Zero-Sum Stochastic Games with Vanishing Stage Duration and Public Signals

Ivan Novikov

Université Paris-Dauphine, CEREMADE

03/12/2025

Introduction

- We wish to examine the behavior of zero-sum stochastic games in continuous-time: the players control a finite continuous-time Markov chain and continuously receive the payoff depending on the current state.
- In a game with stage duration h, the players engage in a continuous-time game, but they can act only at specific times $0, h, 2h, \ldots$
- This yields a game with stage duration h. This is a discrete-time game!
- To study the behavior of continuous-time stochastic games, we can consider games with stage duration h, and see what happens when h vanishes.
- My thesis is dedicated to the study of games with stage duration in the context of public signals.

Table of contents

Definition (1)

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

- A is a finite set of signals;
- Ω is a finite set of states;
- I is a finite set of actions of player 1;
- J is a finite 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;
- $f : I \times J \times \Omega \times \Omega \to \Delta(A)$ is the signaling function.

The function f gives a random signal for each action profile and each pair of the current and the next state.

Definition (2)

The game $(A, \Omega, f, I, J, g, P)$ proceeds in stages as follows. Before the first stage, an initial state $\omega_1 \in \Omega$ and an initial signal $\alpha_1 \in A$ are chosen according to some probability law $p_0 \in \Delta(A \times \Omega)$. At each stage $n \in \mathbb{N}^*$:

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

Remark about the definition

- Particular case: full observation. $A = \Omega$ and $f(i, j, \omega, \omega') = \omega'$.
- Particular case: state-blind case. A is a singleton.
- Our definition is not completely standard.
- It gives more information.
- $P(\omega'|i, j, \omega) = 0$ means that after playing (i, j), the probability to change the state from ω to ω' is 0.
- But we still need to specify a distribution $f(i, j, \omega, \omega') \in \Delta(A)$.
- This is useful for games with stage duration: in a continuous-time game, the players remain in each state ω for some time. $f(i, j, \omega, \omega)$ specifies the signal that the players receive.

Strategies and total payoff

- Strategies σ, τ of players consist of choosing at each stage a mixed action, depending on the past;
- Let $\lambda \in (0,1]$. The λ-discounted payoff:

$$E^{\omega}_{\sigma,\tau}\left(\sum_{i=1}^{\infty}\lambda(1-\lambda)^{i-1}g_i\right);$$

- Depends on the initial state ω and strategies of the players;
- In general, we can fix any non-increasing positive sequence $B = \{b_m\}$, $\sum b_i < \infty$, and consider the total payoff $E^{\omega}_{\sigma,\tau}\left(\sum_{i=1}^{\infty} b_i g_i\right)$.

Value

- Value $v_\lambda : \Omega \to \mathbb{R}$:

$$v_\lambda(\omega) = \sup_\sigma \inf_\tau E^\omega_{\sigma,\tau}\left(\sum_{m=1}^\infty \lambda(1-\lambda)^{m-1} g_i\right)$$
$$= \inf_\tau \sup_\sigma E^\omega_{\sigma,\tau}\left(\sum_{m=1}^\infty \lambda(1-\lambda)^{m-1} g_i\right).$$

- The value exists in our finite case by Shapley 53.

Table of contents

Continuous-time stochastic games

- Finite state space Ω, action spaces I, J of two players;
- Instantaneous payoff function g.
- Infinitesimal generator of the game
  $q : I \times J \to \{\text{matrices } |\Omega| \times |\Omega| \text{ satisfying property } *\}$.
- Matrix $A = (a_{ij})$ satisfies property $*$, if $a_{ij} \geq 0$ for all $i \neq j$, $a_{ii} \leq 0$, and $\sum_{j=1}^{|\Omega|} a_{ij} = 0$ for all i.
- If the state at time t is ω' and the players play (i, j) in the interval $[t, t+h]$, then at the time $t+h$ the state is distributed according to $e^{hq(i,j)}(\omega', \cdot)$.

Continuous-time Markov games (2)

- Players choose their (Markov) strategies $\sigma : \Omega \times [0, \infty) \to \Delta(I)$, $\tau : \Omega \times [0, \infty) \to \Delta(J)$.
- There are some measurability conditions.
- Initial state is ω_0.
- λ-discounted payoff: $E^{\omega_0}_{\sigma,\tau}\left(\int_0^\infty \lambda e^{-\lambda t} g(i_t, j_t, \omega_t) dt\right)$.
- The value is defined as before.

Stochastic games with stage duration

- We want to approximate a continuous-time game by stochastic games.
- Consider a family of stochastic games G_h, parametrized by $h \in (0,1]$.
- h represents stage duration.
- Players can play only at times $0, h, 2h, \ldots$
- State can change only at times $h, 2h, \ldots$
- State space Ω and action spaces I and J of Player 1 and Player 2 are independent of h.
- Payoff function g_h of Player 1 and transition probability P_h depend on h.

Two different models

- Model S (Sorin): $P_h(\omega'|i,j,\omega) = e^{hq(i,j)}(\omega,\omega')$.
- Payoff at n-th stage is $\int_{(n-1)h}^{nh} \lambda e^{-\lambda t} g_t dt$.
- Total payoff is $\int_0^\infty \lambda e^{-\lambda t} g_t dt$.
- The value is $v^S_{\lambda,h}$.

- Model N (Neyman): $P_h(\omega'|i,j,\omega) = [Id + hq(i,j)](\omega,\omega')$.
- Payoff at n-th stage is $\lambda h(1-\lambda h)^{n-1} g_n$.
- Total payoff is $\lambda h \sum_{k=1}^\infty (1-\lambda h)^{k-1} g_k$.
- The value is $v^N_{\lambda,h}$.

- Sorin 2018: If $h \to 0$ then $v^S_{\lambda,h}$ approaches the λ-discounted value of the continuous-time game.
- Novikov 2025: If $h \to 0$ then $v^N_{\lambda,h}$ approaches the λ-discounted value of the continuous-time game.

Papers about games with stage duration

Model S:

- “Limit Value of Dynamic Zero-Sum Games with Vanishing Stage Duration” by Sylvain Sorin (2018);
- “Markov Games with Frequent Actions and Incomplete Information—The Limit Case” by Pierre Cardaliaguet, Catherine Rainer, Dinah Rosenberg, Nicolas Vieille (2016);
- “Continuous-time limit of dynamic games with incomplete information and a more informed player” by Fabien Gensbittel (2016).

Model N:

- “Stochastic games with short-stage duration” by Abraham Neyman (2013);
- “Operator approach to values of stochastic games with varying stage duration” by Sylvain Sorin and Guillaume Vigeral (2016).

Stochastic games with stage duration (general payoff)

- We can consider more general payoffs.
- Let $k : [0, T] \to \mathbb{R}$ be a nonincreasing continuous positive function with $\int_0^T k(t)dt = 1$.
- We can define analogues of Model S and Model N for this more general case, in which the players can act only at times $0, h, 2h, \ldots$, and the payoff is computed with $\int_{nh}^{(n+1)h} k(t) g_t dt$.
- For such a case, we can define the values $v^S_{k,h}$ (for Model S) and $v^N_{k,h}$ (for Model N).
- Sorin 2018 studied the limit $\lim_{h\to 0} v^S_{k,h}$.
- In my article [1], I studied the limit $\lim_{h\to 0} v^N_{k,h}$.

[1] I.N., Zero-Sum State-Blind Stochastic Games with Vanishing Stage Duration. Dyn Games Appl 15, 1094–1115 (2025).

The limit as $h \to 0$

Theorem (I.N., 2025)

The limit $\lim_{h\to 0} v^N_{k,h}$ *exists and is a unique viscosity solution of*

$$0 = \frac{d}{dt} v(t,\omega) + \mathtt{Val}_{I\times J}[k(t) g(i,j,\omega) + \langle q(i,j)(\omega,\cdot)\,, v(t,\cdot)\rangle],$$

with the boundary condition $v(T,\omega) = 0$ *for all* ω.

- $\langle f(\cdot), g(\cdot)\rangle = \sum_{x\in X} f(x) g(x)$;
- $\mathtt{Val}_{I\times J}(G)$ *is the value of the one-shot game* G *with action spaces* I, J.

The analogous result is true for $\lim_{h\to 0} v^S_{k,h}$ (Sorin 2018).
Proof: we use Shapley operators to check that $\|v^S_{k,h} - v^N_{k,h}\|_\infty \to 0$ as $h \to 0$.

Table of contents

Model S with signals

- State space Ω, signaling function f, and action spaces I and J of Player 1 and Player 2 are independent of h.

- Model S: $P_h(\omega'|i,j,\omega) = e^{hq(i,j)}(\omega,\omega')$.
- Payoff at n-th stage is $\int_{(n-1)h}^{nh} e^{-\lambda t} g_t dt$.
- The value is $v_{\lambda,h}^S$.

- For the state-blind case, the limit $\lim_{h\to 0} v_{\lambda,h}^S$ was studied in Sorin 2018.

Model N with signals

- State space Ω, signaling function f, and action spaces I and J of Player 1 and Player 2 are independent of h.

- Model N: $P_h(\omega'|i,j,\omega) = [Id + hq(i,j)](\omega,\omega')$.
- Payoff at n-th stage is $\lambda h(1-\lambda h)^{n-1} g_n$.
- Total payoff is $\lambda h \sum_{k=1}^{\infty} (1-\lambda h)^{k-1} g_k$.
- The value is $v^N_{\lambda,h}$.

- My article [1] studies the limit $\lim_{h\to 0} v^N_{k,h}$ in the state-blind case.

[1] I.N., Zero-Sum State-Blind Stochastic Games with Vanishing Stage Duration. Dyn Games Appl 15, 1094–1115 (2025).

The limit as $h \to 0$ (state-blind case)

Theorem (I.N. 2025)

The limit $\lim_{h\to 0} v^N_{\lambda,h}$ *exists and is a unique viscosity solution of*

$$\lambda v(p) = \mathtt{Val}_{I\times J}[\lambda g(i,j,p) + \langle p * q(i,j), \nabla v(p)\rangle],$$

where

$$(p * q(i,j))\,(\omega) := \sum_{\omega'\in\Omega} p(\omega')\cdot q(i,j)(\omega',\omega);$$

$$\langle f(\cdot), g(\cdot)\rangle := \sum_{x\in X} f(x)g(x).$$

A similar result is proven for the general case with the total payoff $\int_0^T k(t)g_t dt$.

Sketch of the proof

- Here, we cannot evaluate $\|v^S_{k,h} - v^N_{k,h}\|_\infty$ as $h \to 0$.
- The proof is similar to the proof of the analogous result for Model S in the paper of Sylvain Sorin (2018) "Limit Value of Dynamic Zero-Sum Games with Vanishing Stage Duration";
- Namely, we consider the family $\{v_{h,\lambda}(p)\}_{h\in(0,1]}$. We prove that it is equilipschitz-continuous and equibounded;
- Hence by the Arzelà–Ascoli theorem the limit $\lim_{h\to 0} v_{h,\lambda}(p)$ has at least one accumulation point;
- Afterwards we write the Shapley equation to prove that each accumulation point is a viscosity solution of the above differential equation;
- It can be proven that this differential equation has a unique solution.

Table of contents

The limit as $\lambda \to 0$

- $v_\lambda(\omega) = \sup_\sigma \inf_\tau E^\omega_{\sigma,\tau}\left(\lambda \sum_{i=1}^\infty (1-\lambda)^{i-1} g_i\right)$;
- One can ask: what happens if players become more and more patient?
- Mathematically, it means that $\lambda \to 0$;
- Thus, one is interested in the limit $\lim_{\lambda \to 0} v_\lambda(\omega)$;
- This limit is called the *asymptotic value*.
- Asymptotic value always exists in the finite framework (Bewley & Kohlberg 1976), but may fail to exist in a more general setting (Vigeral 2013, Ziliotto 2016).

The limit as $\lambda \to 0$ in games with public signals

- Discounted case in Model N: total payoff $\lambda h \sum_{k=1}^{\infty} (1-\lambda h)^{k-1} g_k$ and value $v^N_{\lambda,h} = v_{\lambda,h}$.

- We consider $\lim\limits_{\lambda \to 0} v_{h,\lambda}$;
- Even in the finite setting, $\lim\limits_{\lambda \to 0} v_{h,\lambda}$ may not exist;
- First example of nonexistence is in Ziliotto 2016. A similar counterexample is in Ziliotto & Renault 2020.

The limit as $\lambda \to 0$ in games with stage duration (full observation)

Proposition (Sorin & Vigeral 2016)

In a game with full state observation, $\lim_{\lambda\to 0}\lim_{h\to 0} v_{h,\lambda}$ *exists if and only if* $\lim_{\lambda\to 0} v_{1,\lambda}$ *exists, and in the case of existence we have* $\lim_{\lambda\to 0}\lim_{h\to 0} v_{h,\lambda} = \lim_{\lambda\to 0} v_{1,\lambda}$.

- $\lim_{\lambda\to 0} v_{1,\lambda}$ should be considered as the asymptotic value of the discrete-time stochastic game, whereas $\lim_{\lambda\to 0}\lim_{h\to 0} v_{h,\lambda}$ should be considered as the asymptotic value of the analogous continuous-time game.
- The analogous result is not true in the case with public signals. My article [2] is dedicated to the study of the asymptotic value in this framework.

[2] I.N., Asymptotic Value in Zero-Sum Stochastic Games with Vanishing Stage Duration and Public Signals. Preprint (2024).

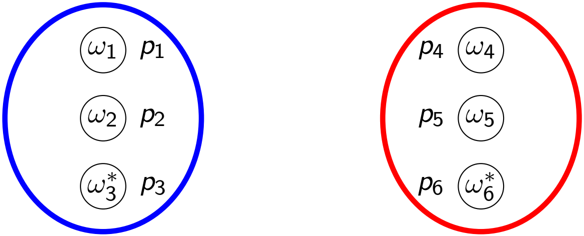
The limit as $\lambda \to 0$ in games with stage duration (public signals 1)

Theorem (I.N., 2024)

There is a stochastic game G (from Ziliotto & Renault 2020) with public signals in which the limit $\lim_{\lambda\to 0}\lim_{h\to 0} v_{h,\lambda}$ *exists, but the limit* $\lim_{\lambda\to 0} v_{1,\lambda}$ *does not exist.*

- the asymptotic value in continuous time ($\lim_{\lambda\to 0}\lim_{h\to 0} v_{h,\lambda}$) exists, but the asymptotic value in discrete time ($\lim_{\lambda\to 0} v_{1,\lambda}$) does not.
- This is impossible in the case with full observation of the state.

The game



Signal MINUS

Payoff -1

Player 1's actions: T, B, Q

Player 2's actions: L, R

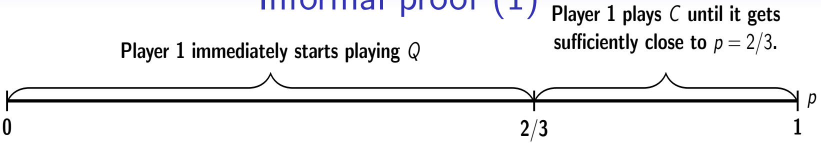
Signal PLUS

Payoff $+1$

Player 1's actions: T, M, B

Player 2's actions: L, M, R, Q

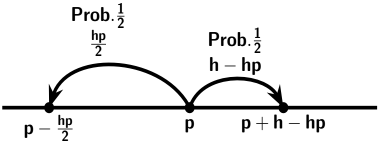
Informal proof (1)



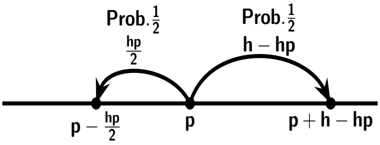
Player 1 plays C until it gets sufficiently close to $p = 2/3$.

Player 1 immediately starts playing Q

Figure: Continuous case (i.e. $h \approx 0$) with small λ. If his belief p that the current state is ω_2 is less than $2/3$, then Player 1 should immediately start playing Q. Otherwise, his belief $\widetilde{p}$ will start to increase until it becomes $\widetilde{p} = 2/3$, which is bad for Player 1. If his belief p that the current state is ω_2 is bigger than $2/3$, then Player 1 can very quickly decrease his belief $\widetilde{p}$ until it becomes $\widetilde{p} \approx 2/3$, which is good for him.



(a) $p > 2/3$ and Player 1 plays not Q. $E(\widetilde{p} - p) = \frac{1}{2}(h - hp) + \frac{1}{2} \cdot \frac{-hp}{2} = \frac{h}{4}(2 - 3p) < 0$.

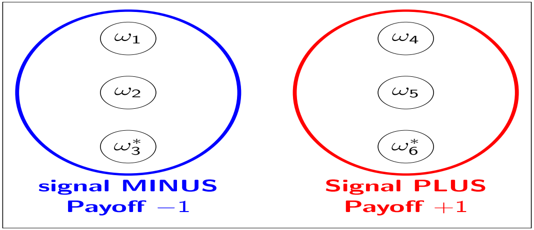


(b) $p < 2/3$ and Player 1 plays not Q. $E(\widetilde{p} - p) = \frac{1}{2}(h - hp) + \frac{1}{2} \cdot \frac{-hp}{2} = \frac{h}{4}(2 - 3p) > 0$.

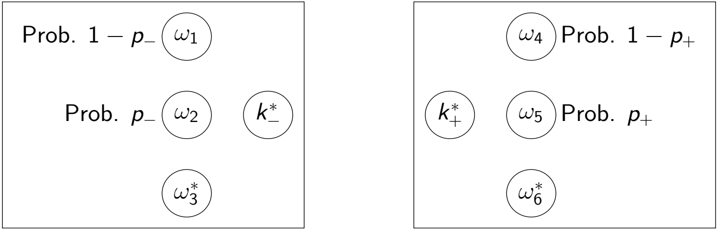
Informal proof (2)

- Thus there is a threshold $p = 2/3$ which Player 1 cannot cross;
- So, the state is going to quickly get absorbed with prob. $2/3$;
- Similarly, there is a threshold $p = 3/4$ which Player 2 cannot cross;
- So, the state is going to get absorbed with prob. $3/4$;
- In the discrete case, the state is absorbed with a very low probability. The oscillation happens because there is a lot of back and forth between the two nonabsorbing states.
- Thus there is no oscillation as $\lambda \to 0$.

Formal proof (1)



↓ Game G with two public signals ↓



(a) State-blind "half-game" $G^-(k_-)$, where $k_- \in [-1, 1]$.

(b) State-blind "half-game" $G^+(k_+)$, where $k_+ \in [-1, 1]$.

Formal proof (2)

- We only need to find the values $v^-_{h,\lambda}(k,p)$ and $v^+_{h,\lambda}(k,p)$ of these two "half-games"!
- In this case we can deduce $\lim_{\lambda\to 0}\lim_{h\to 0} v_{h,\lambda}$ for initial states ω_2 and ω_5 by solving a system of two equations with variables k_- and k_+. We have $\begin{cases} v_{h,\lambda}(\omega_2) = v^-_{h,\lambda}(v_{h,\lambda}(\omega_5),\omega_2) \\ v_{h,\lambda}(\omega_5) = v^+_{h,\lambda}(v_{h,\lambda}(\omega_2),\omega_5) \end{cases}$.
- Later we can find $\lim_{\lambda\to 0}\lim_{h\to 0} v_{h,\lambda}(p)$ for any initial p by replacing k_- or k_+ with values that were just found.

Formal proof (3)

- The main question: how to find $\lim_{h\to 0} v^-_{h,\lambda}(k,p)$?
- We know that it is a unique solution of

$$\lambda v(p) = \mathtt{val}_{I\times J}[\lambda g(i,j,p) + \langle p * q(i,j), \nabla v(p)\rangle];$$

- It is not clear how to find the solution of this equation;
- We can heuristically deduce the solution and verify that it satisfies the above equation.

Open question

Theorem (I.N., 2024)

There is a stochastic game G with public signals in which the limit $\lim_{\lambda\to 0}\lim_{h\to 0} v_{h,\lambda}$ exists, but the limit $\lim_{\lambda\to 0} v_{1,\lambda}$ does not exist.

Open question: For the considered game G, can we say that

1. For any fixed $h \in (0,1]$, the limit $\lim\limits_{\lambda\to 0} v_{h,\lambda}$ does not exist?
2. We have $\left|\limsup\limits_{\lambda\to 0} v_{h,\lambda}(p) - \liminf\limits_{\lambda\to 0} v_{h,\lambda}(p)\right| \to 0$ as $h \to 0$, uniformly in p?

Other open questions

Question 1: Can we say that for any finite state-blind stochastic game, the limit $\lim_{\lambda \to 0} \lim_{h \to 0} v_{h,\lambda}$ exists?

Question 2: Can we say that for any finite stochastic game with public signals, the existence of $\lim_{\lambda \to 0} v_{1,\lambda}$ implies the existence of $\lim_{\lambda \to 0} \lim_{h \to 0} v_{h,\lambda}$?

Question 3: Can we say something about the one-player state-blind case?

This is all.

Thank you!