Zero-Sum Stochastic Games with Frequent Actions and Vanishing Discount Factors

Ivan Novikov

Université Paris 1 Panthéon-Sorbonne

29/06/2026

Introduction

- We study stochastic games: dynamic games in which the players act at stages, receiving a payoff at each stage.
- Our stochastic game is parametrized by two parameters: discount factor λ (players' patience parameter) and stage duration h (frequency parameter).
- Our goal is to discuss some known and open questions about the behavior of the value as both $h \to 0$ and $\lambda \to 0$.
- We consider the case of imperfect state observation since the perfect observation case is trivial.

Table of contents

Stochastic Games with public signals (1)

A *zero-sum stochastic game with public signals* is a 7-tuple $(A, \Omega, f, I, J, g, P)$, where:

- A is a set of signals;
- Ω is a set of states;
- $f : \Omega \to A$ is a signaling function;
- I is a set of actions of Player 1;
- J is a set of actions of Player 2;
- $g : I \times J \times \Omega \to \mathbb{R}$ is the stage payoff function of Player 1;
- $P : I \times J \times \Omega \to \Delta(\Omega)$ is the transition probability function.

We assume that I, J, Ω, A are finite.

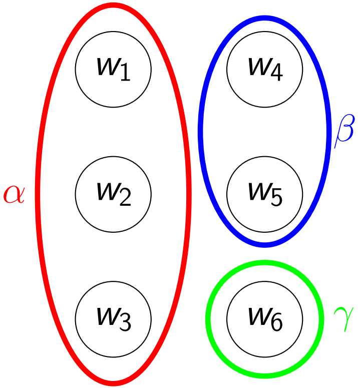
Stochastic Games with public signals (2)

The game $(A, \Omega, f, I, J, g, P)$ proceeds in stages as follows. At each stage n:

1. The current state is ω_n. Players do not observe it, but they observe the signal $\alpha_n = f(\omega_n) \in A$ and the actions of each other at the previous stage;
2. Players choose their mixed actions, $x_n \in \Delta(I)$ and $y_n \in \Delta(J)$;
3. Pure actions $i_n \in I$ and $j_n \in J$ are chosen according to $x_n \in \Delta(I)$ and $y_n \in \Delta(J)$;
4. Player 1 obtains a payoff $g_n = g(i_n, j_n, \omega_n)$, while Player 2 obtains the payoff $-g_n$;
5. The new state ω_{n+1} is chosen according to the probability law $P(\cdot \mid \omega_n, i_n, j_n)$. The new signal is $\alpha_{n+1} = f(\omega_{n+1})$.

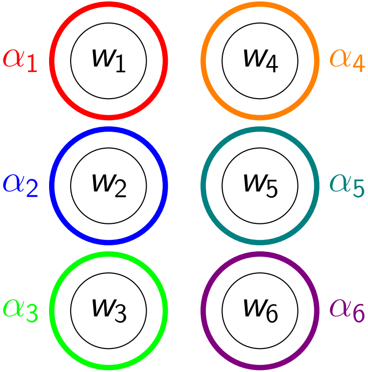
The above description of the game is known to the players. Players do not observe the payoff.

An example of the partition function f (1)



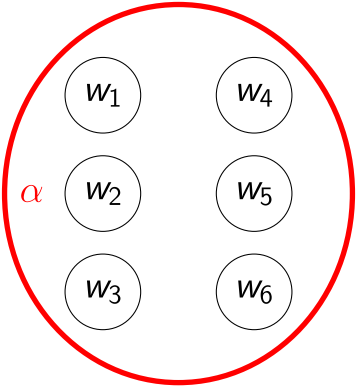
There are 3 public signals, and $f(\omega_1) = f(\omega_2) = f(\omega_3) = \alpha$, $f(\omega_4) = f(\omega_5) = \beta$, $f(\omega_6) = \gamma$.

Examples of the partition function f (2)



The full observation of the state, i.e. there are 6 public signals $\alpha_1, \dots, \alpha_6$; and $f(w_i) := \alpha_i$.

Examples of the partition function f (3)



The state-blind case. There is only one signal α, and $f(w_i) := \alpha$

Table of contents

Strategies and total payoff

- Strategies σ, τ of players consist in choosing at each stage a mixed action;
- The players can take into account the previous actions of players, as well as the current and previous signals.
- λ-discounted total payoff: $E^p_{\sigma,\tau}\left(\lambda \sum_{i=1}^{\infty}(1-\lambda)^{i-1} g_i\right)$;
- Depends on $\lambda \in (0,1)$, initial distribution p, and strategies of the players;
- Value $v_\lambda : \Delta(\Omega) \to \mathbb{R}$:

$$
\begin{aligned}
v_\lambda(p) &= \sup_\sigma \inf_\tau E^p_{\sigma,\tau}\left(\lambda \sum\nolimits_{i=1}^{\infty}(1-\lambda)^{i-1} g_i\right) \\
&= \inf_\tau \sup_\sigma E^p_{\sigma,\tau}\left(\lambda \sum\nolimits_{i=1}^{\infty}(1-\lambda)^{i-1} g_i\right).
\end{aligned}
$$

1-player case

- Suppose that J is a singleton;
- Player 2 cannot influence the game;
- The value is $v_\lambda(p) = \sup\limits_{\sigma} E^p_\sigma \left(\lambda \sum_{i=1}^{\infty}(1-\lambda)^{i-1} g_i\right)$.
- It is an MDP (Markov decision process);
- If I and J are singletons, then the game is a Markov chain.

Limit of λ-discounted game Γ^λ

- $v_\lambda(p) = \sup_\sigma \inf_\tau E^p_{\sigma,\tau}\left(\lambda \sum_{i=1}^\infty (1-\lambda)^{i-1} g_i\right)$;
- One can ask: what happens if the players become more and more patient? I.e., players are willing to wait a lot to obtain a big payoff;
- Mathematically, it means that $\lambda \to 0$;
- Thus, one is interested in the limit $\lim_{\lambda \to 0} v_\lambda(p)$;
- This limit is called *asymptotic value*.
- Asymptotic value always exists for the full observation case, but may fail to exist in a more general signaling setting.
- Asymptotic value always exists for the one player setting, even with signals.

Table of contents

Stage duration (1)

- Fix a base stochastic game $G_1 = (S, I, J, \Omega, f, g, P)$.
- In the game G_h with stage duration h, the leaving probabilities and the discount factor are scaled by $h \in (0, 1)$.
- $P_h(\cdot \mid \omega, i, j) = hP(\cdot \mid \omega, i, j) + (1 - h)\delta_\omega(\cdot)$, where

$$\delta_\omega(\omega') = \begin{cases} 1, & \text{if } \omega' = \omega; \\ 0, & \text{otherwise.} \end{cases}$$

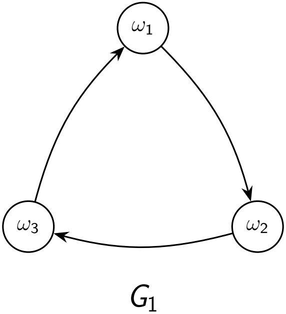
- The λ-discounted payoff of G_h is

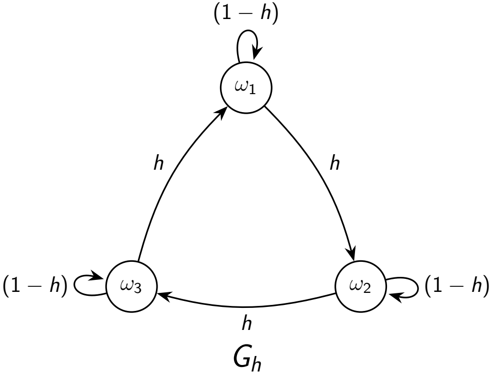
$$\mathbb{E}^h_{\sigma,\tau}\left(\sum_{i=1}^{\infty} \lambda h (1 - \lambda h)^{i-1} g_i\right).$$

- It depends on the stage duration h, the strategies σ, τ, and the initial probability distribution p.
- $V_\lambda(h)$ is the λ-discounted value of G_h.

Stage duration (2)

- h can be thought of as a time period between actions: when h is small, G_h approximates the continuous-time behavior, and $\lim_{h \to 0} V_\lambda(h)$ can be seen as the λ-discounted value of the continuous-time stochastic game.





Stage duration (3)

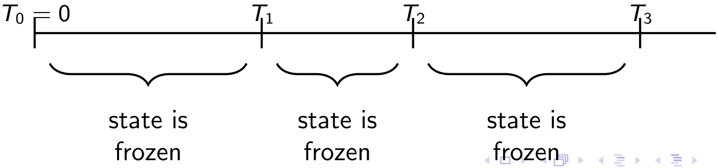
- X_i is the random variable with

$$X_i \sim Bernoulli(h) \iff$$
$$\Big(\mathbb{P}(X_i = 1) = h \quad \text{and} \quad \mathbb{P}(X_i = 0) = 1 - h\Big)$$

- X_i indicates whether the transition at the end of the i-th stage is governed by P (if $X_i = 1$) or by δ_ω (if $X_i = 0$).
- T_i is the random variable

$$T_0 = 0;$$
$$T_i = \inf\{n > T_{i-1} \mid X_n = 1\} \quad \text{for } i > 0.$$

- T_i is the stage number at which the state transition is governed by P for the i-th time.



Papers on the stage duration

- Neyman A, "Stochastic games with short-stage duration" (2013);
- Sorin S and Vigeral G, "Operator approach to values of stochastic games with varying stage duration" (2016).
- I. N., Zero-Sum State-Blind Stochastic Games with Vanishing Stage Duration (2025);
- I.N., Asymptotic Value in Zero-Sum Stochastic Games with Vanishing Stage Duration and Public Signals (2024).
- I.N., Strategies in POMDPs with Stage Duration (2026).

Similar model:

- Sorin S, "Limit Value of Dynamic Zero-Sum Games with Vanishing Stage Duration" (2018);
- Cardaliaguet P, Rainer C, Rosenberg D, Vieille N, "Markov Games with Frequent Actions and Incomplete Information—The Limit Case" (2016);
- Gensbittel F, "Continuous-time limit of dynamic games with incomplete information and a more informed player" (2016).

Table of contents

Our goal

- We want to know what happens when the players are patient and the actions are frequent.
- This means that $\lambda \to 0$ and $h \to 0$.
- That is, we consider the limits

$$\lim_{\lambda \to 0} \lim_{h \to 0} V_\lambda(h) \quad \text{and} \quad \lim_{h \to 0} \lim_{\lambda \to 0} V_\lambda(h).$$

$$\lim_{\lambda \to 0} \lim_{h \to 0} V_\lambda(h) = \text{ the asymptotic value of the continuous-time game.}$$

$$\lim_{h \to 0} \lim_{\lambda \to 0} V_\lambda(h) = \text{ the continuous-time limit of the game with asymptotic payoff.}$$

The case of full state observation

- Sorin and Vigeral (2016) have proven that

$$V_\lambda(h) = V_{\frac{\lambda}{1+\lambda-\lambda h}}(1).$$

- This makes most of the following questions trivial!

General case with signals

We have the following facts

Known fact (I.N. 2026)

$\lim_{h\to 0} \lim_{\lambda \to 0} V_\lambda(h)$ exists for MDPs.

Known fact (I.N. 2025; and Sylvain Sorin 2016)

$\lim_{h\to 0} V_\lambda(h)$ exists for state-blind games, for any fixed λ.

Asymptotic value for MDPs

Known fact (Rosenberg, Solan, and Vieille 2002)

The asymptotic value $\lim_{\lambda \to 0} V_\lambda(1)$ of a discrete-time MDP with signals exists.

Question 1: The existence of the asymptotic value in continuous time

Does the limit $\lim_{\lambda \to 0} \lim_{h \to 0} V_\lambda(h)$ exist?

Question 1 trivially holds for MDPs with full state observation.

Asymptotic value for games

Known fact (Ziliotto 2016)

There is a finite stochastic game with signals, in which the asymptotic value $\lim_{\lambda \to 0} V_\lambda(1)$ in discrete time does not exist.

Note that there are other games with the same properties (in Renault and Ziliotto 2020; in Sorin and Vigeral 2015), but they are all conceptually the same.

Known fact (I.N. 2024) + Computer experiments

For the game from Ziliotto 2016, we have

1. $\lim_{\lambda \to 0} V_\lambda(h)$ does not exist for any h.
2. $\lim_{\lambda \to 0} \lim_{h \to 0} V_\lambda(h)$ exists.
3. $\left| \limsup_{\lambda \to 0} V_\lambda(h) - \liminf_{\lambda \to 0} V_\lambda(h) \right| \to 0$ as $h \to 0$.

Asymptotic value for games (2)

Meaning of the above result: While the asymptotic value does not exist in discrete time, the oscillations become smaller as h becomes smaller. In the continuous-time limit, the asymptotic value begins to exist.
We may ask now:

Question 2

Is there an example of a finite stochastic game with signals, in which the asymptotic value $\lim_{\lambda\to 0}\lim_{h\to 0} V_\lambda(h)$ in continuous time does not exist?

Question 3

Can we say that the limit $\lim_{\lambda\to 0}\lim_{h\to 0} V_\lambda(h)$ always exists for any finite stochastic game with signals?

Question 4

Can we say that the existence of the limit $\lim_{\lambda\to 0} V_\lambda(h)$ implies the existence of the limit $\lim_{\lambda\to 0}\lim_{h\to 0} V_\lambda(h)$?

Questions 3 and 4 trivially hold for games with full observation.

Permutation of the limits

For games, there is in general no limit permutation:

Known fact (I.N. 2024) + Computer experiments

There is a finite stochastic game with signals in which the limit $\lim_{\lambda\to 0}\lim_{h\to 0} V_\lambda(h)$ exists, but the limit $\lim_{h\to 0}\lim_{\lambda\to 0} V_\lambda(h)$ does not.

However, such a counterexample is impossible for MDPs. We can ask:

Question 5

Is it true that $\lim_{\lambda\to 0}\lim_{h\to 0} V_\lambda(h) = \lim_{h\to 0}\lim_{\lambda\to 0} V_\lambda(h)$ for MDPs with signals?

The following question tries to correct the problem with the example.

Question 6

Is it true that $\lim_{\lambda\to 0}\lim_{h\to 0} V_\lambda(h) = \lim_{h\to 0}\lim_{\lambda\to 0} V_\lambda(h)$ for stochastic games with signals, provided that the limit $\lim_{\lambda\to 0} V_\lambda(h)$ exists for all h small enough?

Questions 5 and 6 trivially hold for games with full observation.

Tauberian theorem

- λ-discounted payoff: $\lambda \sum_{i=1}^{\infty} (1-\lambda)^{i-1} g_i$. The value V_λ.
- Payoff of horizon T: $\frac{1}{T} \sum_{i=1}^{T} g_i$. The value V_T.

Tauberian theorem (Lehrer and Sorin 1992) for MDPs

For any finite MDP with signals, if at least one of the limits converges uniformly:
1) $\lim_{\lambda \to 0} V_\lambda$;
2) $\lim_{T \to \infty} V_T$,
then these two limits are equal to each other.

Tauberian theorem (Ziliotto 2016) for stochastic games

For any finite stochastic game with signals, if at least one of the limits converges uniformly:
1) $\lim_{\lambda \to 0} V_\lambda$;
2) $\lim_{T \to \infty} V_T$,
then these two limits are equal to each other.

Tauberian theorem with frequent actions

- λ-discounted payoff: $\sum_{i=1}^{\infty} \lambda h(1-\lambda h)^{i-1} g_i$. The value $V_\lambda(h)$.
- Payoff of horizon T: $\frac{h}{T} \sum_{i=1}^{\lfloor T/h \rfloor} g_i$. The value $V_T(h)$.

Question 7: Tauberian theorem in continuous time

For any finite game (or MDP) with signals, if at least one of the limits converges uniformly:
1) $\lim_{\lambda \to 0} \lim_{h \to 0} V_\lambda(h)$;
2) $\lim_{T \to \infty} \lim_{h \to 0} V_T(h)$,
can we say that these two limits are equal to each other?

Question 7 holds for games with full observation.

This is all.

Thank you!