Learning and Games, day 4

Price of Anarchy and Game Dynamics

Éva Tardos, Cornell

Learning and Games

Price of Anarchy and Game Dynamics

Day 4: Can learning do better than worst Nash?

Main question:
Quality of Selfish outcome

Selfish outcome = result of Learning behavior

Our Question: quality of learning outcomes?

which correlated equilibrium do users coordinate on?

Answer: depends on which learning…

Theorem: $\forall$ correlated equilibrium is the limit point of no-regret play

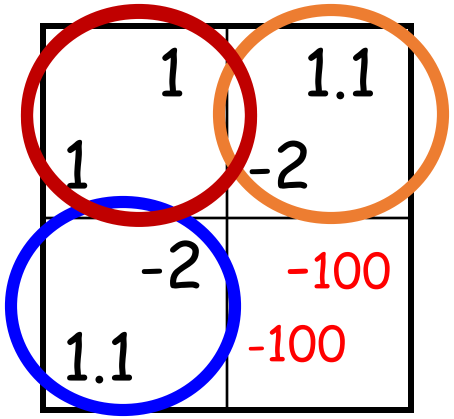
Correlated eq. = learning outcome?

Proof: Intelligent designer algorithm

Take a coarse correlated equilibrium

assume probabilities p rational

Design a sequence of moves that has desired distribution (½;¼,¼,0)

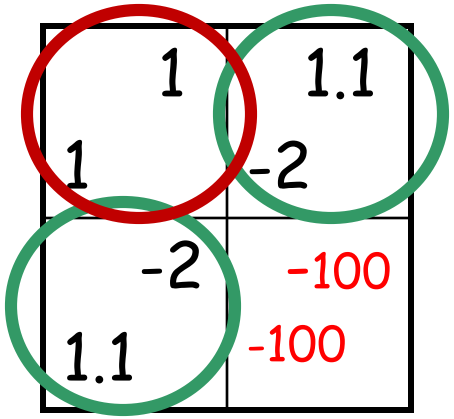


Sequence

(1,1), (2,1), (1,1), (1,2)

Repeat!

Correlated eq. = learning outcome?



Sequence
(1,1), (2,1), (1,1), (1,2)
Repeat!

Intelligent designer algorithm

- Follow the designed sequence as long as all other players do.
- If anyone deviates: switch to smoothed fictitious play

This is no regret!

Quality of Learning Outcome

Price of Anarchy [Koutsoupias-Papadimitriou'99]

$$PoA = \max_{a\ Nash} \frac{cost(a)}{Opt}$$

Assuming **no-regret learners** in fixed game: [Blum, Hajiaghayi, Ligett, Roth'08, Roughgarden'09]

$$PoA = \lim_{T\to\infty} \frac{\sum_{t=1}^{T} cost(a^t)}{T\ Opt}$$

[Lykouris, Syrgkanis, T. 2016] dynamic population

$$PoA = \lim_{T\to\infty} \frac{\sum_{t=1}^{T} cost(a^t, v^t)}{\sum_{t=1}^{T} Opt(v^t)}$$

where v^t is the vector of player types at time t

A Game with Bad PoA

Personal objective: minimize

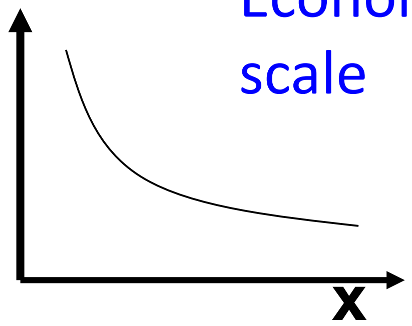
c_P(f) = sum of costs of edges along **P** (wrt. flow **f**)

Overall objective:

C(f) = total cost of a flow **f**: = $\Sigma_e f_e \bullet c_P(f_e)$

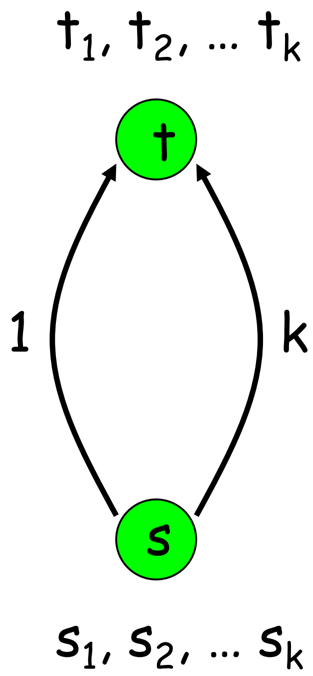
= - social welfare

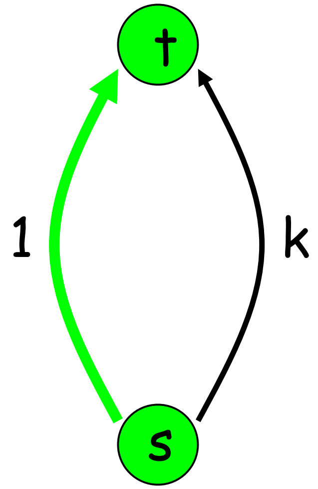
or total/average cost



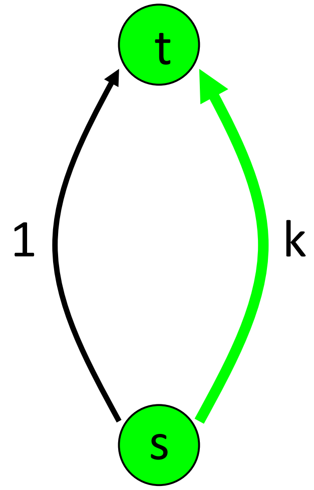
Economy of scale

Cost-sharing: a bad example: $c_e(x) = c_e/x$





One NE:
each player
pays 1/k



Another NE:
each player
pays 1

Claim: this is the worst case

Maybe Best Nash is good?

We know price of anarchy is bad, but Price of Stability is better.

$$\text{Price of Stability} = \frac{\text{cost of best selfish outcome}}{\text{“socially optimum” cost}}$$

Theorem [Anshelevich, Dasgupta, Kleinberg, Tardos, Wexler, Roughgarden FOCS'04]
Price of Stability is at most H_k =O(log k) for k players, while price of anarchy is at most k

Selfish Outcome= non-cooperative?

Nash equilibrium: non-cooperative outcome

- Current strategy “best response” for all players
- no single user has incentive to deviate

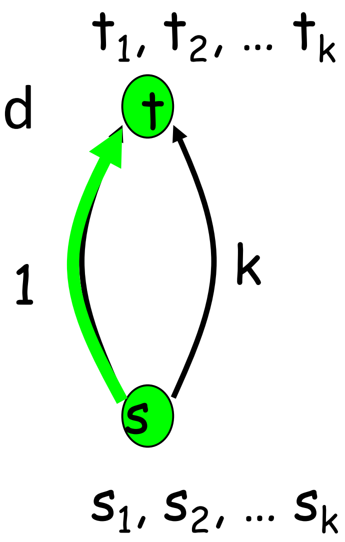
How about groups of players?

Strong Nash equilibrium: no group of players has incentive to deviate [Aumann’59]

Cooperative game?

We can use: Strong Nash equilibrium

- No subset players can coordinate a deviation and improve for every player in the set



[Epstein, Feldman, Mansour EC'07]

the strong price of anarchy is H_k =O(log k)

(but strong Nash may not exists…)

Open problems

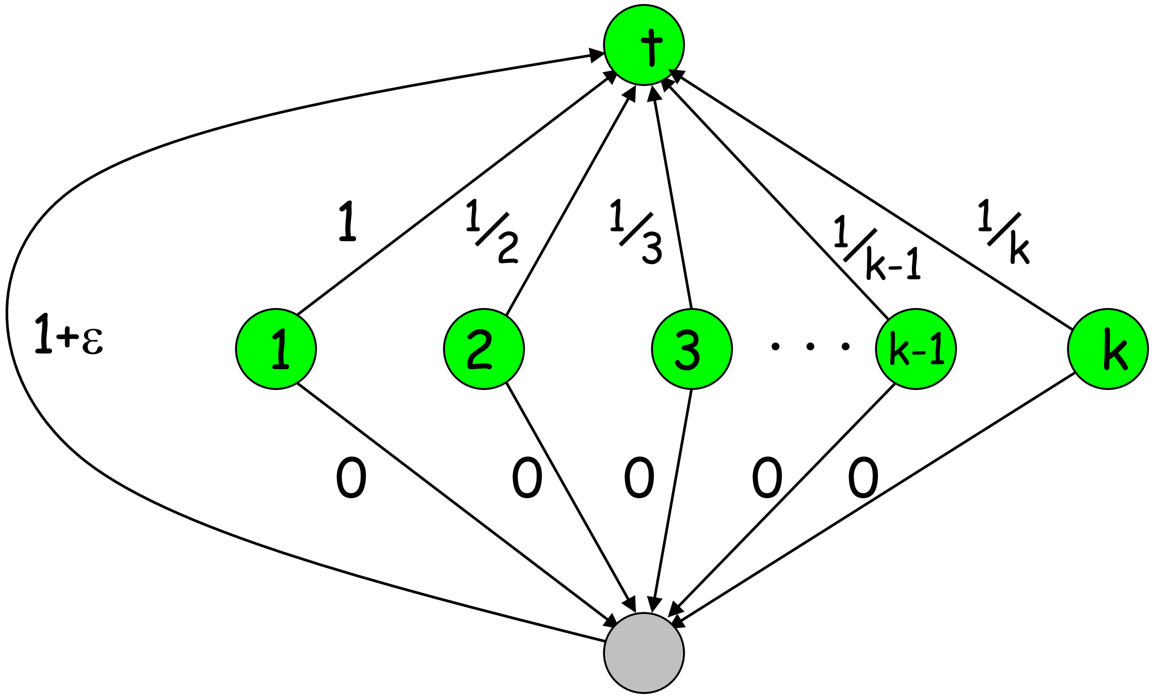
Is there a simple dynamic that leads to such better outcomes?

- Learning or best response (random best response?) from a random start? Or from users arriving one-by-one?

- What is a cooperative dynamic?

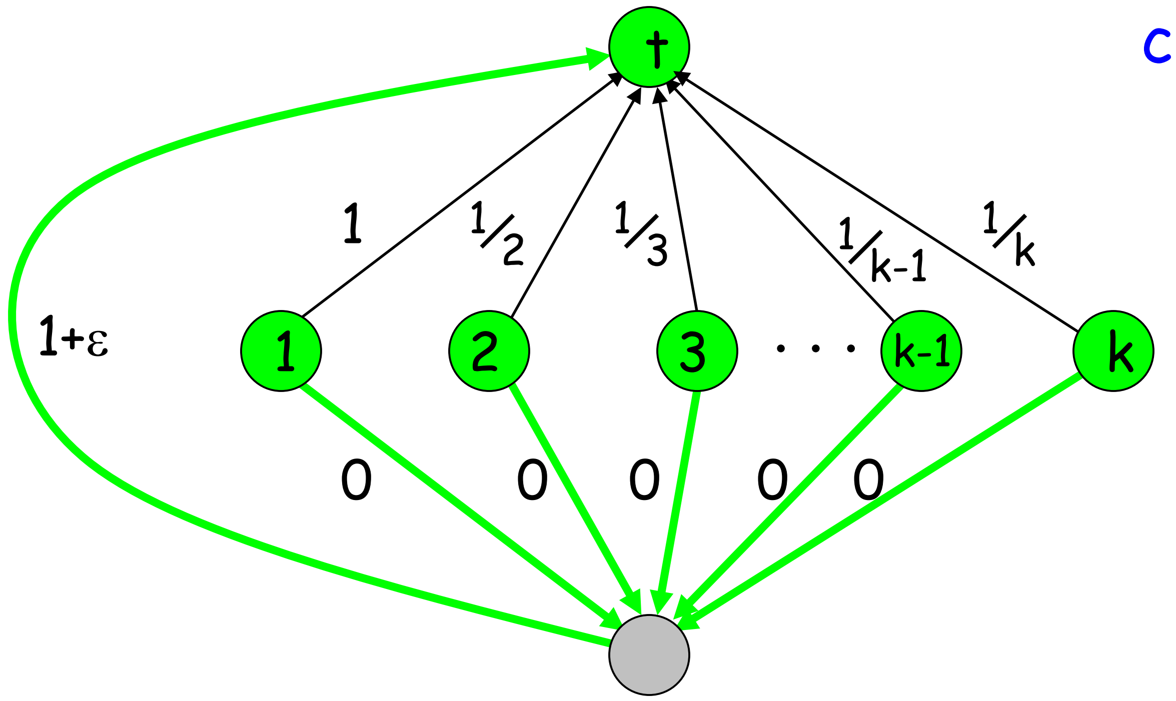
Illustrative Example:

All noes 1… k want to connect to terminal t



Example:

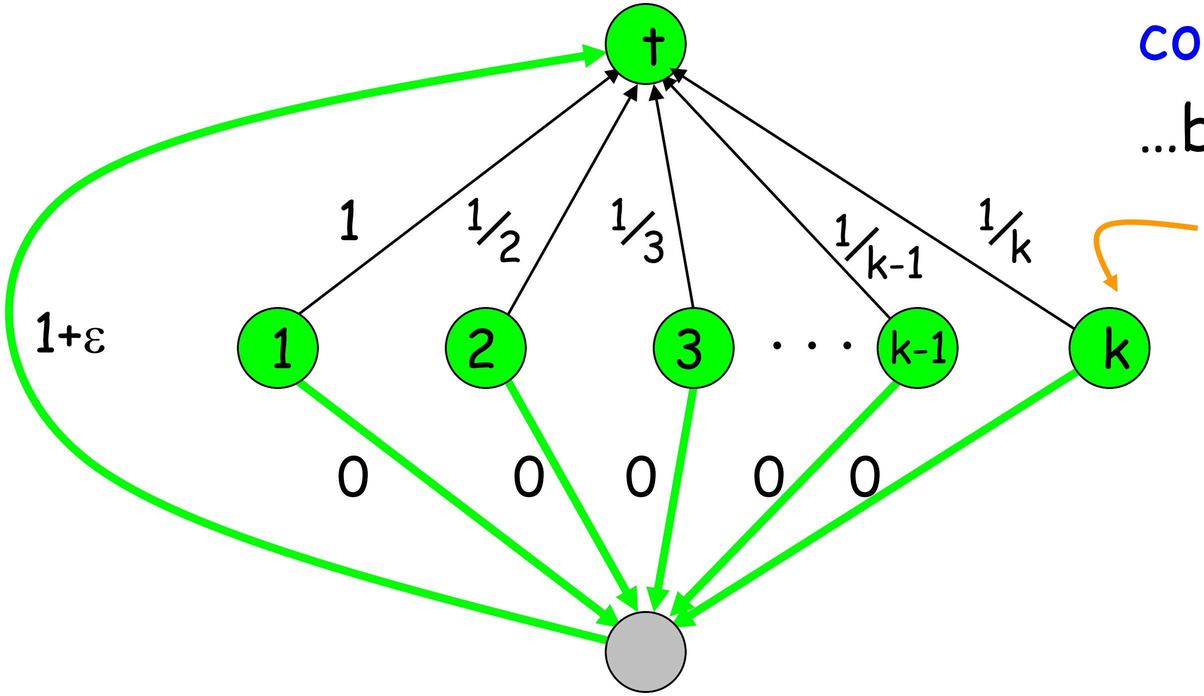
All noes 1… k want to connect to terminal t



cost(OPT) = 1+ε

Example:

All noes 1… k want to connect to terminal t



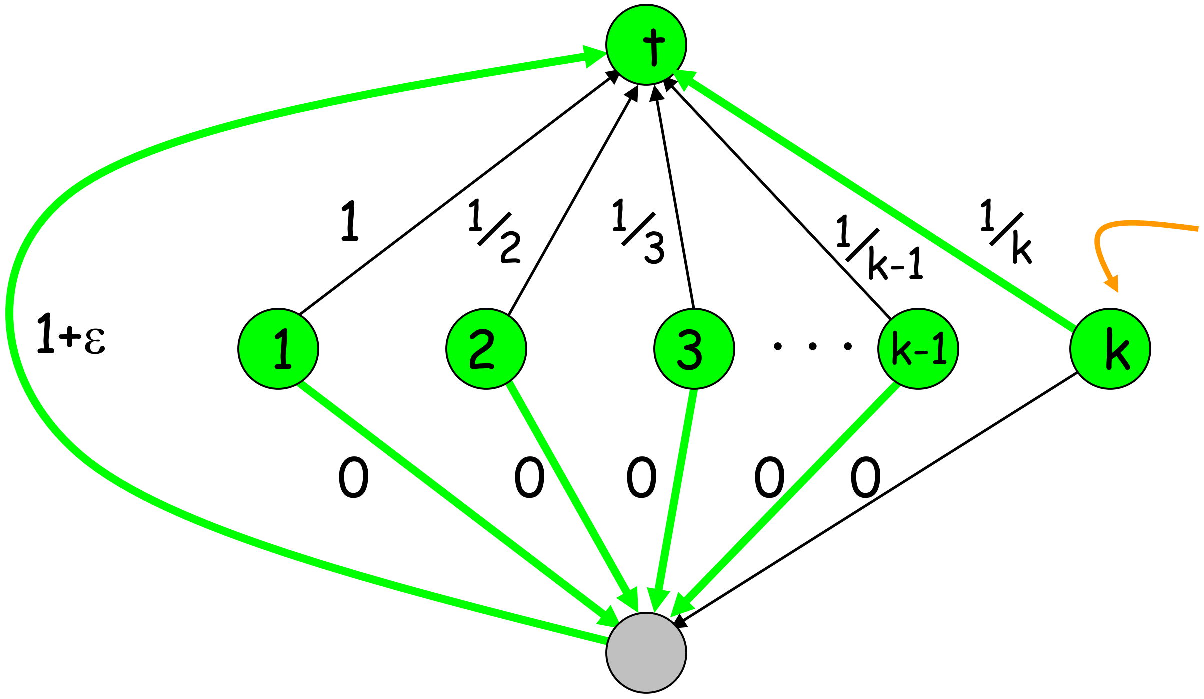
cost(OPT) = 1+ε

…but not a NE:

player k pays (1+ε)/k, could pay 1/k

Example:

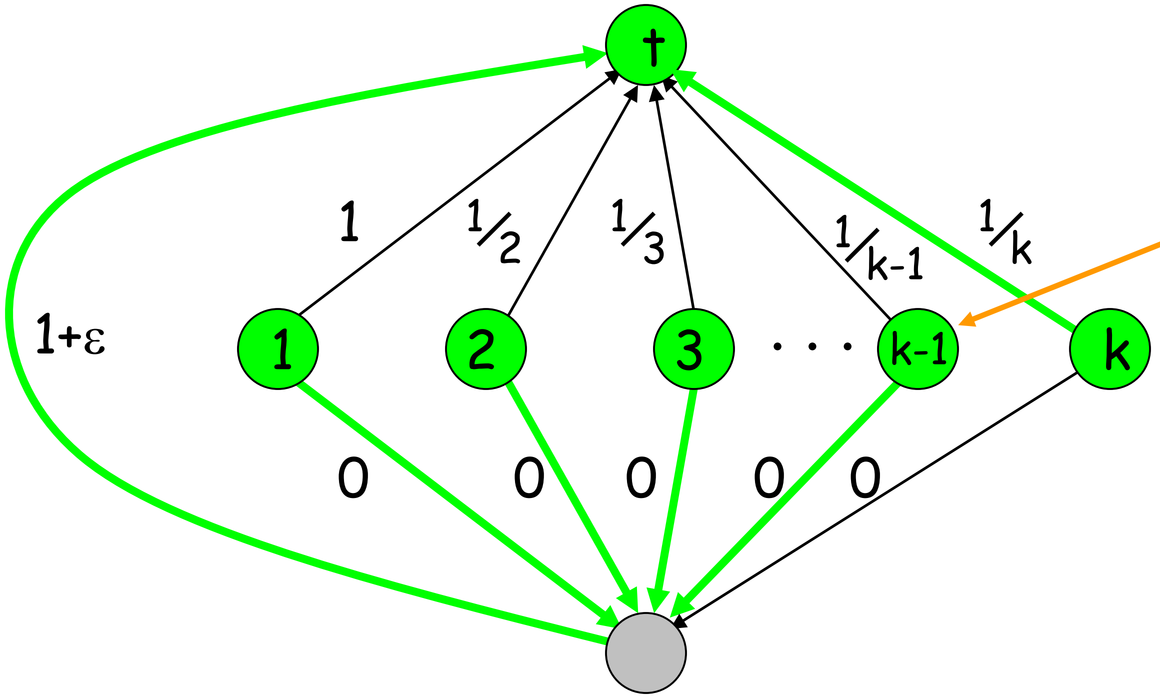
All noes 1… k want to connect to terminal t



so player k would deviate

Example:

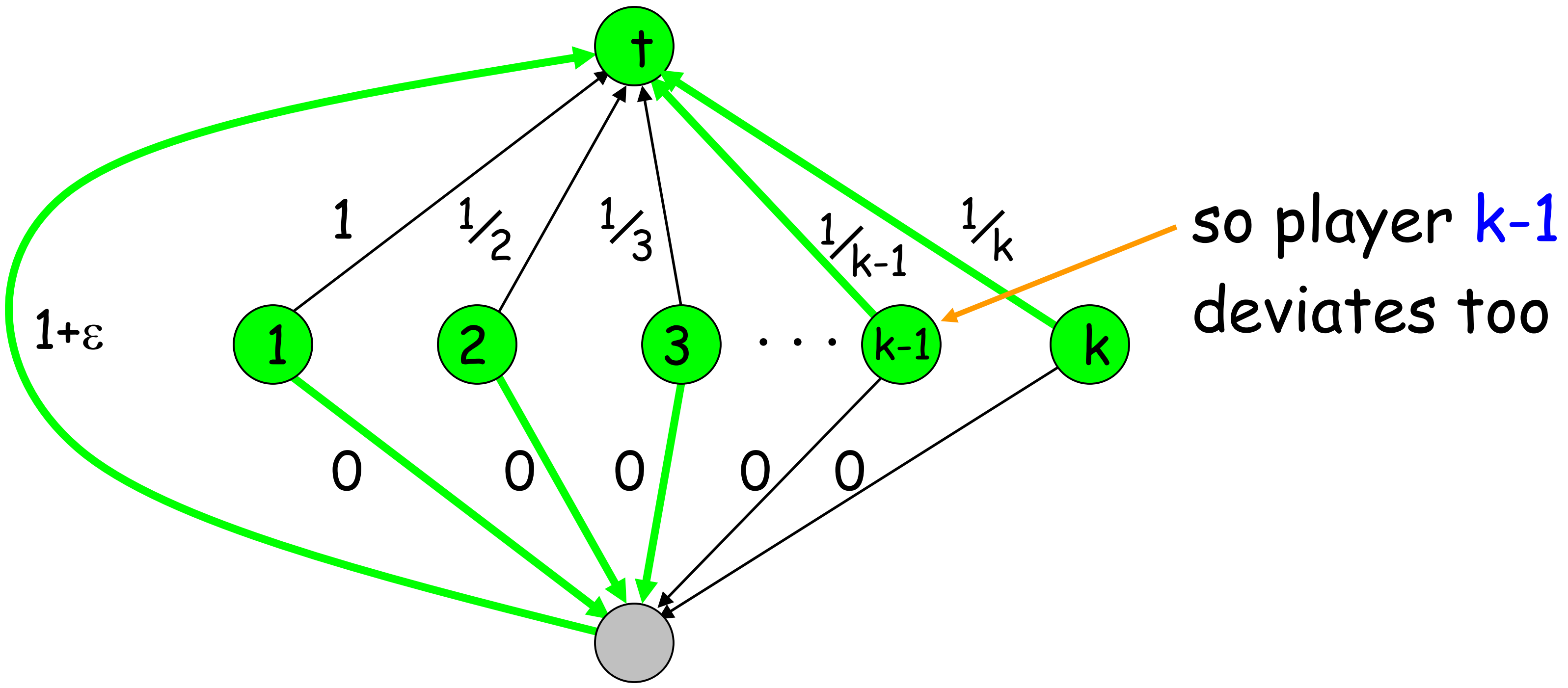
All noes 1… k want to connect to terminal t



now player k-1 pays (1+ε)/(k-1), could pay 1/(k-1)

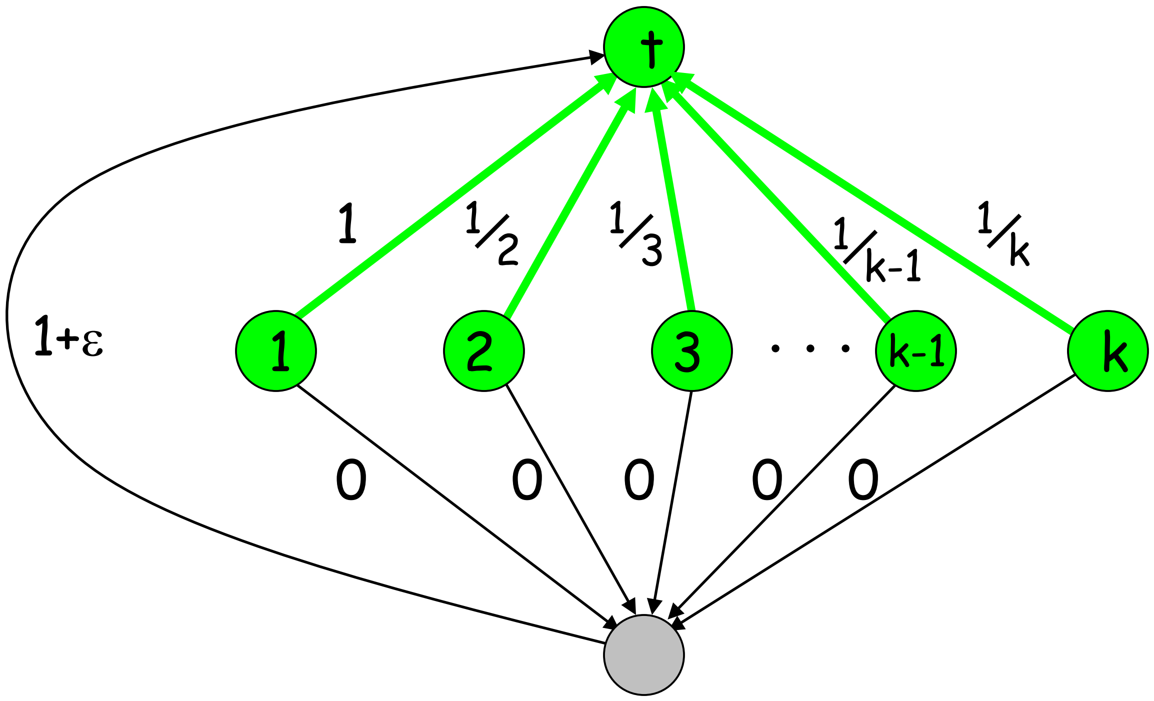
Example:

All noes 1… k want to connect to terminal t



Example:

All noes 1… k want to connect to terminal t



Continuing this process, all players defect.

This is a NE!

(the only Nash)

cost = $1 + \frac{1}{2} + \ldots + \frac{1}{k}$

In fact, a strong Nash

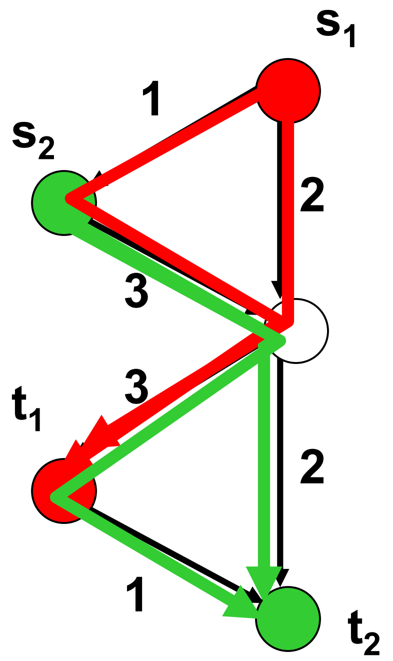
Price of Stability is $H_k = \Theta(\log k)$!

But: strong Nash ∃?

[Epstein, Feldman, Mansour EC’07]
the strong price of anarchy is O(log k)

But ∃?: No!!

Nash unique: cost of 5 each



It is not strong! As there is a solution better for both

cost of 4 each

It’s a “prisoner dilemma”

	C	D
C	4 / 4	3.5 / 5.5
D	5.5 / 3.5	5 / 5

⇒ no strong Nash exists ∃

Proof idea: congestion games have potentials

Φ(f) = Σ_e ($\mathbf{c}_e$(1)+…+ $\mathbf{c}_e$(f_e)) = $\Sigma_e \Phi_e$

[in non-atomic game $\Phi = \sum_e \int_0^{f_e} c_e(\xi) d\xi$]

Theorem (Rosenthal) if player i moves from path P to a new path Q

Improving her cost by Δ, then potential decreases by Δ

Proof: if player i, was using path P and now leaves the game, **Φ(f)** decreases by $\sum_{e \in P} c_e(f_e)$, which is player i's cost.

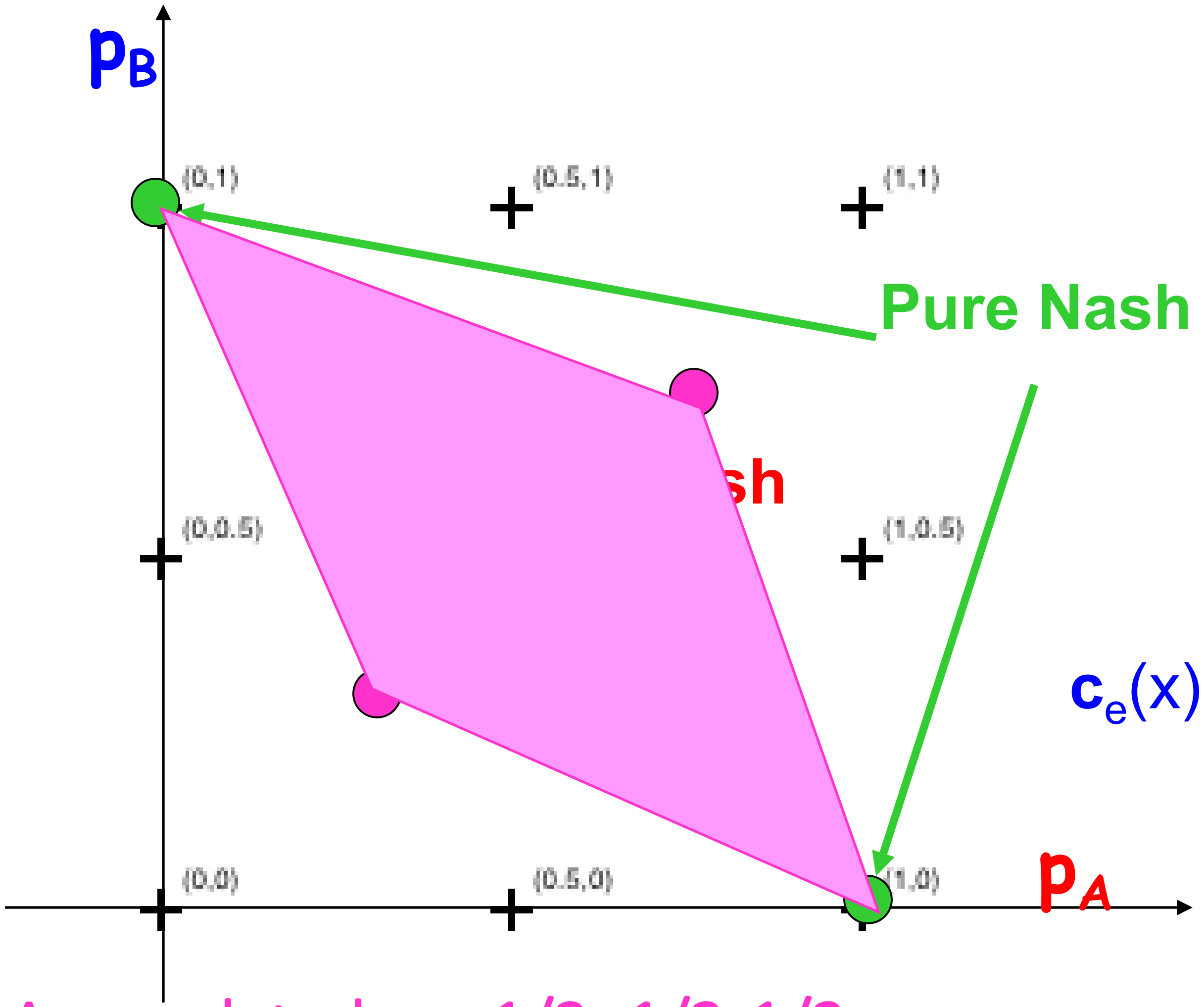
Now she re-enters on path Q . Use the same argument.

Congestion games are potential games

This implies a few useful things

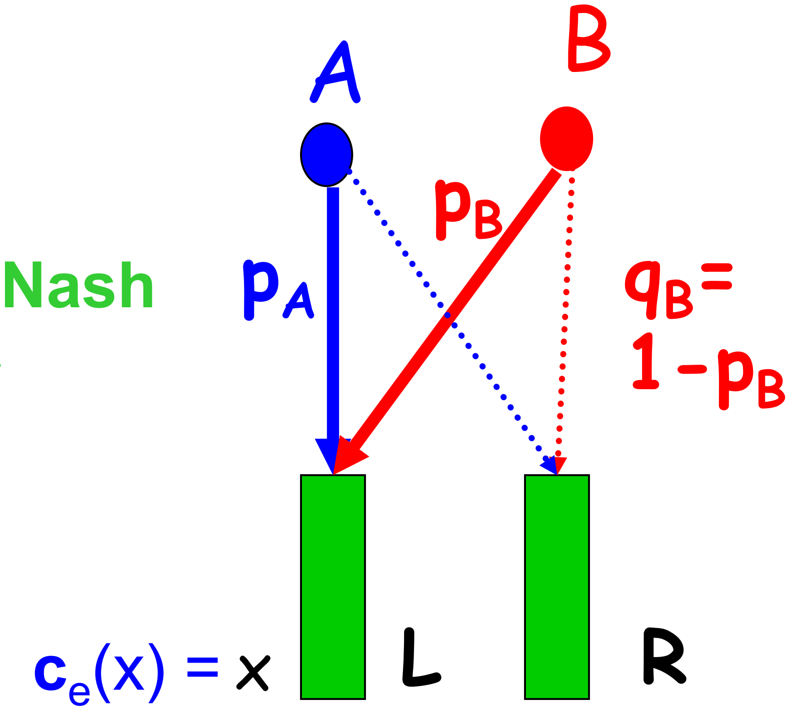
- Nash = local minima of potential Φ
- Repeated best response leads to a Nash equilibrium: decreases potential Φ
- Learning also leads to Nash equilibria (not to correlated equilibria!)

Nash for 2-ball & 2-bin



A correlated eq. 1/3–1/3-1/3

The set of correlated equilibria



2, 2	1, 1
1, 1	2, 2

Using potential Φ...

- Consider the Nash with minimum value of Φ
- This Nash has,

$$\Phi(\text{Nash}) < \Phi(\text{OPT}).$$

Suppose that we also know for any solution

$$\Phi \leq \text{cost} \leq \mathbf{A}\,\Phi$$

$\rightarrow$ $\text{cost}(\text{Nash}) \leq \mathbf{A}\,\Phi(\text{Nash}) \leq \mathbf{A}\,\Phi(\text{OPT}) \leq \mathbf{A}\,\text{cost}(\text{OPT})$.

$\rightarrow$ There is a good Nash!

Results for Cost sharing

proof:

Recall: $\Phi(f) = \Sigma_e (c_e(1)+\ldots+ c_e(f_e)) = \Sigma_e \Phi_e$

$f_e \leq k$ users on edge e then

- true cost is c_e with any >0 users
- Potential is $\Phi_e = c_e + c_e/2 + c_e/3+\ldots+c_e/f_e$

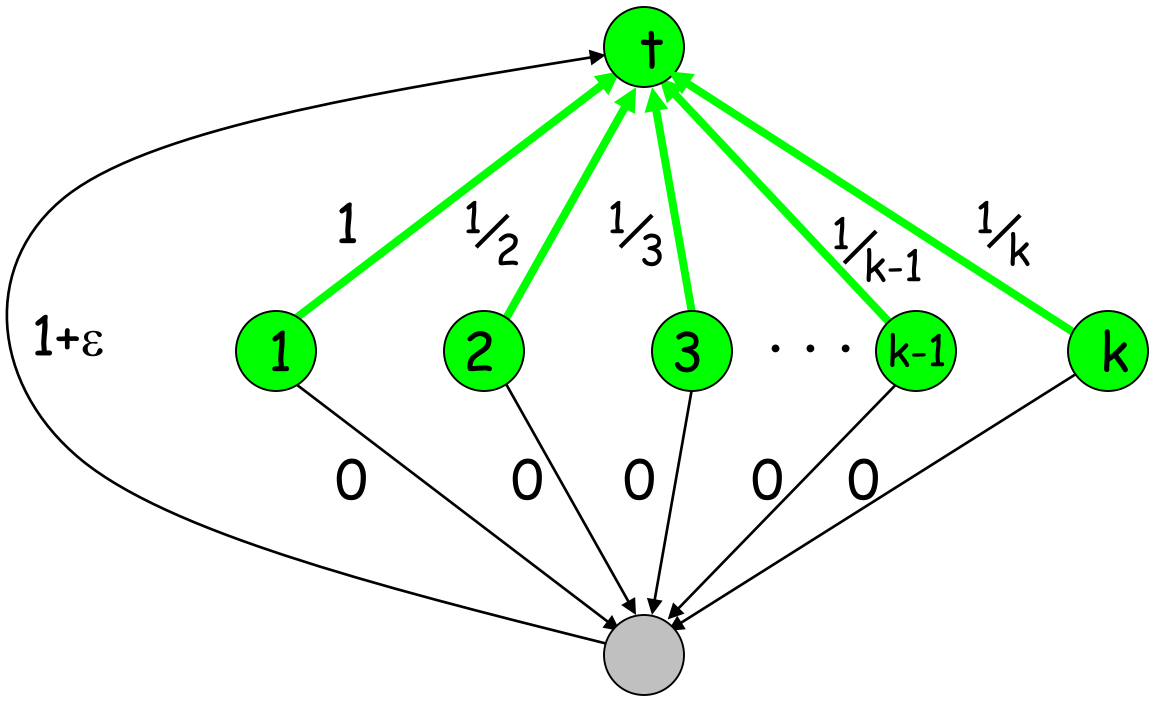
  $\leq c_e \cdot (1+1/2+1/3+\ldots+1/k) = c_e H_k$

- cost $\leq \Phi \leq$ cost$\cdot H_k$

- → Nash optimizing Φ cost at most H_k above the optimum

Example

All noes 1… k want to connect to terminal t



Continuing this process, all players defect.

This is a NE!

(the only Nash)

cost = $1 + \frac{1}{2} + \ldots + \frac{1}{k}$

In fact, a strong Nash

Price of Stability is $H_k = \Theta(\log k)$!

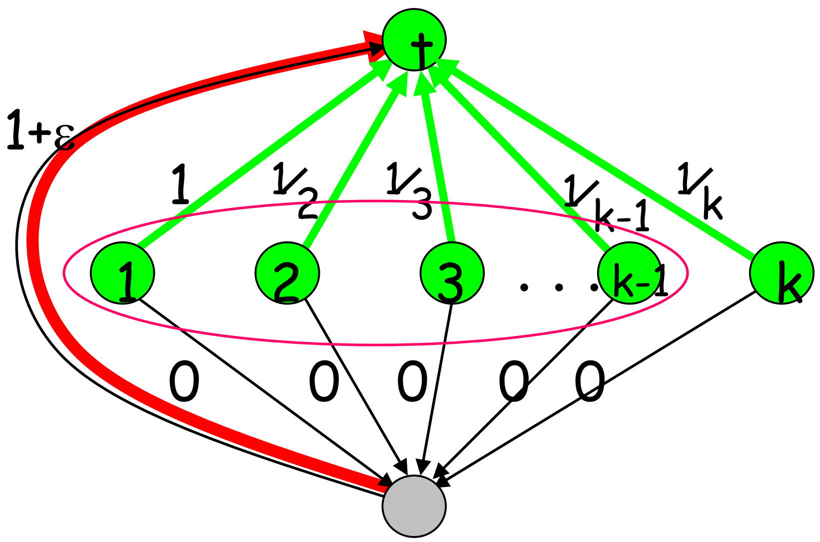
Strong Price of Anarchy?

SE = strong Nash, **Opt**

As a group not all players want to move to Opt:

$\Rightarrow$There exists player, say last player k, that is better off in current solution

$\Rightarrow$ $cost_k(SE) \leq cost_k(Opt)$



Consider remaining k-1 players.

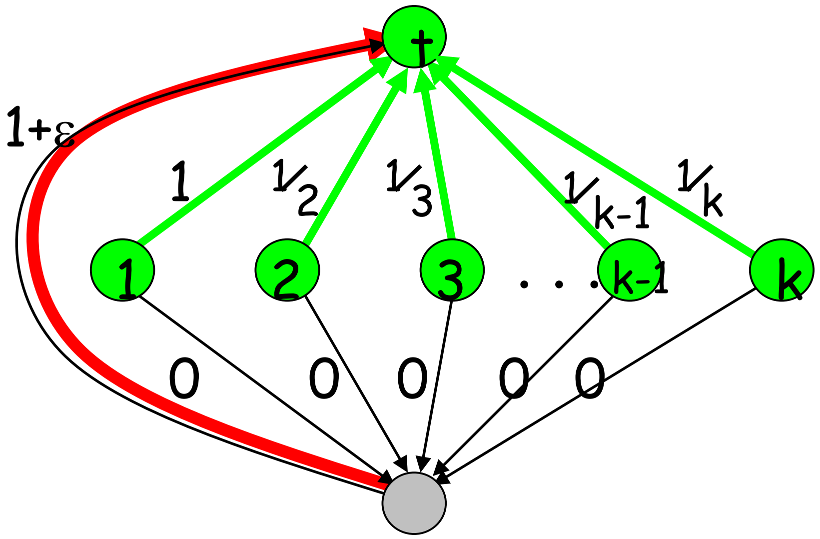
Opt_{k-1} = Opt restricted to remaining k-1 players

As a group the remaining k-1 players also don't want to move to Opt_{k-1} $\Rightarrow$ there is a player, say k-1

$Cost_{k-1}(SE) \leq cost_{k-1}(Opt_{k-1})$

Strong Price of Anarchy

SE = strong Nash, **Opt**,



Continue…

Opt_i = Opt restricted to remaining i

We get: $cost_i(SE) \leq cost_i(Opt_i)$

Lemma: In potential games: $cost_i(Opt_i) = \Phi(Opt_i) - \Phi(Opt_{i-1})$

Proof: consider first i players only, and selfish move of player i of "not playing":

- Cost to player i: $cost_i(Opt_i)$
- potential change $\Phi(Opt_i) - \Phi(Opt_{i-1})$

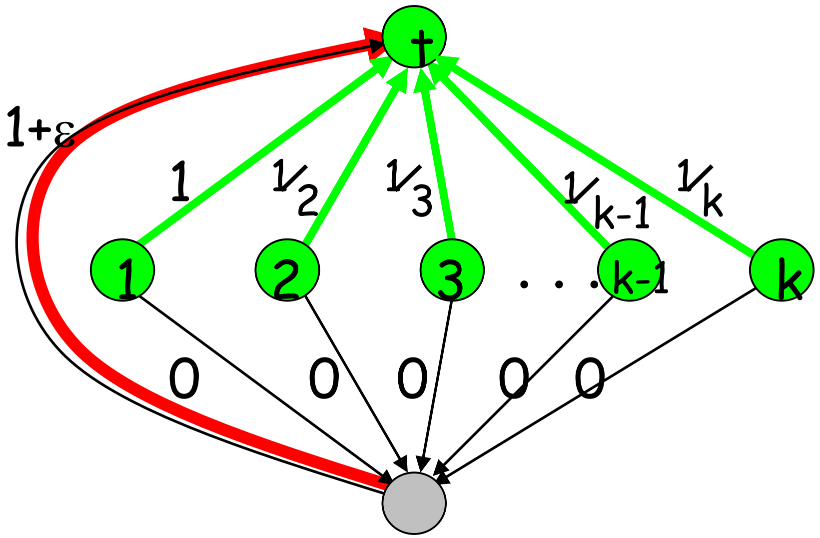
Strong Price of Anarchy

SE = strong Nash, **Opt**,

Opt_i = Opt restricted to first i

set 1…i doesn't want to move

$cost_i(SE) \leq cost_i(Opt_i)$



Potential game: $cost_i(Opt_i) = \Phi(Opt_i) - \Phi(Opt_{i-1})$

We get: $cost_i(SE) \leq cost_i(Opt_i) = \Phi(Opt_i) - \Phi(Opt_{i-1})$

$\sum_i cost_i(SE) \leq \sum_i \Phi(Opt_i) - \Phi(Opt_{i-1}) = \Phi(Opt)$

In cost-sharing game $\Phi(Opt) \leq H_k\, cost(Opt)$

Dynamic with cooperation?

Cooperation: group of users deviate together to improve their welfare

Cooperative game theory…

- No great model for outcome for most games
- Strong Nash: outcome when collusion is not useful.
- But what happens when no such outcome exists: collusion is useful?
- Bargaining: agreement when everyone colludes
  - different bargaining "games" characterized by axioms

Does learning lead to better Nash?

- Idea 1: with a uniformly random start?

- Idea 2: with each player arriving one-by-one, while others are repeatedly best responding
- Charikar, Karloff, C. Mathieu, J. Naor, Saks, SPAA'08: $O(\log^3 n)$ PoA if
  - single source
  - all players arrive before any best response
- idea: arrival phase is $\approx$ online Steiner tree, then use potential function

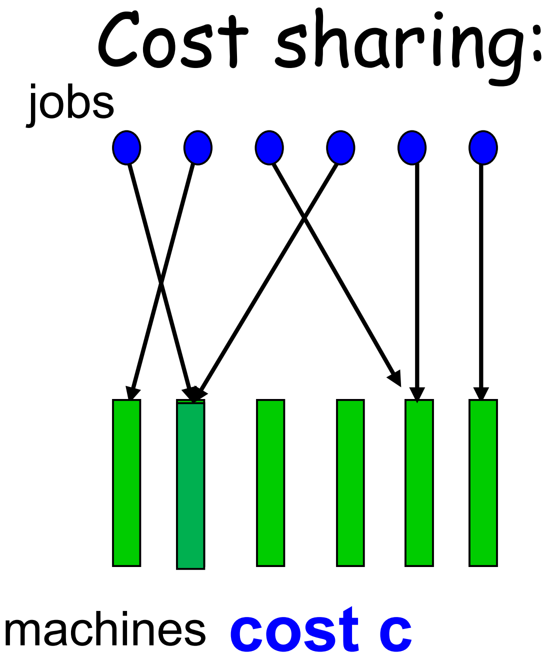
Outcome of Multiplicative Weights

Theorem: R. Kleinberg-Piliouras-Tardos multiplicative weight like processes with small ϵ converge to pure Nash in almost all congestion games

Recall

In congestion games learning converges a Nash (decreases potential).

Which one? Uniform random has cost $\left(1-\frac{1}{e}\right)nc \approx 0.63\, nc$, while Opt=c



Continuous limit of multiplicative weights

Multiplicative weight with ε ~ 0 :

- probability of playing action x is $p_x^t \leftarrow w_x^t / \sum_{s_i} w_{s_i}^t$
- Update $w_x^{t+1} \leftarrow w_x^t \alpha^{c_i(x, s_{-i}^t)}$

Limit as update gets smaller.

$$p_x^{t+1} = \frac{p_x^t \, \alpha^{c_i(x, s_{-i}^t)}}{\sum_y p_y^t \alpha^{c_i(y, s_{-i}^t)}}$$

Limit as update get slower $\alpha = 1 - \epsilon$

- $\lim_{\epsilon \to 0} \frac{p_x^{t+1} - p_x^t}{\epsilon} = \mathrm{p}_\mathrm{x}^\mathrm{t} \left(\sum_y p_y^t c_i(y, s_{-i}^t) - c_i(x, s_{-i}^t)\right)$

Limit = $\frac{dp_x^{t+1}}{d\epsilon}\Big|_{\epsilon=0}$

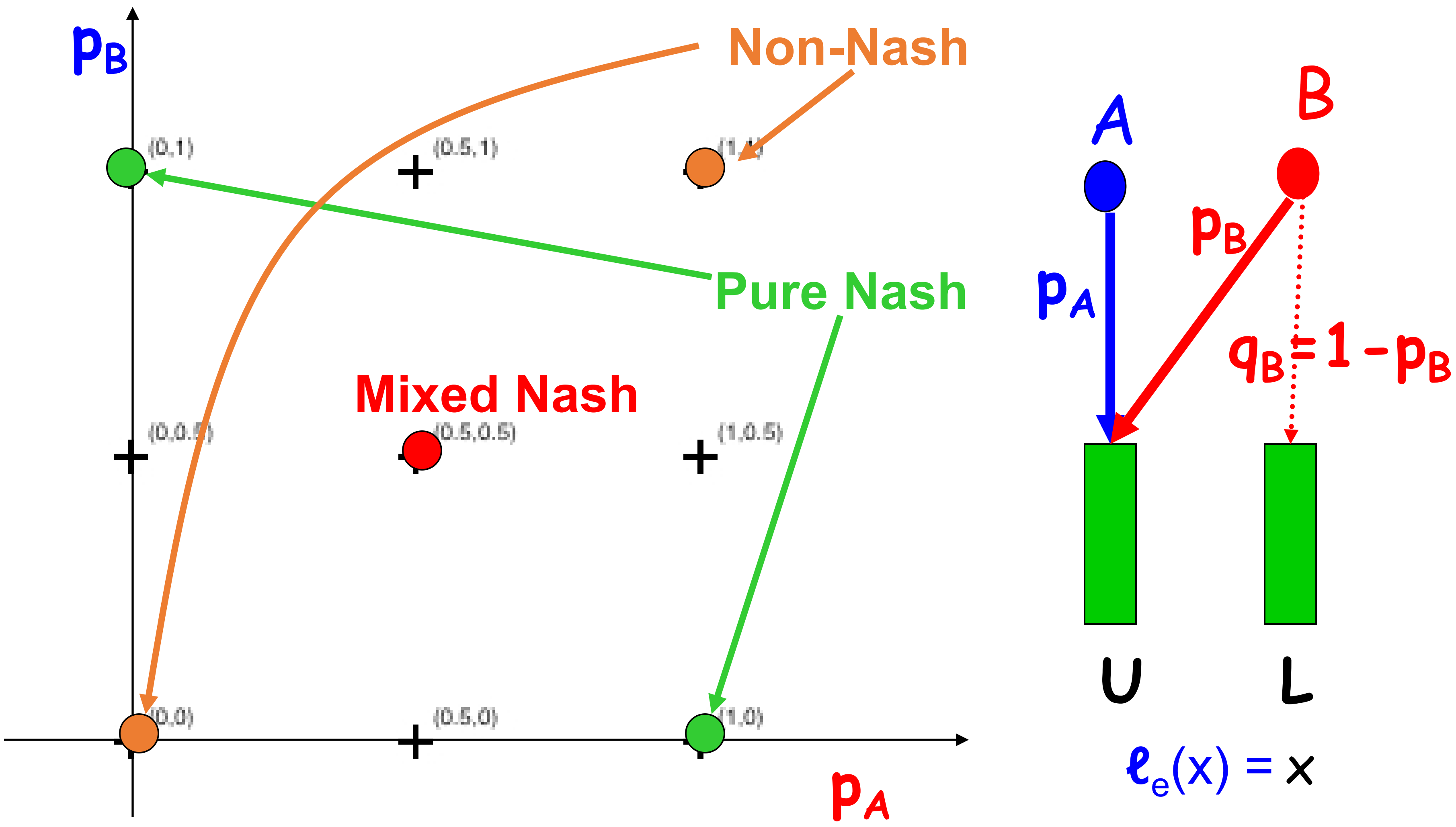
Limit =replicator dynamic: dynamical system

$$\dot{p_x} = \mathrm{p}_\mathrm{x} \left(\sum_y p_y c_i(y, s_{-i}) - c_i(x, s_{-i})\right)$$

What are weakly stable points

Weakly stable for ODE= neighborhood no direction has pull away from fixed point.

Stable points in 2-ball 2-bin



What are weakly stable points

Weakly stable for ODE= neighborhood no direction has pull away from fixed point.

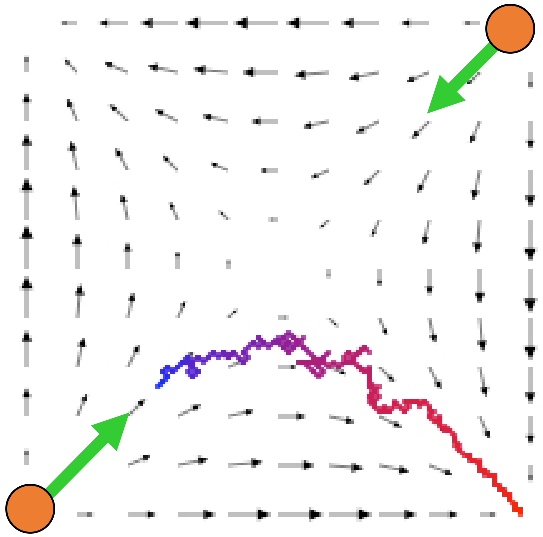
Lemma Weakly stable fixed points are Nash equilibria

Why?

fixed point: $p_e > 0 \Rightarrow$ Exp(cost)=cost(e)

If Exp(cost) > cost(e) (for some $p_e = 0$)

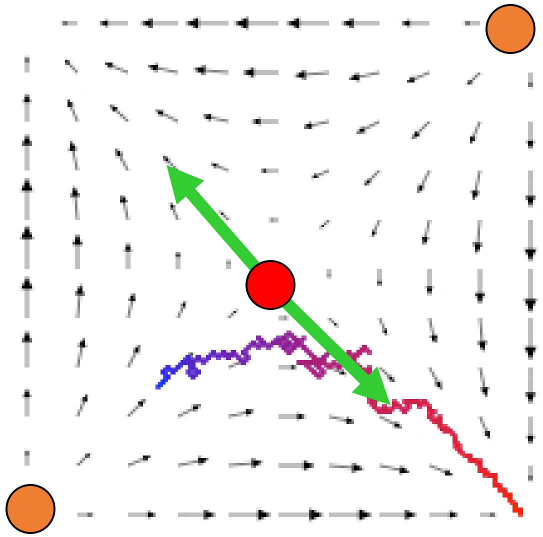
$\Rightarrow$ not stable



Example: 2-balls 2-bins

Weakly stable for ODE= neighborhood no direction has pull away from fixed point.

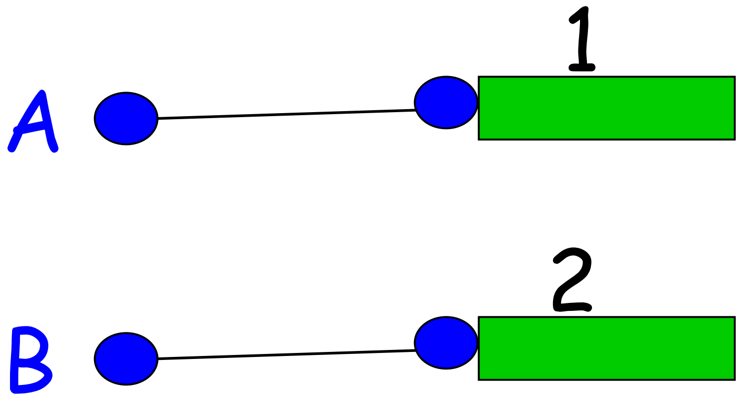
Fact: Mixed Nash in 2-balls 2-bins game not stable



Learning as a symmetry breaking

Simple case

2 balls & 2 bins



Players choose different bins ⇒

- they "learn" that chosen bin better
  other bin would have bigger congestion
- reinforcing the decision

Weakly Stable Nash?

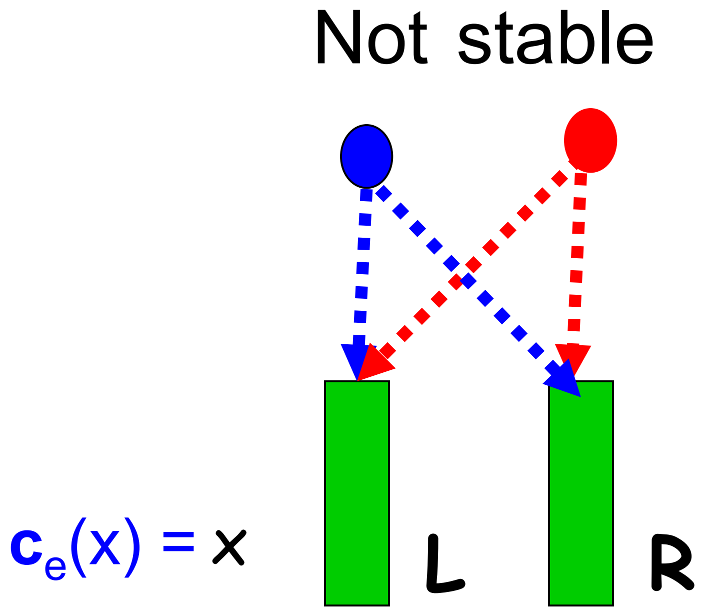
Weakly stable in games: each player remains indifferent between the strategies when one other player chooses a fixed strategy

Example: balls & bins

Weakly stable ⇒ at most one random player in each bin

Random Nash stable:

1 ball and 2 bins



Summary from this week

simple games and variants:

- matching pennies,
- coordination,
- prisoner's dilemma,
- Rock-paper-scissor

Learning algorithms

- Fictitious play, and smoothed versions

No-regret as outcome of learning or as a behavioral model

Price of Anarchy and learning outcomes (including changing environments) in

- Congestion games, such as traffic routing
- Auction games

Learning in multi-item auctions is hard,

Alternate learning we can do instead

Best Nash in congestion games, and what learning does in such games