is equilibrium" are correlated more if β^a and $\beta^{a'}$ are similar.
  - $\beta_i^a = \frac{1}{g_i} \sum_j g_{ij} a_j$.
- Idea: divide all profiles a into "groups" with similar β^a.

Theorem 2

- Proof: surprisingly complicated.
- W.l.o.g., we want to show that, with a large probability, there is no profile a st $\mathrm{Av}(a) > x_{\max}$ and a is an equilibrium.
- Problem: there are too many candidate profiles a.
- Observation I: the above proof treats events "a is equilibrium" for all as as disjoint, whereas they are often correlated.
- Observation II: events "a is equilibrium" and "a' is equilibrium" are correlated more if β^a and $\beta^{a'}$ are similar.
  - $\beta_i^a = \frac{1}{g_i} \sum_j g_{ij} a_j$.
- Idea: divide all profiles a into "groups" with similar β^a.

Theorem 2

- Proof: surprisingly complicated.
- W.l.o.g., we want to show that, with a large probability, there is no profile a st $\text{Av}(a) > x_{\max}$ and a is an equilibrium.
- Problem: there are too many candidate profiles a.
- Observation I: the above proof treats events "a is equilibrium" for all as as disjoint, whereas they are often correlated.
- Observation II: events "a is equilibrium" and "a' is equilibrium" are correlated more if β^a and $\beta^{a'}$ are similar.
  - $\beta_i^a = \frac{1}{g_i} \sum_j g_{ij} a_j$.
- Idea: divide all profiles a into "groups" with similar β^a.

Theorem 2

- The correlation is stronger if $\beta^a \sim \beta^{a'}$, where β^a is a profile of "neighborhood fractions $\beta_i^a = \frac{1}{g_i}\sum_{j\neq i} g_{ij} a_j$), or

$$d\left(\beta_i^a, \beta_i^{a'}\right) = \sqrt{\frac{1}{\sum g_i^2}\sum g_i^2 \left(\beta_i^a - \beta_i^{a'}\right)^2} \text{ is small.}$$

- We show that for each a_0 st. $\mathrm{Av}(a_0) > x$, if δ is sufficiently small and $d(g) \leq \delta$, then

  $$\mathrm{Prob}\left(\{a : d(\beta^a, \beta^{a_0}) \leq \delta\} \text{ contains an equilibrium}\right) \leq \exp(-\delta N).$$

Theorem 2

- Set of "neighborhood fraction" profiles

$$\mathcal{B} = \{\beta^a : a \text{ is a profile}\}.$$

- $\mathcal{N}(\mathcal{B}, \delta)$ is the smallest n such that there exists $b_1, ..., b_n \in \mathcal{B}$ st. $\mathcal{B}$ can be covered with balls radius δ and centers at b_i (metric entropy).
- For some constant $c > 0$,

$$\mathcal{N}(\mathcal{B}, \delta) \leq \exp\left(c\frac{1}{\delta^2} d(g) N\right).$$

Theorem 2

$$
\begin{aligned}
&\operatorname{Prob}\left(\left\{a : d\left(\beta^{a}, \beta\right) \leq \delta\right\} \text{ contains an equilibrium}\right) \\
\leq &\mathcal{N}\left(\mathcal{B}, \delta\right) \cdot \sup_{a_0 : \operatorname{Av}(a_0) > x} \operatorname{Prob}\left(\left\{a : d\left(\beta^{a}, \beta^{a_0}\right) \leq \delta\right\} \text{ contains an equilibrium} \right. \\
\leq &\exp\left(c \frac{1}{\delta^2} d\left(g\right) N - \delta N\right),
\end{aligned}
$$

which is small if $d\left(g\right)$ is small enough.

Theorem 3

Random utility dominant outcome

- So far, we characterized a tight upper bound on the equilibrium set.
- Next, we turn to a lower bound.

Theorem 3

Random utility dominant outcome

- Define *random utility (RU-) dominant* outcome

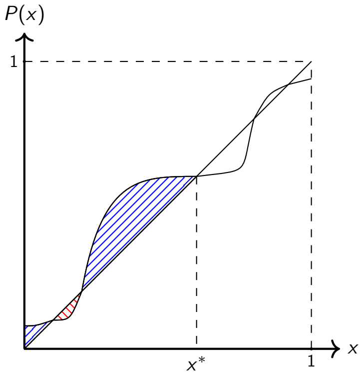
$$x^* \in \arg\max_x \int_0^x \left(y - P^{-1}(y)\right) dy.$$

Theorem 3

Random utility dominant outcome

- Define *random utility (RU-) dominant* outcome

$$x^* \in \arg\max_x \int_0^x \left(y - P^{-1}(y)\right) dy.$$



- RU-outcome is generically a strictly stable fixed point of P.

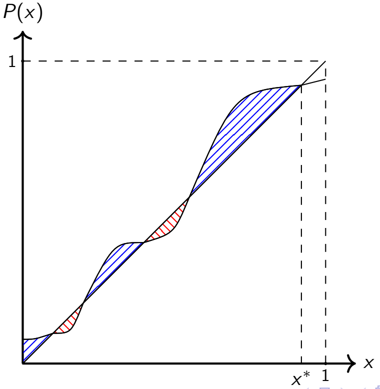
Theorem 3

Random utility dominant outcome

- Define *random utility (RU-) dominant* outcome

$$x^* \in \arg\max_x \int_0^x \left(y - P^{-1}(y)\right) dy.$$

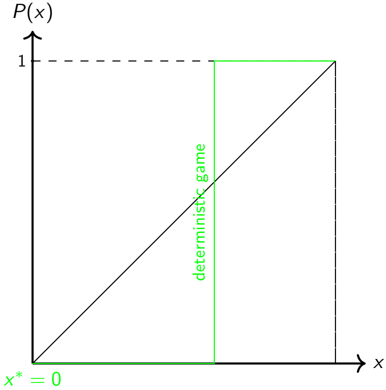
- RU-outcome can be $x_{\min}$ or $x_{\max}$.



Theorem 3

Random utility dominant outcome

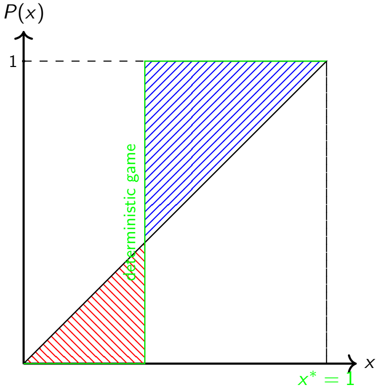
- When game is deterministic, RU-dominance is equivalent to Harsanyi-Selten risk-dominance



Theorem 3

Random utility dominant outcome

- When game is deterministic, RU-dominance is equivalent to Harsanyi-Selten risk-dominance



Theorem 3

Random utility dominant outcome

- Formula

$$x^* \in \arg\max_x \int_0^x \left(y - P^{-1}(y)\right) dy$$

appears in Morris and Shin (06).

  - continuum toy model,
  - observe that the coordination game has a potential,
  - the above outcome maximizes potential,
  - hence it is robust to incomplete information.

Theorem 3

Random utility dominant selection

Theorem

Assume $0 < P(0) < P(1) < 1$. *There exists a sequence of networks* g^n *st. for each* $\eta > 0$,

$$\lim_{n \to \infty} Prob\left(Eq\left(g^n, \varepsilon\right) =_\eta \left\{x^*\right\}\right) \geq 1 - \eta.$$

- For some networks, x^* is the unique average equilibrium behavior.
- The assumption ensures that, for each action, there is a positive probability that the action is dominant.

Theorem 3

Proof

- Network sequence: 2-dimensional lattices
  - line (1-dimensional lattice) is not enough
- Static result, but proof based on best response dynamics.
  - review of contagion arguments (Ellison 93, Blume 93, Morris 00),
  - contagion wave on "toy" line,
  - why line is not enough and why 2-dimensional lattice is.

Theorem 3

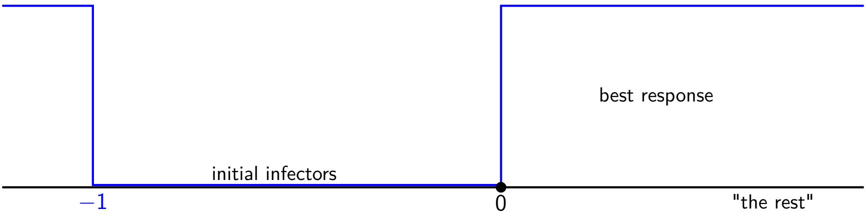
Proof: Review of contagion arguments

- Start with deterministic case, but with small group of initial infectors.
- Assume 0 is risk-dominant.
- We want to show that 0 is the only equilibrium.
- -> contagion.

Theorem 3

Proof: Review of contagion arguments

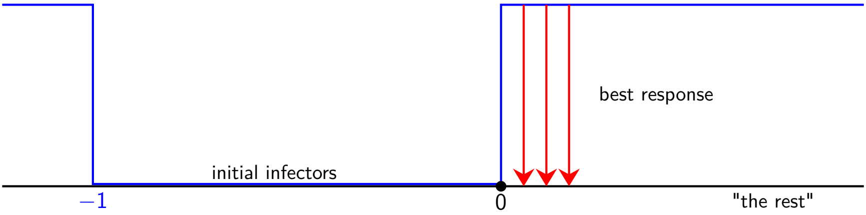
- Ellison 93: suppose that action 0 is risk-dominant,
- initial infectors $-1 \leq i \leq 0$ play 0; the rests play 1,



Theorem 3

Proof: Review of contagion arguments

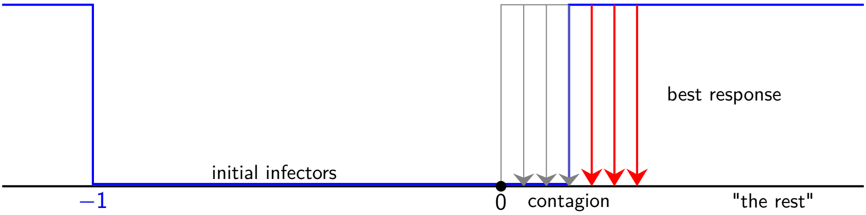
- Ellison 93: suppose that action 0 is risk-dominant,
- initial infectors $-1 \leq i \leq 0$ play 0; the rests play 1,
- best response dynamics



Theorem 3

Proof: Review of contagion arguments

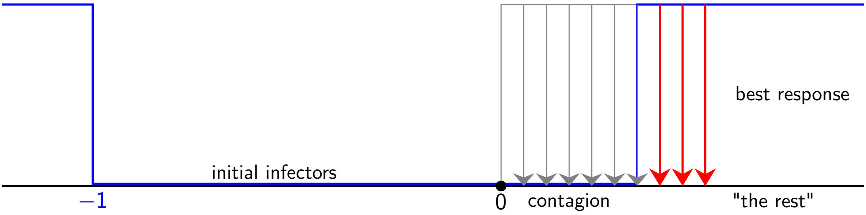
- Ellison 93: suppose that action 0 is risk-dominant,
- initial infectors $-1 \leq i \leq 0$ play 0; the rests play 1,
- best response dynamics -> contagion



Theorem 3

Proof: Review of contagion arguments

- Ellison 93: suppose that action 0 is risk-dominant,
- initial infectors $-1 \leq i \leq 0$ play 0; the rests play 1,
- best response dynamics -> contagion



Theorem 3

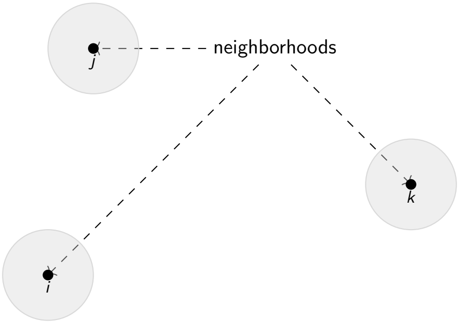
Proof: Review of contagion arguments

- Blume 93, Morris 00 - the same mechanics works on other networks, like 2 (or higher)-dimensional lattices.
- Key step: half of neighbors of “threshold agents” must be infected to spread contagion.

Theorem 3

Proof: Review of contagion arguments

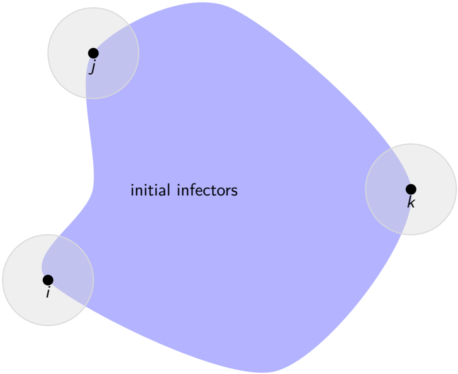
- Blume 93, Morris 00 - the same mechanics works on other networks, like 2 (or higher)-dimensional lattices.
- Key step: half of neighbors of "threshold agents" must be infected to spread contagion.



Theorem 3

Proof: Review of contagion arguments

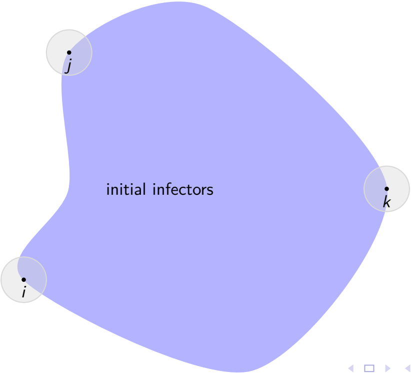
- Blume 93, Morris 00 - the same mechanics works on other networks, like 2 (or higher)-dimensional lattices.
- Key step: half of neighbors of "threshold agents" must be infected to spread contagion.



Theorem 3

Proof: Review of contagion arguments

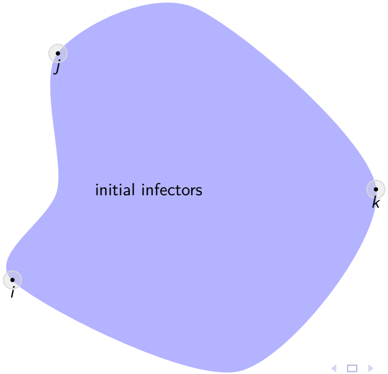
- Blume 93, Morris 00 - the same mechanics works on other networks, like 2 (or higher)-dimensional lattices.
- Key step: half of neighbors of “threshold agents” must be infected to spread contagion
- - $>$ initial infectors must be large enough relative to neighborhoods.



Theorem 3

Proof: Review of contagion arguments

- Blume 93, Morris 00 - the same mechanics works on other networks, like 2 (or higher)-dimensional lattices.
- Key step: half of neighbors of “threshold agents” must be infected to spread contagion
- - $>$ initial infectors must be large enough relative to neighborhoods.



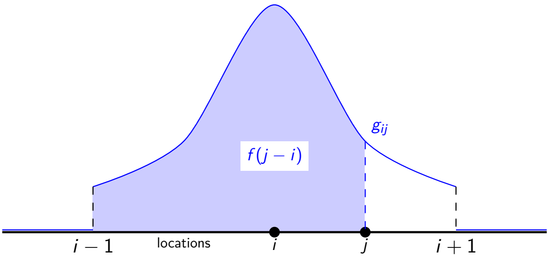
Theorem 3

Proof: Contagion wave on toy line

- Random utility payoffs (so, not deterministic)
- Toy line: Continuum of agents in each location.

Theorem 3

Proof: Contagion wave on line, RU case



- Toy line: agents in location i are connected with agents in location j
  - connection density $g_{ij} = g_{ji} = g_{i+l,j+l}$ for any l,
  - $g_{ij} = 0$ for $j > i+1$,
  - $f(j-i) = \frac{1}{g_i} \int_{i-1}^{j} g_{il} dl$,
  - $f(x) + f(1-x) = 1$.

Theorem 3

Proof: Contagion wave on line, RU case

- For simplicity, assume that $x^* = 0$ is *RU*-dominant, i.e.

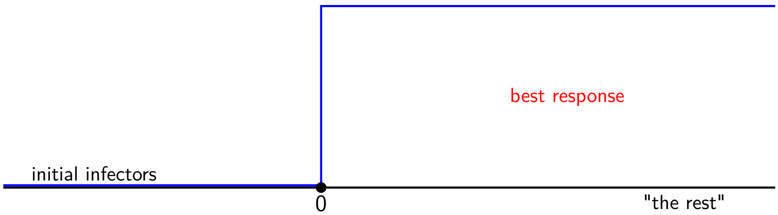
$$\int_0^x \left(y - P^{-1}(y)\right) dy < 0 \text{ for each } x > 0.$$

- Starting from arbitrary profile with a group of initial infectors playing x^*, best response dynamics will spread x^* to the whole line.

Theorem 3

Proof: Contagion wave on line, RU case

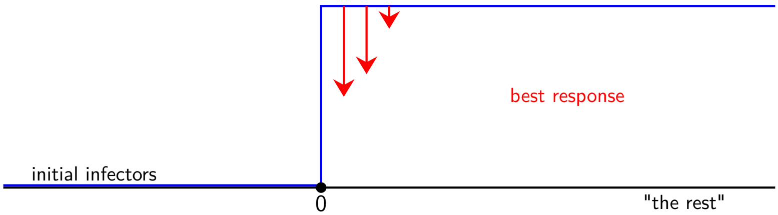
- Initial infectors play $x^* = 0$.



Theorem 3

Proof: Contagion wave on line, RU case

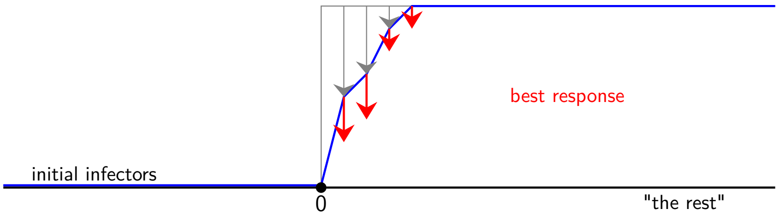
- Initial infectors play $x^* = 0$.



Theorem 3

Proof: Contagion wave on line, RU case

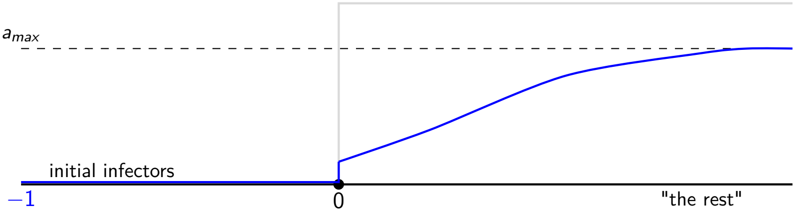
- Initial infectors play $x^* = 0$.



Theorem 3

Proof: Contagion wave on line, RU case

- Suppose that stops before spreading everywhere.



Theorem 3

Proof: Contagion wave on line, RU case

- If the contagion stops, then at each location $i > 0$,

$$a_i \leq P\left(\int a_{i+k} df\,(k)\right).$$

- Taking inverse and integrating by parts

$$P^{-1}(a_i) \leq \int a_{i+k} df\,(k) = \int_0^{a_{\max}} f\,(i-j)\, da_j.$$

- Integrate over $a_i \in [0, a_{\max}]$,

$$\int_0^{a_{\max}} P^{-1}(a_i)\, da_i \leq \int_0^{a_{\max}} \int_0^{a_{\max}} f\,(i-j)\, da_j da_i.$$

Theorem 3

Proof: Contagion wave on line, RU case

- If the contagion stops, then at each location $i > 0$,

$$a_i \leq P\left(\int a_{i+k} df(k)\right).$$

- Taking inverse and integrating by parts

$$P^{-1}(a_i) \leq \int a_{i+k} df(k) = \int_0^{a_{\max}} f(i-j)\, da_j.$$

- Integrate over $a_i \in [0, a_{\max}]$,

$$\int_0^{a_{\max}} P^{-1}(a_i)\, da_i \leq \int_0^{a_{\max}} \int_0^{a_{\max}} f(i-j)\, da_j da_i.$$

Theorem 3

Proof: Contagion wave on line, RU case

- If the contagion stops, then at each location $i > 0$,

$$a_i \leq P\left(\int a_{i+k} df(k)\right).$$

- Taking inverse and integrating by parts

$$P^{-1}(a_i) \leq \int a_{i+k} df(k) = \int_0^{a_{\max}} f(i-j)\, da_j.$$

- Integrate over $a_i \in [0, a_{\max}]$,

$$\int_0^{a_{\max}} P^{-1}(a_i)\, da_i \leq \int_0^{a_{\max}} \int_0^{a_{\max}} f(i-j)\, da_j da_i.$$

Theorem 3

Proof: Contagion wave on line, RU case

- Integrate over $a_i \in [0, a_{\max}]$,

$$\int_0^{a_{\max}} P^{-1}(a_i)\, da_i \leq \int_0^{a_{\max}} \int_0^{a_{\max}} f(i-j)\, da_j da_i$$

Theorem 3

Proof: Contagion wave on line, RU case

- Integrate over $a_i \in [0, a_{\max}]$,

$$
\begin{aligned}
& \int_0^{a_{\max}} P^{-1}(a_i)\, da_i \\
\leq & \int_0^{a_{\max}} \int_0^{a_{\max}} f(i-j)\, da_j da_i \\
= & \frac{1}{2} \int_0^{a_{\max}} f(i-j)\, da_j da_i + \frac{1}{2} \int_0^{a_{\max}} \int_0^{a_{\max}} f(j-i)\, da_j da_i
\end{aligned}
$$

Theorem 3

Proof: Contagion wave on line, RU case

- Integrate over $a_i \in [0, a_{\max}]$,

$$
\begin{aligned}
&\int_0^{a_{\max}} P^{-1}(a_i)\, da_i \\
\leq &\int_0^{a_{\max}} \int_0^{a_{\max}} f(i-j)\, da_j da_i \\
= &\frac{1}{2} \int_0^{a_{\max}} f(i-j)\, da_j da_i + \frac{1}{2} \int_0^{a_{\max}} \int_0^{a_{\max}} f(j-i)\, da_j da_i \\
= &\frac{1}{2} \int_0^{a_{\max}} \int_0^{a_{\max}} \left[f(i-j) + f(j-i) \right] da_j da_i
\end{aligned}
$$

Theorem 3

Proof: Contagion wave on line, RU case

- Integrate over $a_i \in [0, a_{\max}]$,

$$\begin{aligned}
&\int_0^{a_{\max}} P^{-1}(a_i)\, da_i \\
\leq &\int_0^{a_{\max}} \int_0^{a_{\max}} f(i-j)\, da_j da_i \\
= &\frac{1}{2}\int_0^{a_{\max}} f(i-j)\, da_j da_i + \frac{1}{2}\int_0^{a_{\max}} \int_0^{a_{\max}} f(j-i)\, da_j da_i \\
= &\frac{1}{2}\int_0^{a_{\max}} \int_0^{a_{\max}} \left[f(i-j) + f(j-i)\right] da_j da_i
\end{aligned}$$

- Recall that $f(i-j) + f(j-i) = 1$.

Theorem 3

Proof: Contagion wave on line, RU case

- Integrate over $a_i \in [0, a_{\max}]$,

$$
\begin{aligned}
\int_0^{a_{\max}} P^{-1}(a_i)\, da_i \leq \int_0^{a_{\max}} \int_0^{a_{\max}} f(i-j)\, da_j da_i \\
= \frac{1}{2}\int_0^{a_{\max}} f(i-j)\, da_j da_i + \frac{1}{2}\int_0^{a_{\max}} \int_0^{a_{\max}} f(j-i)\, da_j da_i \\
= \frac{1}{2}\int_0^{a_{\max}} \int_0^{a_{\max}} [f(i-j) + f(j-i)]\, da_j da_i \\
= \frac{1}{2}\int_0^{a_{\max}} \int_0^{a_{\max}} da_j da_i = \int_0^{a_{\max}} a da.
\end{aligned}
$$

- We get contradiction with $\int_0^{a_{\max}} \left(y - P^{-1}(y)\right) dy < 0$.

Theorem 3

Proof: Contagion wave on line, RU case

- Integrate over $a_i \in [0, a_{\max}]$,

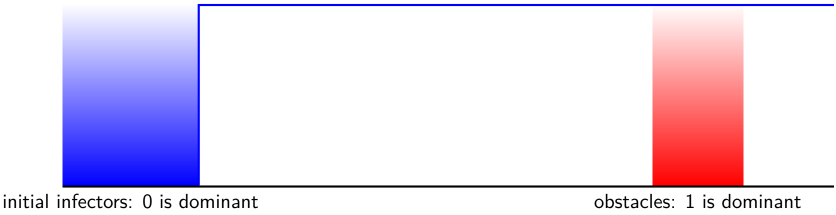
$$\begin{aligned}
\int_0^{a_{\max}} P^{-1}(a_i)\, da_i \leq & \int_0^{a_{\max}} \int_0^{a_{\max}} f(i-j)\, da_j da_i \\
=& \frac{1}{2}\int_0^{a_{\max}} f(i-j)\, da_j da_i + \frac{1}{2}\int_0^{a_{\max}} \int_0^{a_{\max}} f(j-i)\, da_j da_i \\
=& \frac{1}{2}\int_0^{a_{\max}} \int_0^{a_{\max}} \left[f(i-j) + f(j-i)\right] da_j da_i \\
=& \frac{1}{2}\int_0^{a_{\max}} \int_0^{a_{\max}} da_j da_i = \int_0^{a_{\max}} a da.
\end{aligned}$$

- We get contradiction with $\int_0^{a_{\max}} \left(y - P^{-1}(y)\right) dy < 0$.

Theorem 3

Proof: Contagion wave on line, RU case

- Hence the contagion has to spread to the entire line.
- But! - so far we assumed that locations contain continuum.
- Contagion can be also stopped by unusual payoff shocks, like those that make 1 dominant.



Theorem 3

Proof: Contagion wave on line, RU case

- Hence the contagion has to spread to the entire line.
- But! - so far we assumed that locations contain continuum.
- Contagion can be also stopped by unusual payoff shocks, like those that make 1 dominant.

Theorem 3

Proof: Contagion wave on line, RU case

- We can compare the relative likelihood of infectors vs obstacles.
- On line, the latter can be more frequent.
- But not on 2-dimensional lattices.

Theorem 3

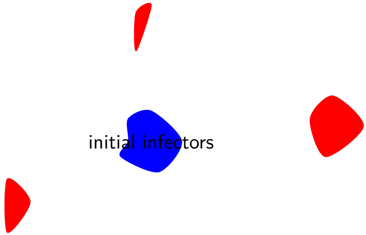
Proof: Contagion wave on line, RU case

- We can compare the relative likelihood of infectors vs obstacles.
- On line, the latter can be more frequent.
- But not on 2-dimensional lattices.

Theorem 3

Proof: Contagion wave on line, RU case

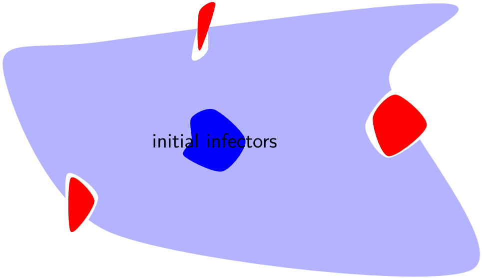
- We can compare the relative likelihood of infectors vs obstacles.
- On line, the latter can be more frequent.
- But not on 2-dimensional lattices.



Theorem 3

Proof: Contagion wave on line, RU case

- We can compare the relative likelihood of infectors vs obstacles.
- On line, the latter can be more frequent.
- But not on 2-dimensional lattices.



Theorem 4

- So far, we showed that there are networks g such that $\mathrm{Eq}\,(g, \varepsilon) \subseteq_\eta \{x^*\}$ with a large probability.
- Next, we show that if $d\,(g) = \max_{i,j} \frac{g_{ij}}{g_i}$ is sufficiently small, than $\{x^*\} \subseteq_n \mathrm{Eq}\,(g, \varepsilon)$.

Theorem 4

Theorem

For any sequence of graphs g_n, if $d(g_n) \to 0$, then

$$\lim_n Prob\left(\{x^*\} \subseteq_\eta Eq(g_n)\right) = 1.$$

Theorem 4

- Hence $\{x^*\}$ is the smallest equilibrium set.
- Equilibrium selection theory: no matter what network, there is an equilibrium with aggregate behavior,
  - the proof tries to make this idea more precise.
- Analog of a result from Morris "Contagion": if all but finitely many agents play risk-dominant action, the best response dynamics won't move towards risk-dominated action.

Theorem 4

Proof: Morris "Contagion"

- Morris: "Contagion":
- Initial profile a^0: all but fintely many play risk-dominant action 0
- Consider a best response dynamics $a^0 < a^1 < a^2 < \ldots$
  - each "round" only one agent changes action
- For each profile a, define *capacity to infect*:

$$\mathcal{F}_0(a) = \sum_{i,j:a_i=1,a_j=0} g_{ij}.$$

Theorem 4

Proof: Morris "Contagion"

- Morris: "Contagion":
- Initial profile a^0: all but fintely many play risk-dominant action 0
- Consider a best response dynamics $a^0 < a^1 < a^2 < \ldots$
  - each "round" only one agent changes action
- For each profile a, define *capacity to infect*:

$$\mathcal{F}_0\left(a\right) = \sum_{i,j:a_i=1,a_j=0} g_{ij}.$$

Theorem 4

Proof: Morris "Contagion"

- Capacity must go down at each round:
  - if i changes action from 0 to 1 as a best response, the capacity changes by by
  $$\sum_{j:a_j=0} g_{ij} - \sum_{j:a_j=1} g_{ij}.$$
  - but, if 1 is a best response and 0 is risk-dominant, then
  $$\sum_{j:a_j=0} g_{ij} < \frac{1}{2} < \sum_{j:a_j=1} g_{ij}.$$
  - So, the capacity goes down every single infection!
- Because the capacity cannot be negative, contagion has to stop.
- If the initial profile was close to 0, the capacity was small and the contagion will stop very soon, with most agents not changing their actions.

Theorem 4

Proof: Morris "Contagion"

- Capacity must go down at each round:
  - if i changes action from 0 to 1 as a best response, the capacity changes by by
$$\sum_{j:a_j=0} g_{ij} - \sum_{j:a_j=1} g_{ij}.$$
  - but, if 1 is a best response and 0 is risk-dominant, then
$$\sum_{j:a_j=0} g_{ij} < \frac{1}{2} < \sum_{j:a_j=1} g_{ij}.$$
  - So, the capacity goes down every single infection!
- Because the capacity cannot be negative, contagion has to stop.
- If the initial profile was close to 0, the capacity was small and the contagion will stop very soon, with most agents not changing their actions.

Theorem 4

Proof: Morris "Contagion"

- Capacity must go down at each round:
  - if i changes action from 0 to 1 as a best response, the capacity changes by by
$$\sum_{j:a_j=0} g_{ij} - \sum_{j:a_j=1} g_{ij}.$$
  - but, if 1 is a best response and 0 is risk-dominant, then
$$\sum_{j:a_j=0} g_{ij} < \frac{1}{2} < \sum_{j:a_j=1} g_{ij}.$$
  - So, the capacity goes down every single infection!
- Because the capacity cannot be negative, contagion has to stop.
- If the initial profile was close to 0, the capacity was small and the contagion will stop very soon, with most agents not changing their actions.

Theorem 4

Proof: Morris "Contagion"

- Capacity must go down at each round:
  - if i changes action from 0 to 1 as a best response, the capacity changes by by
    $$\sum_{j:a_j=0} g_{ij} - \sum_{j:a_j=1} g_{ij}.$$
  - but, if 1 is a best response and 0 is risk-dominant, then
    $$\sum_{j:a_j=0} g_{ij} < \frac{1}{2} < \sum_{j:a_j=1} g_{ij}.$$
  - So, the capacity goes down every single infection!
- Because the capacity cannot be negative, contagion has to stop.
- If the initial profile was close to 0, the capacity was small and the contagion will stop very soon, with most agents not changing their actions.

Theorem 4

Proof: Morris "Contagion"

- Capacity must go down at each round:
  - if i changes action from 0 to 1 as a best response, the capacity changes by by
  $$\sum_{j:a_j=0} g_{ij} - \sum_{j:a_j=1} g_{ij}.$$
  - but, if 1 is a best response and 0 is risk-dominant, then
  $$\sum_{j:a_j=0} g_{ij} < \frac{1}{2} < \sum_{j:a_j=1} g_{ij}.$$
  - So, the capacity goes down every single infection!
- Because the capacity cannot be negative, contagion has to stop.
- If the initial profile was close to 0, the capacity was small and the contagion will stop very soon, with most agents not changing their actions.

Theorem 4

Proof: Morris “Contagion”

- Key feature of a good definition of capacity
  - it decreases along best response dynamics,
  - it is small,
  - cannot be negative.
- The number of stages until the dynamics stops is related to the initial capacity.

Theorem 4

Proof: RU case

- Our proof follows a similar idea.
- Let x^*be *RU*-dominant outcome.
- Construct *initial profile* a^0 st. for each i,

$$a_i^0 \in \arg\max_a u_i\left(a, x^*, \varepsilon_i\right)$$

  - many people play 0 and many play 1
- Consider best response dynamics $a^0 < a^1 < \ldots. < a^T$.
- We show that $\frac{T}{N} \sim O\left(d\left(g\right)\right)$.
- Hence $a_i^0 \in \arg\max_a u_i\left(a, x^*, \varepsilon_i\right)$ is a pretty safe action to take, whatever is the true network.

Theorem 4

Proof: RU case

- Our proof follows a similar idea.
- Let x^*be *RU*-dominant outcome.
- Construct *initial profile* a^0 st. for each i,

$$a_i^0 \in \arg\max_a u_i\left(a, x^*, \varepsilon_i\right)$$

  - many people play 0 and many play 1
- Consider best response dynamics $a^0 < a^1 < .... < a^T$.
- We show that $\frac{T}{N} \sim O\left(d\left(g\right)\right)$.
- Hence $a_i^0 \in \arg\max_a u_i\left(a, x^*, \varepsilon_i\right)$ is a pretty safe action to take, whatever is the true network.

Theorem 4

Proof: RU case

- Our proof follows a similar idea.
- Let x^*be RU-dominant outcome.
- Construct *initial profile* a^0 st. for each i,

$$a_i^0 \in \arg\max_a u_i\left(a, x^*, \varepsilon_i\right)$$

  - many people play 0 and many play 1
- Consider best response dynamics $a^0 < a^1 < .... < a^T$.
- We show that $\frac{T}{N} \sim O\left(d\left(g\right)\right)$.
- Hence $a_i^0 \in \arg\max_a u_i\left(a, x^*, \varepsilon_i\right)$ is a pretty safe action to take, whatever is the true network.

Theorem 4

Proof: RU case

- Definition of capacity

$$\sum_{i,j:a_i=1,a_j=0} g_{ij}.$$

- The problem is that is is too big for the initial profile a^0

Theorem 4

Proof: RU case

- Definition of capacity

$$\sum_{i,j:a_i=1,a_j=0} g_{ij}.$$

- The problem is that is is too big for the initial profile a^0

Theorem 4
Proof: RU case

- Definition of capacity:

$$\sum_{i,j:a_i=1,a_j=0} g_{ij} = \frac{1}{2}\sum_{i,j} g_{ij}\left(a_i - a_j\right)^2 .$$

Theorem 4

Proof: RU case

- Definition of capacity:

$$\frac{1}{2}\sum_{i,j} g_{ij}\left(a_i - a_j\right)^2 .$$

Theorem 4

Proof: RU case

- Definition of capacity:

$$\mathcal{F}(a) = \frac{1}{2} \sum_{i,j} g_{ij} \left(P\left(\beta_i^a\right) - P\left(\beta_j^a\right) \right)^2 .$$

- replace a_i by the "continuum best response" to the neighborhood profile β^a.

Theorem 4

Proof: RU case

- Definition of capacity:

$$\mathcal{F}(a) = \frac{1}{2}\sum_{i,j} g_{ij}\left(P\left(\beta_i^a\right) - P\left(\beta_j^a\right)\right)^2 .$$

  - replace a_i by the “continuum best response” to the neighborhood profile β^a.

Theorem 4

Proof: RU case

- Definition of capacity:

$$\mathcal{F}(a) = \frac{1}{2}\sum_{i,j} g_{ij}\left(P\left(\beta_i^a\right) - P\left(\beta_j^a\right)\right)^2.$$

  - because $x^* = P(x^*)$ and $d(g) \sim 0$,
  - $\beta_i^a \sim x^*$ for most i,
  - $P\left(\beta_i^a\right) \sim P\left(\beta_j^a\right)$ for most i and j,
- capacity is small.

Theorem 4

Proof: RU case

- RU case.
- Assume that *RU*-dominant outcome $x^* = P(x^*) = 0$.
- Definition of capacity: Instead of

$$\frac{1}{2}\sum_{i,j} g_{ij}\left(a_i - a_j\right)^2,$$

we take

$$\mathcal{F}(a) = \frac{1}{2}\sum_{i,j} g_{ij}\left(P\left(\beta_i^a\right) - P\left(\beta_j^a\right)\right)^2.$$

  - recall that β_i^a is the neighborhood fraction, and
  - $P\left(\beta_i^a\right)$ is the "expected" best response of agent i.

Theorem 4

Proof: RU case

- Turns out that this is a good definition
  - capacity is small at a^0 as with a large probability $\beta_i^a \sim \beta_j^a$,
  - and it is a sum of a martingale and a decreasing process. Ignoring (probabilistically) small terms, we get, for each T

$$\mathcal{F}\left(P\left(\beta^0\right)\right) \geq 2\sum_i g_i \left[\int_{x^*}^{P\left(\beta_i^T\right)} \left(P^{-1}\left(y\right) - y\right) dy\right].$$

Conclusion

- Heterogeneous payoffs in coordination games on network.
- We characterized the largest and the smallest possible set of equilibrium average behaviors across all networks.
- Results:
  - The largest set achieved on a collection of complete graphs,
  - partial identification theory,
  - The smallest set achieved on 2-dimensional (but not necessarily 1-dimensional) lattice,
  - equilibrium selection theory.
- Main assumptions:
  - independent payoff shocks,
  - large degree,
  - both assumptions are important.