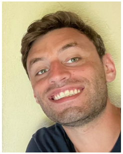
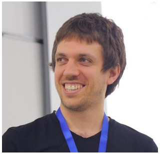
Taming Partial Observation in Stochastic Games: the blind case

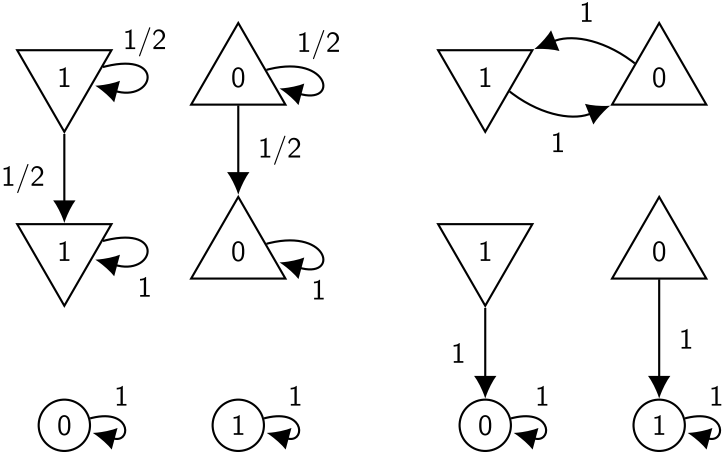

K. Chatterjee[1] D. Lurie[2] **R. Saona**[3] B. Ziliotto[4]

[1]Institute of Science and Technology Austria (ISTA)
[2]Paris-Dauphine University
[3]London School of Economics and Political Science (LSE)
[4]CNRS at Toulouse School of Economics

CGO seminar at LSE — Oct 2025

Can I discretize my continuous space and still study limit properties of the dynamic?

Stochastic Games

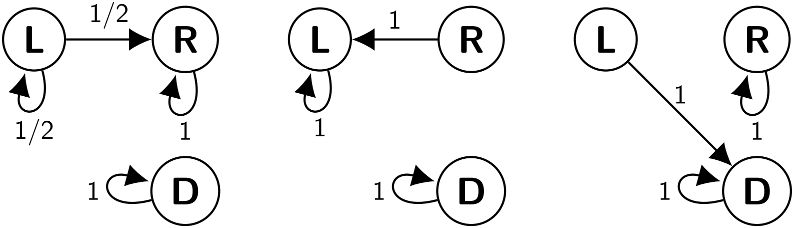


(a) Wait

(b) Commit

Single-controller stochastic game with actions (a) `Wait` and (b) `Commit`.

Simple blind MDP



(a) Approach (b) Restart (c) Commit

Blind MDP with actions (a) `Approach`, (b) `Restart`, and (c) `Commit`.

Definitions

A Blind Stochastic Game is a tuple $\Gamma = (\mathcal{S}, \mathcal{I}, \mathcal{J}, \delta, r, s_1)$ where

- $\mathcal{S}$ is a finite set of **states**.
- $\mathcal{I}$ and $\mathcal{J}$ are finite sets of **actions** for each player.
- $\delta \colon \mathcal{S} \times \mathcal{I} \times \mathcal{J} \to \Delta(\mathcal{S})$ is a probabilistic **transition** function.
- $r \colon \mathcal{S} \to \mathbb{R}$ is a **reward** function.
- $s_1 \in \mathcal{S}$ is an **initial state**.

Players play simultaneously and observe each others actions. Therefore, they have the same belief over the current state.

Limit Value

Denote σ and τ general strategies for the players.
For $\lambda \in (0, 1)$, the λ-objective of the players is to optimize

$$\gamma_\lambda(\sigma, \tau) := \mathbb{E}^{\sigma,\tau} \left((1 - \lambda) \sum_{t=1}^{\infty} \lambda^{t-1} \, r(S_t) \right) .$$

The value is defined as

$$\mathsf{val}_\lambda := \min_\sigma \max_\tau \gamma_\lambda(\sigma, \tau) = \max_\tau \min_\sigma \gamma_\lambda(\sigma, \tau) \,.$$

The limit value is defined as

$$\mathsf{val} := \lim_{\lambda \to 1^-} \mathsf{val}_\lambda \ .$$

Previous results

Mertens' Conjecture

Conjecture (1987, International Congress of Mathematics)

In every (zero-sum) stochastic game, the limit value exists.

Proven in many special cases of stochastic games.

Limit Value: Existence

Theorem (2002, Rosenberg & Solan & Vieille, Annals of Statistics)

Every blind 1-player stochastic game (MDP) has a limit value.

Limit Value: Nonexistence

Theorem (2016, Bruno Ziliotto, Annals of Probability)

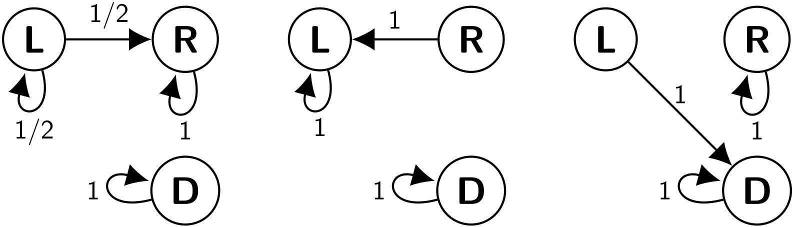
There exists a blind stochastic game where the limit value does not exist.

Limit Value: Undecidability

Theorem (2003, Madani & Hanks & Condon, Artificial Intelligence)

The problem of recognizing bounds ε-apart from the limit value of blind MDPs is undecidable.

Difficulty:
Absorbing states can accumulate arbitrarily small contributions. So, the player(s) behaviour depends on nonapproximable effects because, in the limit value, they are infinitely patient.



(a) Approach (b) Restart (c) Commit

Blind MDP with actions (a) Approach, (b) Restart, and (c) Commit.

Ergodic transitions

Ergodicity: Forgetting where you come from

In Markov Chains, an ergodic transition probability P satisfies

$$\lim_{n\to\infty} P^n = \mathbb{1}\mu^\top .$$

Equivalently, for all $p \in \Delta(\mathcal{S})$, we have that

$$p^\top \lim_{n\to\infty} P^n = \mu^\top .$$

In particular, $s, \widetilde{s}, s' \in \mathcal{S}$

$$\lim_{n\to\infty} \left| P^n_{s,s'} - P^n_{\widetilde{s},s'} \right| = 0 .$$

Coefficient of Ergodicity

Definition (Coefficient of Ergodicity)

Given a matrix $P \in \mathbb{R}^{n \times n}$, define

$$\operatorname{erg}(P) := \max_{s, \widetilde{s} \in [n]} \sum_{s' \in [n]} \left| P^n_{s,s'} - P^n_{\widetilde{s},s'} \right| .$$

Note that

- $\operatorname{erg}(PQ) \leq \operatorname{erg}(P) \ \operatorname{erg}(Q)$.
- $\operatorname{erg}(P) = 0$ if and only if $P = \mathbb{1}\mu^\top$.

Definition (Ergodic blind stochastic game)

For all $\varepsilon > 0$, there exists an integer n_ε such that,
for all $n \geq n_\varepsilon$ and tuples of action pairs $(i_1, j_1, \ldots, i_n, j_n)$,

$$\operatorname{erg}\Big(P(i_1, j_1) \cdots P(i_n, j_n)\Big) \leq \varepsilon \,.$$

Intuitively, the current belief is approximated by considering only the last n_ε actions:
no need to remember your initial distribution!

Verifying Ergodicity

By a counting argument, we get the following result.

Proposition (Paz, 1971, Introduction to Probabilistic Automata)

A blind stochastic game Γ *is ergodic if and only if there exists an integer* $n_1 := \frac{3^{|\mathcal{S}|} - 2^{|\mathcal{S}|+1} + 1}{2}$ *is such that, for every tuples of action pairs* $(i_1, j_1, \ldots, i_{n_1}, j_{n_1})$*,*

$$\operatorname{erg}\left(P(i_1, j_1) \cdots P(i_{n_1}, j_{n_1}) \right) < 1 \,.$$

Our Contributions

Theorem

Every ergodic blind stochastic game has a limit value.

Proof sketch.

- Construct a finite stochastic game based on n_ε steps at a time.
- Belief dynamics remain close between the original and approximated model.
- Finite-stage payoff remain close between the models.

□

Limit Value: Approximability

Theorem

Approxmating the limit value of an ergodic blind stochastic game can be done in 2-EXPSPACE.

Proof sketch.

- The previous construction requires 2-EXP states.
- Approximating the limit value can be done by solving a sentence of the first order theory of the reals, which is PSPACE on the input.

□

Limit Value: Undecidability

Theorem

The problem of recognizing lower and upper bounds of the limit value of ergodic blind MDPs is undecidable.

Proof sketch.

- Consider an arbitrary blind MDP.
- Add a positive transition to a new state and a restart action.
- These modifications do not change the limit value, because the controller is infinitely patient.
- Remarkably, the transitions are now ergodic!

□

Summary of Contributions

Blind Class	Existence	Approximation	Exact
SGs	No	–	–
Ergodic SGs	**Yes**	**2-EXPSPACE**	**Undecidable**
MDPs	Yes	Undecidable	Undecidable
Ergodic MDPs	Yes	**2-EXPSAPCE**	**Undecidable**

Summary of results

Thank you!