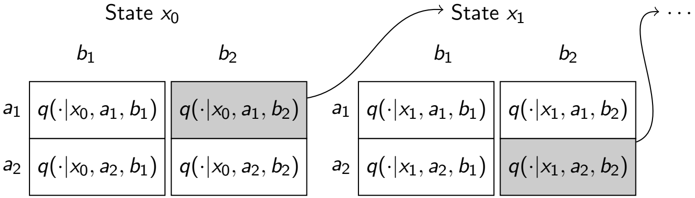
Subgame optimal strategies in zero-sum stochastic games with tolerance levels

János Flesch, Jean-Jacques Herings, Jasmine Maes,
Arkadi Predtetchinski

Maastricht University
The Netherlands

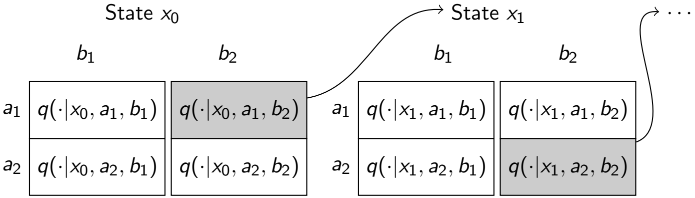
14 March 2019
Mons, Belgium
Theory and Algorithms in Graph and Stochastic Games

Model: Two-player zero-sum stochastic game



Generates a play $p = \underbrace{x_0}_{\text{initial state}} \underbrace{a_1 b_1 x_1}_{t=1} \underbrace{a_2 b_2 x_2}_{t=2} \underbrace{a_3 b_3 x_3}_{t=3} \ldots$

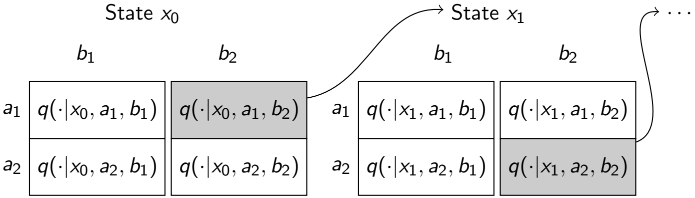
Model: Two-player zero-sum stochastic game



$$\text{Generates a play } p = \underbrace{x_0}_{\text{initial state}} \; \underbrace{a_1 b_1 x_1}_{t=1} \underbrace{a_2 b_2 x_2}_{t=2} \underbrace{a_3 b_3 x_3}_{t=3} \ldots$$

- Finite action spaces $\mathcal{A}$ and $\mathcal{B}$ and countable state space $\mathcal{X}$.

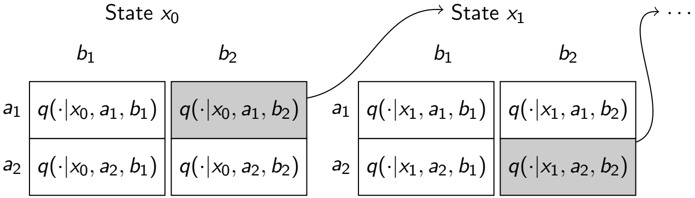
Model: Two-player zero-sum stochastic game



$$\text{Generates a play } p = \underbrace{x_0}_{\text{initial state}} \quad \underbrace{a_1 b_1 x_1}_{t=1} \underbrace{a_2 b_2 x_2}_{t=2} \underbrace{a_3 b_3 x_3}_{t=3} \ldots$$

- Finite action spaces $\mathcal{A}$ and $\mathcal{B}$ and countable state space $\mathcal{X}$.
- Transition probability q

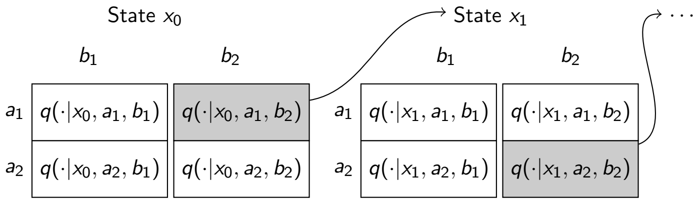
Model: Two-player zero-sum stochastic game



Generates a play $p = \underbrace{x_0}_{\text{initial state}} \underbrace{a_1 b_1 x_1}_{t=1} \underbrace{a_2 b_2 x_2}_{t=2} \underbrace{a_3 b_3 x_3}_{t=3} \ldots$

- Finite action spaces $\mathcal{A}$ and $\mathcal{B}$ and countable state space $\mathcal{X}$.
- Transition probability q
- Set of plays $\mathcal{P}$

Model: Two-player zero-sum stochastic game



Generates a play $p = \underbrace{x_0}_{\text{initial state}} \underbrace{a_1 b_1 x_1}_{t=1} \underbrace{a_2 b_2 x_2}_{t=2} \underbrace{a_3 b_3 x_3}_{t=3} \ldots$

- Finite action spaces $\mathcal{A}$ and $\mathcal{B}$ and countable state space $\mathcal{X}$.
- Transition probability q
- Set of plays $\mathcal{P}$
- Bounded **Borel measurable** payoff function $u : \mathcal{P} \to \mathbb{R}$.
  (can be generalized to universally measurable)

Model: Two player zero-sum stochastic game

- Set of histories: $\mathcal{H}$

  $h = \underbrace{x_0}_{\text{initial state}} \underbrace{a_1 b_1 x_1}_{t=1} \underbrace{a_2 b_2 x_2}_{t=2} \underbrace{a_3 b_3 x_3}_{t=3}$, length $\|h\| = 3$
- Behavioral strategies:

  Player 1 $\sigma : \mathcal{H} \to \Delta(\mathcal{A})$ Player 2 $\tau : \mathcal{H} \to \Delta(\mathcal{B})$
- Probability space $(\mathcal{P}, \mathcal{F}, \mathbb{P}_{h,\sigma,\tau})$
- Value exists (Maitra, Sudderth (1998), Martin (1998))

  $$v(h) = \sup_{\sigma \in \mathcal{S}_1} \inf_{\tau \in \mathcal{S}_2} \mathbb{E}_{h,\sigma,\tau}[u] = \inf_{\tau \in \mathcal{S}_2} \sup_{\sigma \in \mathcal{S}_1} \mathbb{E}_{h,\sigma,\tau}[u]$$

Model: Two player zero-sum stochastic game

- Set of histories: $\mathcal{H}$

  $h = \underbrace{x_0}_{\text{initial state}} \quad \underbrace{a_1 b_1 x_1}_{t=1} \underbrace{a_2 b_2 x_2}_{t=2} \underbrace{a_3 b_3 x_3}_{t=3}$, length $\|h\| = 3$
- Behavioral strategies:

  Player 1 $\sigma : \mathcal{H} \to \Delta(\mathcal{A})$ Player 2 $\tau : \mathcal{H} \to \Delta(\mathcal{B})$
- Probability space $(\mathcal{P}, \mathcal{F}, \mathbb{P}_{h,\sigma,\tau})$
- Value exists (Maitra, Sudderth (1998), Martin (1998))

  $$v(h) = \sup_{\sigma \in \mathcal{S}_1} \inf_{\tau \in \mathcal{S}_2} \mathbb{E}_{h,\sigma,\tau}[u] = \inf_{\tau \in \mathcal{S}_2} \sup_{\sigma \in \mathcal{S}_1} \mathbb{E}_{h,\sigma,\tau}[u]$$

Take the perspective of maximizing player

Main concept:

Subgame ϕ-optimal strategies

Strategies that perform "good enough" across all subgames.

Main concept: Subgame ϕ-optimal strategy

Optimal strategy

The strategy $\sigma \in \mathcal{S}_1$ is an **optimal strategy** if for every $\tau \in \mathcal{S}_2$:

$$\mathbb{E}_{\sigma,\tau}\left[u\right] \geq v.$$

Main concept: Subgame ϕ-optimal strategy

Optimal strategy

The strategy $\sigma \in \mathcal{S}_1$ is an **optimal strategy** if for every $\tau \in \mathcal{S}_2$:

$$\mathbb{E}_{\sigma,\tau}[u] \geq v.$$

Subgame optimal strategy

The strategy $\sigma \in \mathcal{S}_1$ is a **subgame optimal strategy** if for every $\tau \in \mathcal{S}_2$ and for every history $h \in \mathcal{H}$:

$$\mathbb{E}_{h,\sigma,\tau}[u] \geq v(h).$$

Main concept: Subgame ϕ-optimal strategy

Optimal strategy

The strategy $\sigma \in \mathcal{S}_1$ is an **optimal strategy** if for every $\tau \in \mathcal{S}_2$:

$$\mathbb{E}_{\sigma,\tau}[u] \geq v.$$

Subgame optimal strategy

The strategy $\sigma \in \mathcal{S}_1$ is a **subgame optimal strategy** if for every $\tau \in \mathcal{S}_2$ and for every history $h \in \mathcal{H}$:

$$\mathbb{E}_{h,\sigma,\tau}[u] \geq v(h).$$

Subgame ϕ-optimal strategy

The strategy $\sigma \in \mathcal{S}_1$ is a **subgame ϕ-optimal strategy** if for every $\tau \in \mathcal{S}_2$ and for every history $h \in \mathcal{H}$:

$$\mathbb{E}_{h,\sigma,\tau}[u] \geq v(h) - \phi(h).$$

Special cases and equilibria

Subgame ϕ-optimal strategy

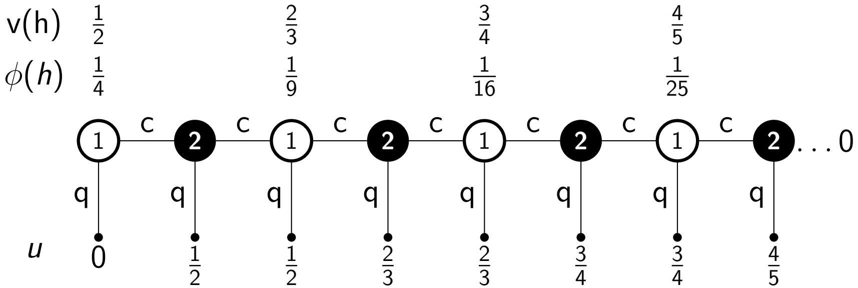
The strategy $\sigma \in \mathcal{S}_1$ is a **subgame ϕ-optimal strategy** if for every $\tau \in \mathcal{S}_2$ and for every history $h \in \mathcal{H}$:

$$\mathbb{E}_{h,\sigma,\tau}[u] \geq v(h) - \phi(h).$$

- $\phi(h) = 0$ everywhere $\Rightarrow$ subgame optimal strategy.
  $\sim$ Subgame perfect equilibrium
- $\phi(h) = \epsilon$ everywhere $\Rightarrow$ subgame ϵ-optimal strategy.
  $\sim$ Subgame perfect ϵ-equilibrium
- Tolerance function $\phi : \mathcal{H} \to [0, \infty)$
  $\sim$ ϕ-tolerance equilibrium (Flesch, Predtetchinski (2016))

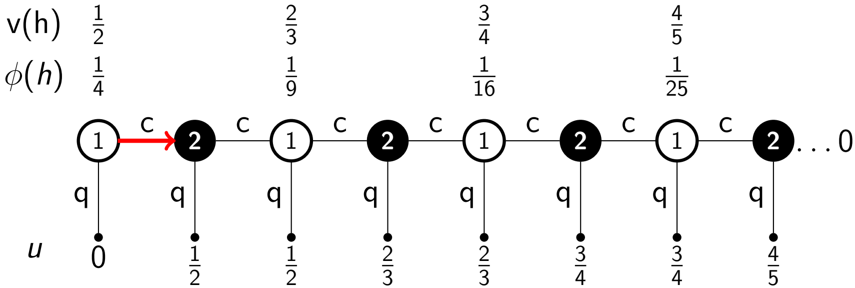
Simplified example

$$\mathbb{E}_{h,\sigma,\tau}[u] = u(\pi(\sigma,\tau;h)) \geq v(h) - \phi(h).$$



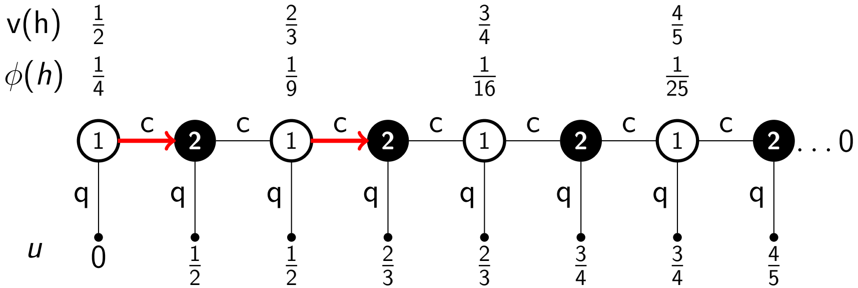
Simplified example

$$\mathbb{E}_{h,\sigma,\tau}[u] = u(\pi(\sigma,\tau;h)) \geq v(h) - \phi(h).$$



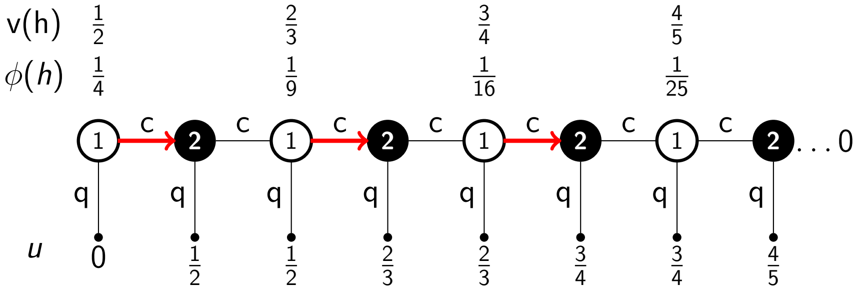
Simplified example

$$\mathbb{E}_{h,\sigma,\tau}[u] = u(\pi(\sigma,\tau;h)) \geq v(h) - \phi(h).$$



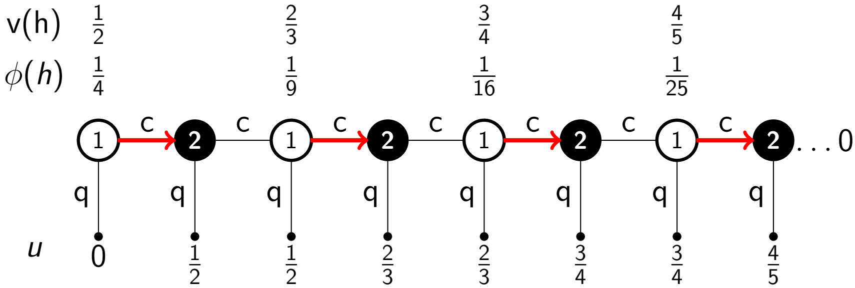
Simplified example

$$\mathbb{E}_{h,\sigma,\tau}\left[u\right] = u(\pi(\sigma,\tau;h)) \geq v(h) - \phi(h).$$



Simplified example

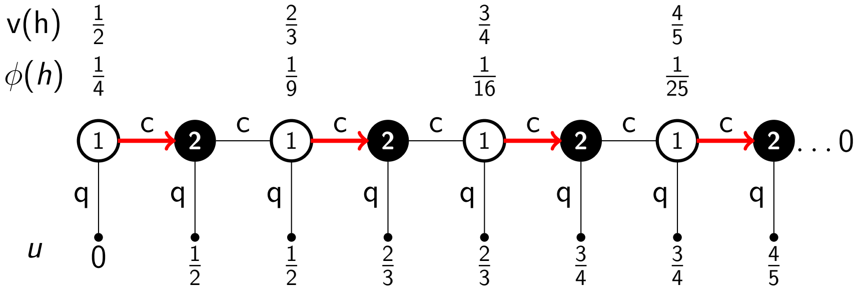
$$\mathbb{E}_{h,\sigma,\tau}[u] = u(\pi(\sigma,\tau;h)) \geq v(h) - \phi(h).$$



- NO subgame ϕ-optimal strategy, (even if $\phi > 0$).

Simplified example

$$\mathbb{E}_{h,\sigma,\tau}[u] = u(\pi(\sigma,\tau;h)) \geq v(h) - \phi(h).$$



- NO subgame ϕ-optimal strategy, (even if $\phi > 0$).
- For every $\epsilon > 0 \Rightarrow$ there exists a subgame ϵ-optimal strategy.

Question 1:

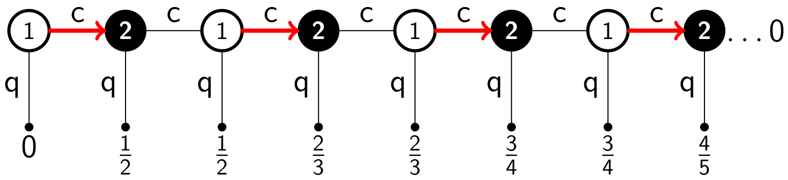
What are the **necessary and sufficient conditions** for a strategy to be a subgame ϕ-optimal strategy?

Characterization for subgame optimal strategies

Theorem (Characterization)

A strategy σ is a **subgame optimal strategy** *for player 1 if and only if for every $\tau \in \mathcal{S}_2$ and for every $h \in \mathcal{H}$ with $\|h\| = t$ we have that:*

1. **(1-day optimal)** $\mathbb{E}_{h,\sigma,\tau}\left[V^{t+1}\right] \geq v(h)$.
2. **(equalizing)** $u \geq \limsup_{n\to\infty} V^n, \qquad \mathbb{P}_{h,\sigma,\tau} - a.s.$



Sufficient conditions for subgame ϕ-optimal strategies

Theorem (Characterization)

A strategy σ is a **subgame optimal strategy** *for player 1 if and only if for every $\tau \in \mathcal{S}_2$ and for every $h \in \mathcal{H}$ with $\|h\| = t$ we have that:*

1. **(1-day optimal)** $\mathbb{E}_{h,\sigma,\tau}\left[V^{t+1}\right] \geq v(h)$.
2. **(equalizing)** $u \geq \limsup_{n\to\infty} V^n, \qquad \mathbb{P}_{h,\sigma,\tau} - a.s.$

Theorem (Sufficient condition)

A strategy σ is a **subgame ϕ-optimal strategy** *for player 1 if for every $\tau \in \mathcal{S}_2$ and for every $h \in \mathcal{H}$ with $\|h\| = t$ there exist $\phi_1(h)$ and $\phi_2(h)$ such that:*

1. $\phi_1(h) + \phi_2(h) = \phi(h)$
2. **(n-day ϕ_1-optimal)** $\mathbb{E}_{h,\sigma,\tau}\left[V^{t+n}\right] \geq v(h) - \phi_1(h), \forall n \in \mathbb{N}$.
3. **(ϕ_2-equalizing)** $u \geq \limsup_{n\to\infty} V^n - \phi_2(h), \qquad \mathbb{P}_{h,\sigma,\tau} - a.s.$

Necessary condition for subgame ϕ-optimal strategies

Theorem (Characterization)

A strategy σ is a **subgame optimal strategy** *for player 1 if and only if for every $\tau \in \mathcal{S}_2$ and for every $h \in \mathcal{H}$ with $\|h\| = t$ we have that:*

1. **(1-day optimal)** $\mathbb{E}_{h,\sigma,\tau}\left[V^{t+1}\right] \geq v(h)$.
2. **(equalizing)** $u \geq \limsup_{n\to\infty} V^n, \qquad \mathbb{P}_{h,\sigma,\tau} - a.s.$

Theorem (Necessary condition)

If a strategy σ is a **subgame ϕ-optimal strategy** *for player 1 then for every $\tau \in \mathcal{S}_2$ and for every $h \in \mathcal{H}$ with $\|h\| = t$:*

1. **(n-day ϕ-optimal)** $\mathbb{E}_{h,\sigma,\tau}\left[V^{t+n}\right] \geq v(h) - \phi(h), \ \forall n \in \mathbb{N}$.
2. **(ϕ-equalizing)** $u \geq \limsup_{n\to\infty}(V^n - \Phi^n), \qquad \mathbb{P}_{h,\sigma,\tau} - a.s.$

Question 2:

When does a subgame ϕ-optimal strategy exist?

(Assume $\phi > 0$)

Existence of subgame ϕ-optimal strategies

Theorem

Player 1 has a subgame ϕ-optimal strategy for $\phi > 0$, if every $p \in \mathcal{P}$ satisfies at least one of the following conditions:

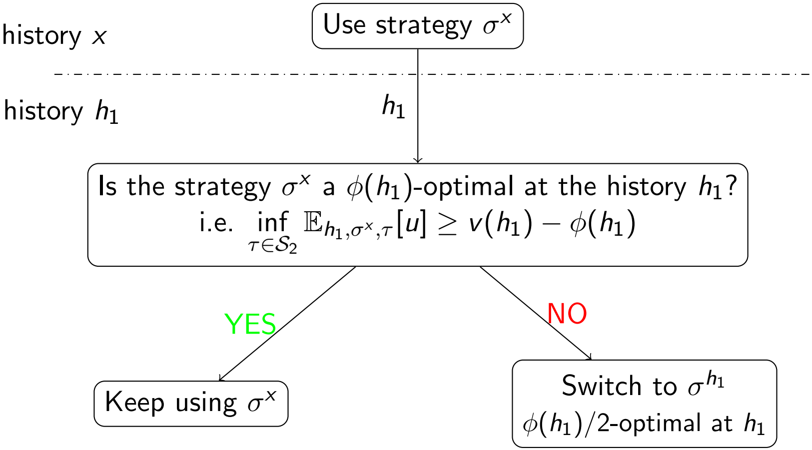
1. **(point of upper-semicontinuity)**
   If $\lim_{t\to\infty} p_t = p$ *then* $u(p) \geq \limsup_{t\to\infty} u(p_t)$,
2. **(positive limit inferior)**
   $\liminf_{t\to\infty} \phi(p_{|t}) > 0$.

Extreme cases:

- The payoff-function is upper-semicontinuous.
  Laraki, Maitra and Sudderth (2013)
- $\phi(h) = \epsilon$ for every $h \in \mathcal{H}$.
  Mashiah-Yaakovi (2015)

Construction: Use a switching strategy σ^{ϕ}

σ^h is a strategy which is $\phi(h)/2$ optimal at history h



Intuition

To be a subgame ϕ-optimal strategy the switching strategy needs to satisfy that for every $\tau \in \mathcal{S}_2$, $h \in \mathcal{H}$ with $\|h\| = t$:

1. $\mathbb{E}_{h,\sigma,\tau}[V^{t+n}] \geq v(h) - \phi(h)/2$, ✓
2. $u \geq \limsup_{n\to\infty} V^n - \phi(h)/2 \quad \mathbb{P}_{h,\sigma,\tau}$-a.s.
    1. Is always fulfilled along plays which are points of upper semicontinuity.
    2. Is fulfilled when we only switch finitely often.
       (Along plays with $\liminf_{t\to\infty} \phi(p_{|t}) > 0$ the probability that we need to switch infinitely often is 0.)

Question 3:

What is the relationship between the existence of subgame optimal strategies and the existence of subgame ϕ-optimal strategies?

Equivalence theorem

Theorem

There exists a subgame ϕ-optimal strategy for every $\phi > 0$.
$\Leftrightarrow$ *There exists a subgame optimal strategy.*

Equivalence theorem

Theorem

There exists a subgame ϕ-optimal strategy for every $\phi > 0$.
$\Leftrightarrow$ *There exists a subgame optimal strategy.*

- **Difference with subgame ϵ-optimal strategies:**
  There exists a subgame ϵ-optimal strategy for every $\epsilon > 0$.
  $\not\Rightarrow$ There exists a subgame optimal strategy.

Equivalence theorem

Theorem

There exists a subgame ϕ-optimal strategy for every $\phi > 0$.
$\Leftrightarrow$ *There exists a subgame optimal strategy.*

- **Difference with subgame ϵ-optimal strategies:**
  There exists a subgame ϵ-optimal strategy for every $\epsilon > 0$.
  $\not\Rightarrow$ There exists a subgame optimal strategy.
- **Not continuity property:** payoff-function u is only assumed to be bounded and Borel-measurable.

Equivalence theorem

Theorem

There exists a subgame ϕ-optimal strategy for every $\phi > 0$.
$\Leftrightarrow$ *There exists a subgame optimal strategy.*

- **Difference with subgame ϵ-optimal strategies:**
  There exists a subgame ϵ-optimal strategy for every $\epsilon > 0$.
  $\not\Rightarrow$ There exists a subgame optimal strategy.
- **Not continuity property:** payoff-function u is only assumed to be bounded and Borel-measurable.
- **Instead:** Use necessary conditions for subgame ϕ-optimal strategy and the characterization for subgame optimal strategies.

Conclusion

1. We study **Subgame ϕ-optimal strategies**, i.e. strategies of the maximizing player that guarantee the value in every subgame up till a subgame dependent tolerance level
2. We give necessary and sufficient conditions for such subgame ϕ-optimal strategies.
3. We provide conditions for the existence of subgame ϕ-optimal strategies.
4. We show that if a subgame ϕ-optimal strategy exists for any positive tolerance function ϕ, then so does a subgame optimal strategy.