

# Stochastic Equilibria under Imprecise Deviations in Terminal-Reward Concurrent Games

Patricia Bouyer, Nicolas Markey and Daniel Stan

CNRS, LSV, ENS Cachan & Université Paris Saclay

GandALF 2016



1 Concurrent framework

2 Existence of equilibria

3 Imprecise deviations

- 1 Concurrent framework
- 2 Existence of equilibria
- 3 Imprecise deviations

# Games with mixed strategies

Concurrent non-zero sum games allow ...

- To modelize heterogeneous systems;
- Several events to occur simultaneously;
- Agents' goals not to be necessarily antagonistic

## Games with mixed strategies

Concurrent non-zero sum games allow ...

- To modelize heterogeneous systems;
- Several events to occur simultaneously;
- Agents' goals not to be necessarily antagonistic

whereas mixed strategies enable ...

- Breaking the symmetry (by randomization)
- Equilibria more likely to occur.

# Games with mixed strategies

Concurrent non-zero sum games allow ...

- To modelize heterogeneous systems;
- Several events to occur simultaneously;
- Agents' goals not to be necessarily antagonistic

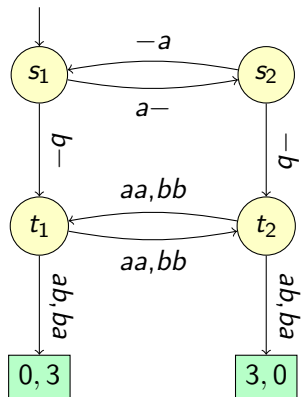
whereas mixed strategies enable ...

- Breaking the symmetry (by randomization)
- Equilibria more likely to occur.

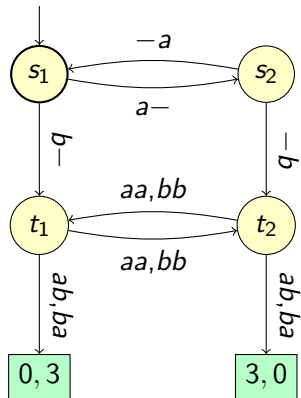
Main goals:

- Synthesizing strategies;
- As simple as possible.

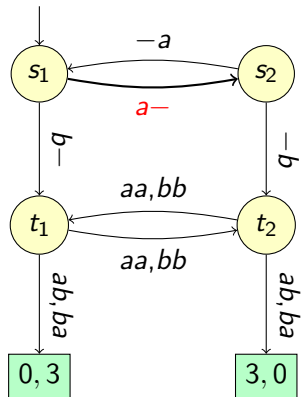
# Concurrent games: an example



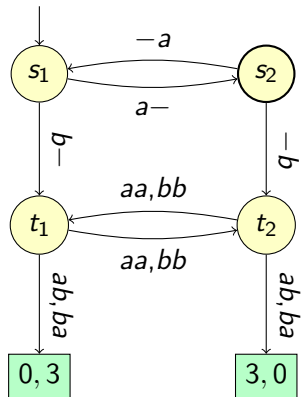
# Concurrent games: an example



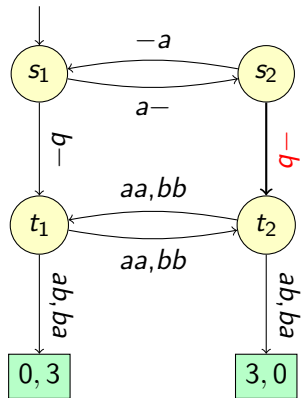
# Concurrent games: an example



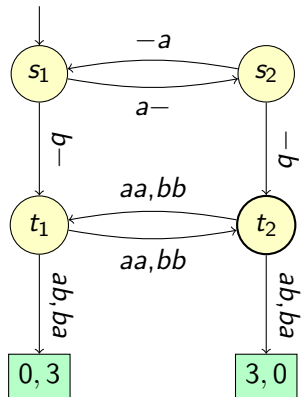
# Concurrent games: an example



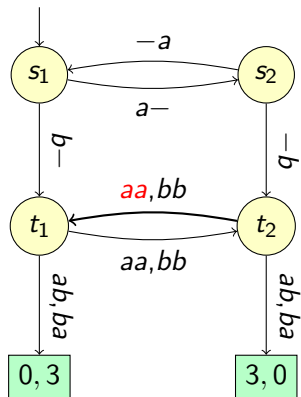
# Concurrent games: an example



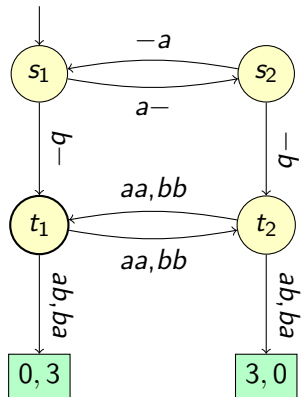
# Concurrent games: an example



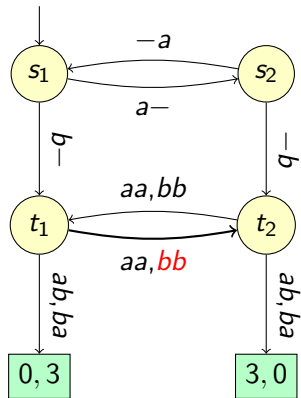
# Concurrent games: an example



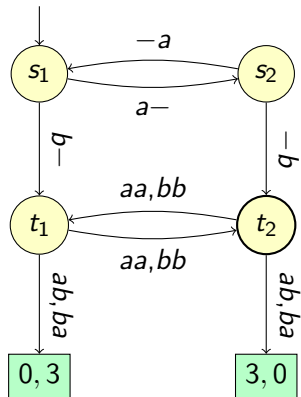
# Concurrent games: an example



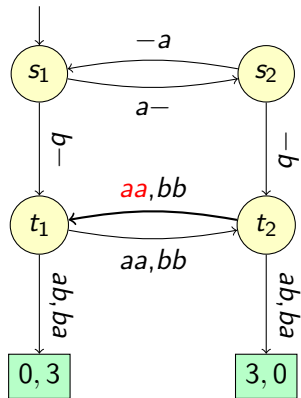
# Concurrent games: an example



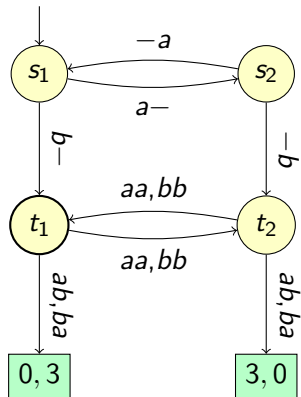
# Concurrent games: an example



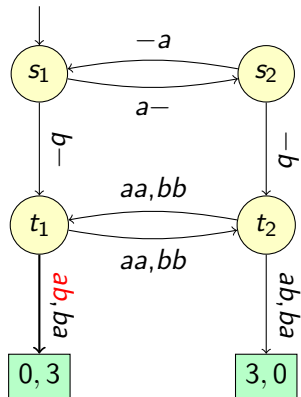
# Concurrent games: an example



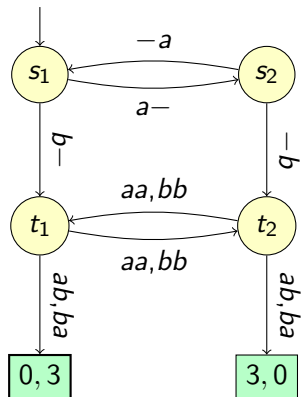
# Concurrent games: an example



# Concurrent games: an example

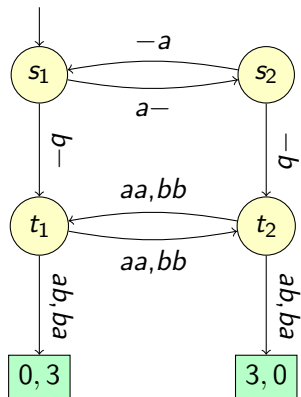


# Concurrent games: an example



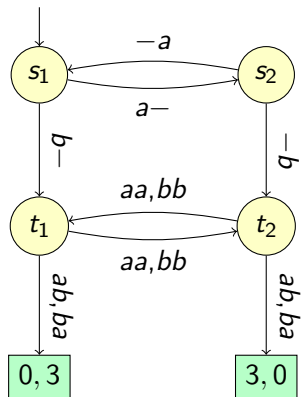
- Game on graph
- Several agents
- For each state  $s \in \text{States}$ , each  $i \in \text{Agt}$ , set of  $\text{Allow}_s(s)$  actions
- Transitions played concurrently
- Terminal rewards

# Concurrent games: an example



- Game on graph
- Several agents
- For each state  $s \in \text{States}$ , each  $i \in \text{Agt}$ , set of  $\text{Allow}_s(s)$  actions
- Transitions played concurrently
- Terminal rewards
- Also: stochastic transitions (players and environment)

# Concurrent games: an example



- Game on graph
- Several agents
- For each state  $s \in \text{States}$ , each  $i \in \text{Agt}$ , set of  $\text{Allow}_s(s)$  actions
- Transitions played concurrently
- Terminal rewards
- Also: stochastic transitions (players and environment)
- Non-zero sum (cycling rewards 0)

## Definition (Strategies)

A strategy for agent  $i$  is given by  $\sigma_i$  such that for all  $h \in \text{States}^+$ ,

$$\sigma_i(h) \in \text{Dist}(\text{Allow}_i(\text{last}(h)))$$

- Depends on sequence of visited states

## Definition (Strategies)

A strategy for agent  $i$  is given by  $\sigma_i$  such that for all  $h \in \text{States}^+$ ,

$$\sigma_i(h) \in \text{Dist}(\text{Allow}_i(\text{last}(h)))$$

- Depends on sequence of visited states
- $\sigma_i \in \mathbb{M}$  is **stationnary** if  $\forall h \sigma_i(h) = \sigma_i(\text{last}(h))$ ;

## Definition (Strategies)

A strategy for agent  $i$  is given by  $\sigma_i$  such that for all  $h \in \text{States}^+$ ,

$$\sigma_i(h) \in \text{Dist}(\text{Allow}_i(\text{last}(h)))$$

- Depends on sequence of visited states
- $\sigma_i \in \mathbb{M}$  is **stationnary** if  $\forall h \sigma_i(h) = \sigma_i(\text{last}(h))$ ;
- $\sigma_i \in \mathcal{S}$  is **pure** if all distributions give probability 1 to some action;

## Definition (Strategies)

A strategy for agent  $i$  is given by  $\sigma_i$  such that for all  $h \in \text{States}^+$ ,

$$\sigma_i(h) \in \text{Dist}(\text{Allow}_i(\text{last}(h)))$$

- Depends on sequence of visited states
- $\sigma_i \in \mathbb{M}$  is **stationnary** if  $\forall h \sigma_i(h) = \sigma_i(\text{last}(h))$ ;
- $\sigma_i \in \mathcal{S}$  is **pure** if all distributions give probability 1 to some action;
- **mixed** otherwise;
- **finite memory** if . . . .

# Nash Equilibrium

## Definition

We denote by  $\phi_i$  reward function for agent  $i$  and  $(\sigma_i)_{i \in \text{Agts}}$  the strategy profile.

# Nash Equilibrium

## Definition

We denote by  $\phi_i$  reward function for agent  $i$  and  $(\sigma_i)_{i \in \text{Agt}}$  the strategy profile.

$\sigma$  is a Nash Equilibrium (NE) if for all agent  $i$  and any other strategy for  $i$  (deviation)  $\sigma'_i$ ,

$$\mathbb{E}^{\sigma^{[i/\sigma'_i]}}(\phi_i) \leq \mathbb{E}^{\sigma}(\phi_i)$$

# Nash Equilibrium

## Definition

We denote by  $\phi_i$  reward function for agent  $i$  and  $(\sigma_i)_{i \in \text{Agt}}$  the strategy profile.

$\sigma$  is a Nash Equilibrium (NE) if for all agent  $i$  and any other strategy for  $i$  (deviation)  $\sigma'_i$ ,

$$\mathbb{E}^{\sigma^{[i/\sigma'_i]}}(\phi_i) \leq \mathbb{E}^{\sigma}(\phi_i)$$

No incentive to deviate to increase its own expected reward.

# Nash Equilibrium

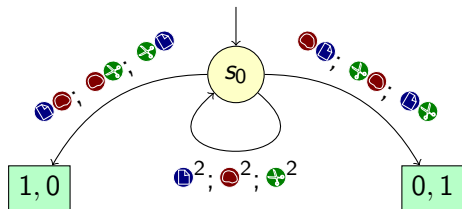
## Definition

We denote by  $\phi_i$  reward function for agent  $i$  and  $(\sigma_i)_{i \in \text{Agt}}$  the strategy profile.

$\sigma$  is a Nash Equilibrium (NE) if for all agent  $i$  and any other strategy for  $i$  (deviation)  $\sigma'_i$ ,

$$\mathbb{E}^{\sigma^{[i/\sigma'_i]}}(\phi_i) \leq \mathbb{E}^{\sigma}(\phi_i)$$

No incentive to deviate to increase its own expected reward.



# Nash Equilibrium

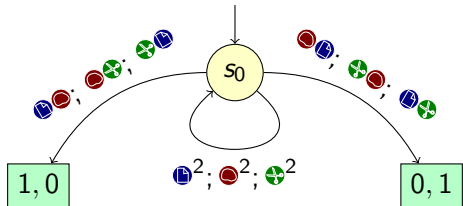
## Definition

We denote by  $\phi_i$  reward function for agent  $i$  and  $(\sigma_i)_{i \in \text{Agt}}$  the strategy profile.

$\sigma$  is a Nash Equilibrium (NE) if for all agent  $i$  and any other strategy for  $i$  (deviation)  $\sigma'_i$ ,

$$\mathbb{E}^{\sigma^{[i/\sigma'_i]}}(\phi_i) \leq \mathbb{E}^{\sigma}(\phi_i)$$

No incentive to deviate to increase its own expected reward.



$$\sigma_i = \frac{1}{3} (\text{red} + \text{blue} + \text{green})$$

# Nash Equilibrium

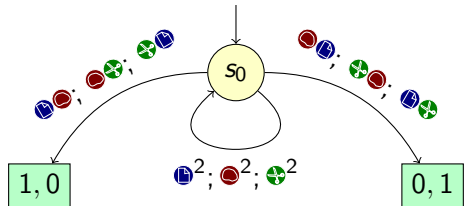
## Definition

We denote by  $\phi_i$  reward function for agent  $i$  and  $(\sigma_i)_{i \in \text{Agt}}$  the strategy profile.

$\sigma$  is a Nash Equilibrium (NE) if for all agent  $i$  and any other strategy for  $i$  (deviation)  $\sigma'_i$ ,

$$\mathbb{E}^{\sigma[i/\sigma'_i]}(\phi_i) \leq \mathbb{E}^{\sigma}(\phi_i)$$

No incentive to deviate to increase its own expected reward.



$$\sigma_i = \frac{1}{3} (\text{red} + \text{blue} + \text{green})$$

$$\Rightarrow \mathbb{E}^{\sigma}(\phi) = \left( \frac{1}{2}, \frac{1}{2} \right)$$

# Nash Equilibrium

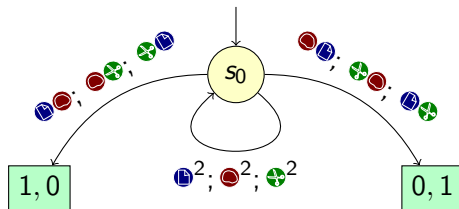
## Definition

We denote by  $\phi_i$  reward function for agent  $i$  and  $(\sigma_i)_{i \in \text{AgT}}$  the strategy profile.

$\sigma$  is a Nash Equilibrium (NE) if for all agent  $i$  and any other strategy for  $i$  (deviation)  $\sigma'_i$ ,

$$\mathbb{E}^{\sigma[i/\sigma'_i]}(\phi_i) \leq \mathbb{E}^{\sigma}(\phi_i)$$

No incentive to deviate to increase its own expected reward.



$$\sigma_1 = \sigma_2 = \frac{1}{3} (\text{red} + \text{blue} + \text{green})$$

is a NE of reward

$$\mathbb{E}^{\sigma}(\phi) = \left(\frac{1}{2}, \frac{1}{2}\right).$$

## Existence of equilibria under mixed strategies

### Theorem (Nash 1950)

*Every one-stage game has a Nash Equilibrium in mixed strategies.*

## Existence of equilibria under mixed strategies

### Theorem (Nash 1950)

*Every one-stage game has a Nash Equilibrium in mixed strategies.*

### Theorem (Secchi and Sudderth 2001)

*NE always exists for safety qualitative objectives in finite memory strategies.*

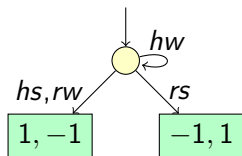
### Theorem (Chatterjee et al. 2004)

*For  $\varepsilon > 0$ ,  $\varepsilon$ -Nash Equilibrium always exists with terminal reward, and stationary strategies.*

$$\mathbb{E}^{\sigma^{[i/\sigma'_i]}}(\phi_i) \leq \mathbb{E}^{\sigma}(\phi_i) + \varepsilon$$

# Does a mixed Nash Equilibrium always exist?

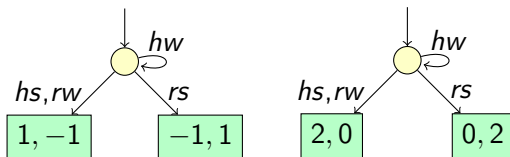
Idea: two player concurrent zero-sum games may not have optimal strategies but only  $\epsilon$ -optimal strategies (for any  $\epsilon > 0$ ).



Hide-or-Run game.

# Does a mixed Nash Equilibrium always exist?

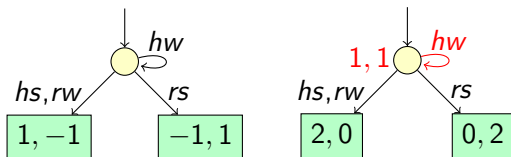
Idea: two player concurrent zero-sum games may not have optimal strategies but only  $\epsilon$ -optimal strategies (for any  $\epsilon > 0$ ).



Hide-or-Run game.

## Does a mixed Nash Equilibrium always exist?

Idea: two player concurrent zero-sum games may not have optimal strategies but only  $\epsilon$ -optimal strategies (for any  $\epsilon > 0$ ).



Hide-or-Run game.

Shifted hide-or-Run game

Value problem in a zero-sum game is not a special case of Nash Equilibrium problem with positive terminal rewards.

## NE on graphs are harder

Theorem (Bouyer et al. [2014])

*The existence problem is undecidable for 3-player concurrent games with non-negative terminal rewards and a constraint.*

## NE on graphs are harder

### Theorem (Bouyer et al. [2014])

*The existence problem is undecidable for 3-player concurrent games with non-negative terminal rewards and a constraint. Also holds on arbitrary terminal rewards without constraints.*

## NE on graphs are harder

### Theorem (Bouyer et al. [2014])

*The existence problem is undecidable for 3-player concurrent games with non-negative terminal rewards and a constraint. Also holds on arbitrary terminal rewards without constraints.*

### Theorem (Ummels [2010])

*There exist games where Nash equilibria require finite memory.*

- 1 Concurrent framework
- 2 Existence of equilibria
- 3 Imprecise deviations

## Equilibria existence proof: general scheme.

- Let  $\mathbb{M}$  be the set of stationary strategy profiles.
- Consider the best response function:

$$BR_i : \mathbb{M} \rightarrow 2^{\mathbb{M}}$$

mapping best strategy of  $i$  against a profile.

- Consider global best response  $BR = (BR_i)_{i \in \text{Agt}}$  and show it is continuous.
- Apply Kakutani fixed-point theorem:  $\exists \sigma \sigma \in BR(\sigma)$ .



## Equilibria existence proof: general scheme.

- Let  $\mathbb{M}$  be the set of stationary strategy profiles.
- Consider the best response function:

$$BR_i : \mathbb{M} \rightarrow 2^{\mathbb{M}}$$

mapping best strategy of  $i$  against a profile.

- Consider global best response  $BR = (BR_i)_{i \in \text{Agt}}$  and show it is continuous.
- Apply Kakutani fixed-point theorem:  $\exists \sigma \sigma \in BR(\sigma)$ .



Continuity of BR is based on termination assumptions.

## Equilibria existence proof: general scheme.

- Let  $\mathbb{M}$  be the set of stationary strategy profiles.
- Consider the best response function:

$$BR_i : \mathbb{M} \rightarrow 2^{\mathbb{M}}$$

mapping best strategy of  $i$  against a profile.

- Consider global best response  $BR = (BR_i)_{i \in \text{Agt}}$  and show it is continuous.
- Apply Kakutani fixed-point theorem:  $\exists \sigma \sigma \in BR(\sigma)$ .



Continuity of BR is based on termination assumptions.

- One-stage termination;

## Equilibria existence proof: general scheme.

- Let  $\mathbb{M}$  be the set of stationary strategy profiles.
- Consider the best response function:

$$BR_i : \mathbb{M} \rightarrow 2^{\mathbb{M}}$$

mapping best strategy of  $i$  against a profile.

- Consider global best response  $BR = (BR_i)_{i \in \text{Agt}}$  and show it is continuous.
- Apply Kakutani fixed-point theorem:  $\exists \sigma \sigma \in BR(\sigma)$ .



Continuity of BR is based on termination assumptions.

- One-stage termination;
- Safety condition, discounted reward;

## Equilibria existence proof: general scheme.

- Let  $\mathbb{M}$  be the set of stationary strategy profiles.
- Consider the best response function:

$$BR_i : \mathbb{M} \rightarrow 2^{\mathbb{M}}$$

mapping best strategy of  $i$  against a profile.

- Consider global best response  $BR = (BR_i)_{i \in \text{Agt}}$  and show it is continuous.
- Apply Kakutani fixed-point theorem:  $\exists \sigma \sigma \in BR(\sigma)$ .



Continuity of BR is based on termination assumptions.

- One-stage termination;
- Safety condition, discounted reward;
- Enforce terminating actions.

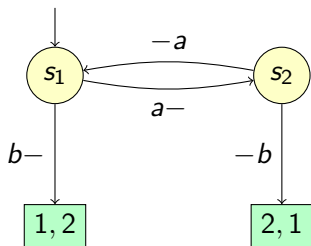
# Limit behaviour

War of wits: first player to play  $b$  gives up the game and loses.

# Limit behaviour

War of wits: first player to play  $b$  gives up the game and loses.

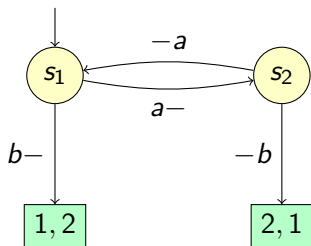
Note  $(x, y) = (\sigma_1(b | s_1), \sigma_2(b | s_2))$ .



# Limit behaviour

War of wits: first player to play  $b$  gives up the game and loses.

Note  $(x, y) = (\sigma_1(b | s_1), \sigma_2(b | s_2))$ .



- NE strategy profiles:  $(x, 0)$  and  $(0, x)$  for any  $1 \geq x > 0$ .

# Limit behaviour

War of wits: first player to play  $b$  gives up the game and loses.

Note  $(x, y) = (\sigma_1(b | s_1), \sigma_2(b | s_2))$ .

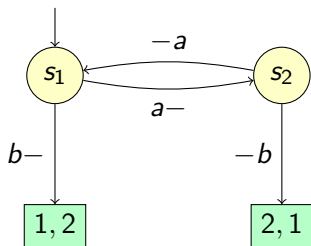


- NE strategy profiles:  $(x, 0)$  and  $(0, x)$  for any  $1 \geq x > 0$ .
- NE payoffs: respectively  $(1, 2)$  and  $(2, 1)$

# Limit behaviour

War of wits: first player to play  $b$  gives up the game and loses.

Note  $(x, y) = (\sigma_1(b | s_1), \sigma_2(b | s_2))$ .



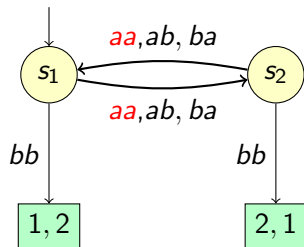
- NE strategy profiles:  $(x, 0)$  and  $(0, x)$  for any  $1 \geq x > 0$ .
- NE payoffs: respectively  $(1, 2)$  and  $(2, 1)$
- BR function graph not continuous in  $(0, 0)$ .

## Forcing termination: non-cycling game

A state is cycling if agents can enforce a cycle, without terminating deviation by any agent.

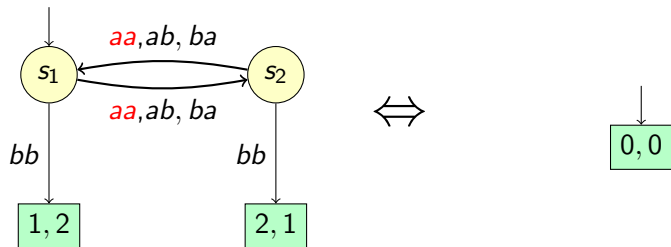
## Forcing termination: non-cycling game

A state is cycling if agents can enforce a cycle, without terminating deviation by any agent.



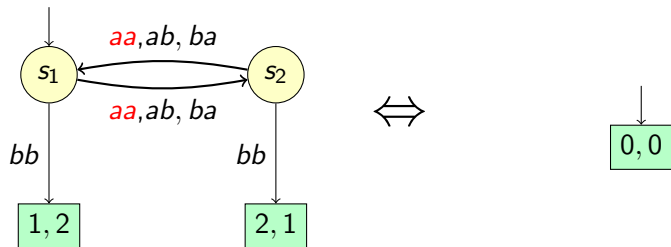
## Forcing termination: non-cycling game

A state is cycling if agents can enforce a cycle, without terminating deviation by any agent.



## Forcing termination: non-cycling game

A state is cycling if agents can enforce a cycle, without terminating deviation by any agent.



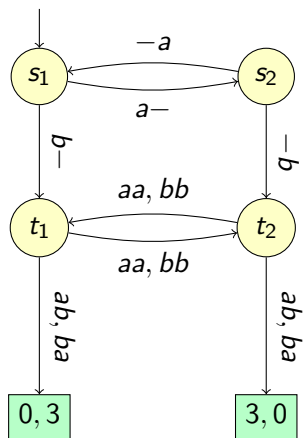
- Any equilibrium in the reduced game is an equilibrium in the original game.
- W.l.o.g, we assume there is always a player that can ensure positive probability termination.

## Cycle constraints: sketch

$C \subseteq \text{States}$  is cycling if there exists a strategy profile ensuring a stay in  $C$  from any state in  $C$ .

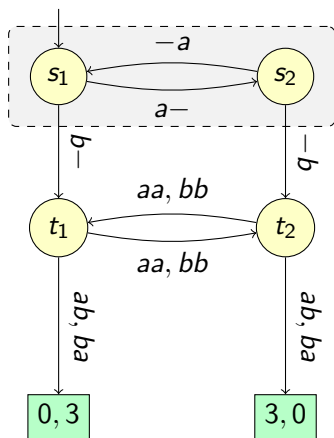
## Cycle constraints: sketch

$C \subseteq \text{States}$  is cycling if there exists a strategy profile ensuring a stay in  $C$  from any state in  $C$ .



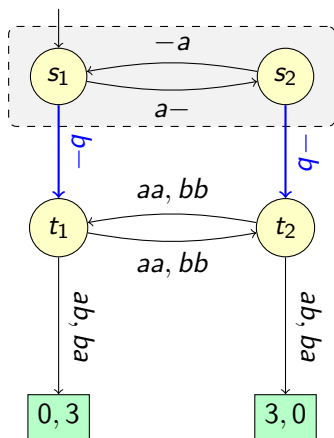
## Cycle constraints: sketch

$C \subseteq \text{States}$  is cycling if there exists a strategy profile ensuring a stay in  $C$  from any state in  $C$ .



## Cycle constraints: sketch

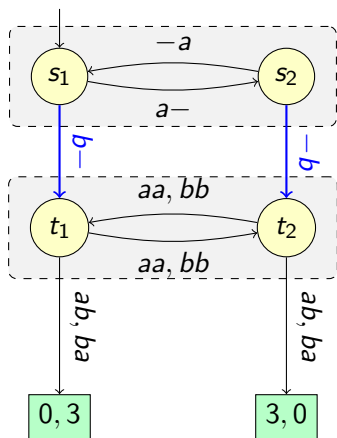
$C \subseteq \text{States}$  is cycling if there exists a strategy profile ensuring a stay in  $C$  from any state in  $C$ .



- $\sigma_1(b \mid s_1) \geq \epsilon; \sigma_2(b \mid s_2) \geq \epsilon$

## Cycle constraints: sketch

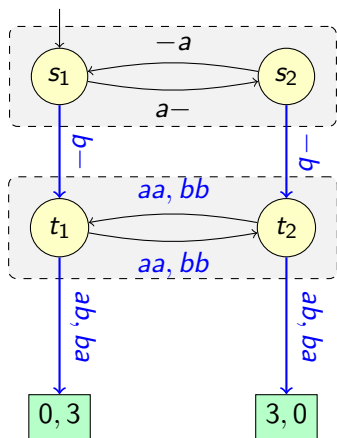
$C \subseteq \text{States}$  is cycling if there exists a strategy profile ensuring a stay in  $C$  from any state in  $C$ .



- $\sigma_1(b \mid s_1) \geq \epsilon; \sigma_2(b \mid s_2) \geq \epsilon$

## Cycle constraints: sketch

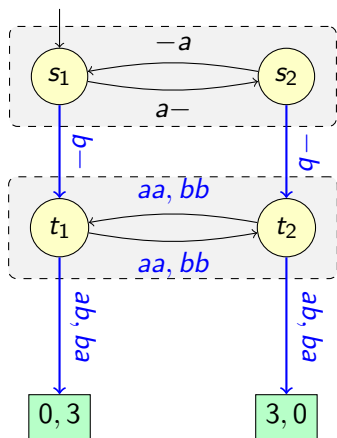
$C \subseteq \text{States}$  is cycling if there exists a strategy profile ensuring a stay in  $C$  from any state in  $C$ .



- $\sigma_1(b \mid s_1) \geq \epsilon; \sigma_2(b \mid s_2) \geq \epsilon$
- $\forall i, j, x \sigma_i(x \mid t_j) \geq \epsilon;$

## Cycle constraints: sketch

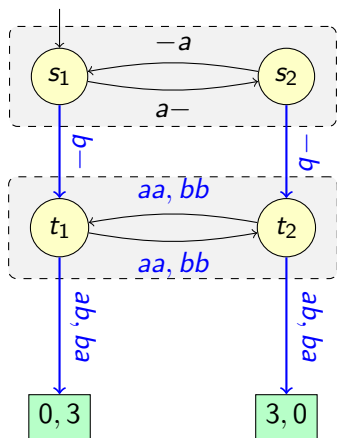
$C \subseteq \text{States}$  is cycling if there exists a strategy profile ensuring a stay in  $C$  from any state in  $C$ .



- $\sigma_1(b \mid s_1) \geq \varepsilon; \sigma_2(b \mid s_2) \geq \varepsilon$
- $\forall i, j, x \sigma_i(x \mid t_j) \geq \varepsilon;$
- Denote with  $\Delta_\varepsilon \subseteq \mathbb{M}$  the set of satisfying stationary strategy profiles.

## Cycle constraints: sketch

$C \subseteq \text{States}$  is cycling if there exists a strategy profile ensuring a stay in  $C$  from any state in  $C$ .



- $\sigma_1(b \mid s_1) \geq \varepsilon; \sigma_2(b \mid s_2) \geq \varepsilon$
- $\forall i, j, x \sigma_i(x \mid t_j) \geq \varepsilon;$
- Denote with  $\Delta_\varepsilon \subseteq \mathbb{M}$  the set of satisfying stationary strategy profiles.
- Remark  $\Delta_\varepsilon = \prod_i \Delta_\varepsilon^{(i)}$ .

# Bounding probability of termination

## Theorem

*For  $\varepsilon > 0$ , there exists  $p > 0$  and  $k \in \mathbb{N}$  such that for any  $\sigma \in \Delta_\varepsilon$ ,*

$$\forall s \in \text{States } \mathbb{P}^\sigma(s \cdot \text{States}^{\leq k} \cdot \mathbb{F}) \geq p$$

*That is to say, after  $k$  iterations, there is a bounded probability that a final state is reached.*

# Bounding probability of termination

## Theorem

*For  $\varepsilon > 0$ , there exists  $p > 0$  and  $k \in \mathbb{N}$  such that for any  $\sigma \in \Delta_\varepsilon$ ,*

$$\forall s \in \text{States } \mathbb{P}^\sigma(s \cdot \text{States}^{\leq k} \cdot \mathbb{F}) \geq p$$

*That is to say, after  $k$  iterations, there is a bounded probability that a final state is reached.*

Ensures almost-sure termination, and even more.

# Fixed-point

## Theorem (Best response function)

Let

$$\text{BR}_\varepsilon(\sigma) = \left\{ \sigma' \in \Delta_\varepsilon^{(i)} \mid \forall i \in \text{Agt} \forall s \in \text{States}, \mathbb{E}^{\sigma'[i/\sigma'_i]}(\phi_i \mid s) \geq \mathbb{E}^\sigma(\phi_i \mid s) \right\}$$

# Fixed-point

## Theorem (Best response function)

Let

$$\text{BR}_\varepsilon(\sigma) = \left\{ \sigma' \in \Delta_\varepsilon^{(i)} \mid \forall i \in \text{Agt} \forall s \in \text{States}, \mathbb{E}^{\sigma'[i/\sigma'_i]}(\phi_i \mid s) \geq \mathbb{E}^\sigma(\phi_i \mid s) \right\}$$

For  $0 < \varepsilon \leq \frac{1}{|\text{Act}|}$ ,  $\text{BR}_\varepsilon$  has a fixed point  $\sigma \in \text{BR}_\varepsilon(\sigma)$  in  $\Delta_\varepsilon$ .

# Fixed-point

## Theorem (Best response function)

Let

$$\text{BR}_\varepsilon(\sigma) = \left\{ \sigma' \in \Delta_\varepsilon^{(i)} \mid \forall i \in \text{Agt} \forall s \in \text{States}, \mathbb{E}^{\sigma'[i/\sigma'_i]}(\phi_i \mid s) \geq \mathbb{E}^\sigma(\phi_i \mid s) \right\}$$

For  $0 < \varepsilon \leq \frac{1}{|\text{Act}|}$ ,  $\text{BR}_\varepsilon$  has a fixed point  $\sigma \in \text{BR}_\varepsilon(\sigma)$  in  $\Delta_\varepsilon$ .

## Sketch.

$\text{BR}_\varepsilon$  graph is continuous over  $\Delta_\varepsilon$ , and for any  $\sigma \in \Delta_\varepsilon$ ,  $\text{BR}_\varepsilon(\sigma)$  is a non-empty closed convex, so fixed point theorem (Kakutani [1941]) holds. □

- 1 Concurrent framework
- 2 Existence of equilibria
- 3 Imprecise deviations**

- Previous fixed point is stationary;
- Not necessary a NE.

- Previous fixed point is stationary;
- Not necessary a NE.

### Definition

$\sigma \in \mathbb{S}$ , is an equilibrium under  $\varepsilon$ -imprecise deviations if for any player  $i$ ,

$$\forall \sigma'_i \exists \sigma''_i d(\sigma'_i, \sigma''_i) \leq \varepsilon, \quad \mathbb{E}^{\sigma^{[i/\sigma'_i]}}(\phi_i | h) \leq \mathbb{E}^{\sigma}(\phi_i | h)$$

with  $d(\sigma, \sigma')$  the maximal distance between distributions.

- Previous fixed point is stationary;
- Not necessary a NE.

### Definition

$\sigma \in \mathbb{S}$ , is an equilibrium under  $\varepsilon$ -imprecise deviations if for any player  $i$ ,

$$\forall \sigma'_i \exists \sigma''_i d(\sigma'_i, \sigma''_i) \leq \varepsilon, \quad \mathbb{E}^{\sigma^{[i/\sigma'_i]}}(\phi_i | h) \leq \mathbb{E}^{\sigma}(\phi_i | h)$$

with  $d(\sigma, \sigma')$  the maximal distance between distributions.

- A NE is an equilibria under  $\varepsilon$ -imprecise deviations.
- $\varepsilon$ -NE and equilibria under imprecise deviations are incomparable.

# Memoryless deviations

Do fixed points of  $BR_\epsilon$  fall in this category of equilibria ?

# Memoryless deviations

Do fixed points of  $BR_\varepsilon$  fall in this category of equilibria ? (Restriction to memoryless deviations)

$$\forall \sigma'_i \in \mathbb{M}_i; \exists \sigma''_i \in \mathbb{M}_i; d(\sigma'_i, \sigma''_i) \leq \varepsilon, \quad \mathbb{E}^{\sigma^{[i/\sigma'_i]}}(\phi_i | h) \leq \mathbb{E}^\sigma(\phi_i | h)$$

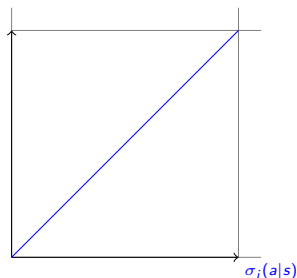
## Memoryless deviations

Consider  $\sigma \in \mathbb{M}$ , and an agent  $i$ , we construct a new game  $\widehat{\mathcal{G}}_{-i}^{\langle \sigma \rangle, \varepsilon}$  of the following form:

- The game is turn-based;
- All players except  $i$  have fixed strategy;
- $i$  is playing against antagonistic player  $\hat{i}$ ;

# Memoryless deviations

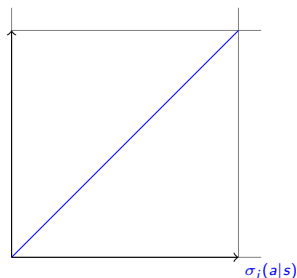
Consider  $\sigma \in \mathbb{M}$ , and an agent  $i$ , we construct a new game  $\widehat{\mathcal{G}}_{-i}^{\langle \sigma \rangle}^\varepsilon$  of the following form:



- The game is turn-based;
- All players except  $i$  have fixed strategy;
- $i$  is playing against antagonistic player  $\hat{i}$ ;
- $i$  suggests a new move;

# Memoryless deviations

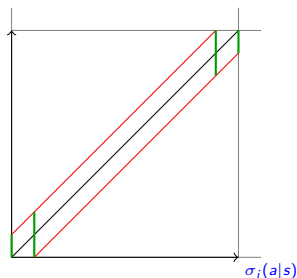
Consider  $\sigma \in \mathbb{M}$ , and an agent  $i$ , we construct a new game  $\widehat{\mathcal{G}}_{-i}^{\langle \sigma \rangle}^\varepsilon$  of the following form:



- The game is turn-based;
- All players except  $i$  have fixed strategy;
- $i$  is playing against antagonistic player  $\hat{i}$ ;
- $i$  suggests a new move;

# Memoryless deviations

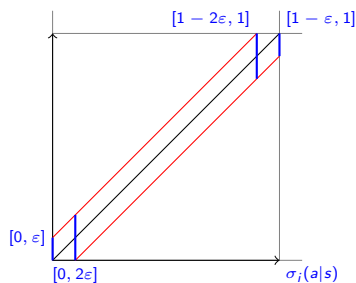
Consider  $\sigma \in \mathbb{M}$ , and an agent  $i$ , we construct a new game  $\widehat{\mathcal{G}}_{-i}^{\langle \sigma \rangle^\varepsilon}$  of the following form:



- The game is turn-based;
- All players except  $i$  have fixed strategy;
- $i$  is playing against antagonistic player  $\hat{i}$ ;
- $i$  suggests a new move;
- Then  $\hat{i}$  has some latitude to deviate.

# Memoryless deviations

Consider  $\sigma \in \mathbb{M}$ , and an agent  $i$ , we construct a new game  $\widehat{\mathcal{G}}_{-i}^{\langle \sigma \rangle, \varepsilon}$  of the following form:



- The game is turn-based;
- All players except  $i$  have fixed strategy;
- $i$  is playing against antagonistic player  $\hat{i}$ ;
- $i$  suggests a new move;
- Then  $\hat{i}$  has some latitude to deviate.



## Other consequence: computing a NE is in PSPACE

### Theorem

*Assume the number of actions fixed. For every  $\varepsilon > 0$ ,  $(x_i)_i, (y_i)_i \in \mathbb{R}^{\text{Agt}}$ , one can decide in polynomial space whether there exists a stationary equilibrium under  $\varepsilon$ -imprecise deviations  $\sigma$  such that for every  $i \in \text{Agt}$ ,  $x_i \leq \mathbb{E}^\sigma(\phi_i \mid s_0) \leq y_i$ .*

## Other consequence: computing a NE is in PSPACE

### Theorem

*Assume the number of actions fixed. For every  $\varepsilon > 0$ ,  $(x_i)_i, (y_i)_i \in \mathbb{R}^{\text{Agt}}$ , one can decide in polynomial space whether there exists a stationary equilibrium under  $\varepsilon$ -imprecise deviations  $\sigma$  such that for every  $i \in \text{Agt}$ ,  $x_i \leq \mathbb{E}^\sigma(\phi_i \mid s_0) \leq y_i$ .*

### Sketch.

As for classical NE (Ummels and Wojtczak [2011]), we compute some initial non-deterministic polynomial time pre-computation, to encode a stability formula  $\varphi \in \text{FO}(\mathbb{R}, \leq)$ .

## Other consequence: computing a NE is in PSPACE

### Theorem

*Assume the number of actions fixed. For every  $\varepsilon > 0$ ,  $(x_i)_i, (y_i)_i \in \mathbb{R}^{\text{Agt}}$ , one can decide in polynomial space whether there exists a stationary equilibrium under  $\varepsilon$ -imprecise deviations  $\sigma$  such that for every  $i \in \text{Agt}$ ,  $x_i \leq \mathbb{E}^\sigma(\phi_i \mid s_0) \leq y_i$ .*

### Sketch.

As for classical NE (Ummels and Wojtczak [2011]), we compute some initial non-deterministic polynomial time pre-computation, to encode a stability formula  $\varphi \in \text{FO}(\mathbb{R}, \leq)$ .

$\varphi$  encodes existence of some kind of stable profile:

**relies on the construction of  $\widehat{\mathcal{G}} \langle \sigma \rangle_{-i}^\varepsilon$ .**



# Overview

- Getting closer to exact NE existence problem (2 players),
- Usage of linear constraints to enforce a non-linear property (termination),
- New notion of equilibria, not equivalent to previous ones,

# Overview

- Getting closer to exact NE existence problem (2 players),
- Usage of linear constraints to enforce a non-linear property (termination),
- New notion of equilibria, not equivalent to previous ones,
- Still hope for exact NE in non-negative terminal reward games,
- NP-hardness can be adapted to this new notion.

# Overview

- Getting closer to exact NE existence problem (2 players),
- Usage of linear constraints to enforce a non-linear property (termination),
- New notion of equilibria, not equivalent to previous ones,
- Still hope for exact NE in non-negative terminal reward games,
- NP-hardness can be adapted to this new notion.

Thank you for your attention.

## Bibliography I

Patricia Bouyer, Nicolas Markey, and Daniel Stan. Mixed Nash equilibria in concurrent games. In Proceedings of the 34th Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS'14), volume 29 of Leibniz International Proceedings in Informatics, pages 351–363. Leibniz-Zentrum für Informatik, December 2014. doi: 10.4230/LIPIcs.FSTTCS.2014.351. URL <http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/BMS-fsttcs14.pdf>.

Krishnendu Chatterjee, Marcin Jurdziński, and Rupak Majumdar. On Nash equilibria in stochastic games. In CSL'04, volume 3210 of LNCS, pages 26–40. Springer, 2004.

Shizuo Kakutani. A generalization of brouwers fixed point theorem. Duke Math. J., 8(3):457–459, 09 1941. doi: 10.1215/S0012-7094-41-00838-4. URL <http://dx.doi.org/10.1215/S0012-7094-41-00838-4>.

## Bibliography II

- John F. Nash. Equilibrium points in  $n$ -person games. Proceedings of the National Academy of Sciences of the United States of America, 36(1): 48–49, 1950.
- Piercesare Secchi and William D. Sudderth. Stay-in-a-set games. International Journal of Game Theory, 30:479–490, 2001.
- Steven A. Lippman Thomas M. Liggett. Short notes: Stochastic games with perfect information and time average payoff. SIAM Review, 11(4): 604–607, 1969. ISSN 00361445. URL <http://www.jstor.org/stable/2029090>.
- Michael Ummels. Stochastic Multiplayer Games: Theory and Algorithms. Ph.D. Thesis, Department of Computer Science, RWTH Aachen, Germany, January 2010. URL <http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/ummels-phd10.pdf>.

# Bibliography III

Michael Ummels and Dominik Wojtczak. The complexity of Nash equilibria in stochastic multiplayer games. Logical Methods in Computer Science, 7(3), 2011.