Zero-Sum State-Blind Stochastic Games with Vanishing Stage Duration

Ivan Novikov

Zero-Sum State-Blind Stochastic Games with Vanishing Stage Duration

Ivan Novikov, Université Paris Dauphine (CEREMADE)

Abstract

In stochastic games with stage duration h, players act at times $h, 2h, 3h, \ldots$, and the payoff and leaving probabilities are proportional to h. When h tends to 0, such games approximate games in continuous time. The behavior of the values when h tends to 0 was already studied in the case of stochastic games with perfect observation of the state.

We consider the same question for the case of state-blind stochastic games. Our main result states that the value of any state-blind stochastic game with stage duration h converges as h tends to 0 to a unique viscosity solution of a partial differential equation.

Keywords: Stochastic games, Zero-sum stochastic games, State-blind stochastic games, Shapley operator, Varying stage duration, Viscosity solution, Continuous-time Markov games

Contents

Notation:

- $\mathbb{N}^*$ is the set of all positive integers;
- $\mathbb{R}_+ := \{x : x \in \mathbb{R} \text{ and } x \geq 0\}$;
- If $f(x)$ is a function defined on a set X, then $\|f(x)\|_\infty := \sup_{x\in X} |f(x)|$;
- If $f(x)$ is a function defined on a finite set X, then $\|f(x)\|_1 := \sum_{x\in X} |f(x)|$;
- If C is a finite set, then $\Delta(C)$ is the set of probability measures on C;
- If X is a finite set, and $f, g : X \to \mathbb{R}$ are two functions, then $\langle f(\cdot), g(\cdot)\rangle := \sum_{x\in X} f(x)g(x)$;
  If $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in X$, then $\langle x_1, x_2\rangle := \sum_{i=1}^n x_i y_i$;
- If X is a finite set, $\zeta \in \Delta(X)$, and μ is a $|X| \times |X|$ matrix, then for any $x \in X$

$$(\zeta * \mu)(x) := \sum_{x'\in X} \zeta(x') \cdot \mu_{x'x};$$

- If I, J are sets and $g : I \times J \to \mathbb{R}$ is a function, then

$$\mathtt{Val}_{I\times J}[g(i,j)] := \sup_{x\in\Delta(I)} \inf_{y\in\Delta(J)} \left(\int_{I\times J} g(i,j)\, dx(i) \otimes dy(j)\right) = \inf_{y\in\Delta(J)} \sup_{x\in\Delta(I)} \left(\int_{I\times J} g(i,j)\, dx(i) \otimes dy(j)\right).$$

  I.e. $\mathtt{Val}_{I\times J}[g(i,j)]$ is the value of a one-shot zero-sum game with action spaces I, J and with payoff function g. Note that we assume that the sets I, J, g are such that the value exists.

1 Introduction

Zero-sum stochastic games with perfect observation of the state were first defined in [Sha53]. Such a game is played in discrete time as follows. At the start of each stage, player 1 and player 2 observe the current state and remember previous players' actions, and choose their mixed actions, after which player 1 receives some payoff, depending on players' actions and the current state. Player 2 receives the opposite of this payoff. The next stage is chosen according to some probability law, depending on players' actions and the current state. For a fixed $\lambda \in (0, 1]$, player 1 wants to maximize the λ-discounted total payoff $E\left(\lambda \sum_{i=1}^{\infty} (1-\lambda)^{i-1} g_i\right)$, where g_i is i-th stage payoff; player 2 wants to minimize it. Under some weak conditions maxmin and minmax coincide, and the resulting quantity is called the value denoted by v_λ.

An interesting question is what happens if the players are patient, i.e., what happens with v_λ when λ tends to 0. Convergence of v_λ as λ tends to 0 was proved in particular for finite games in [BK76].

A similar model of stochastic games, in which players cannot observe the current state, is called state-blind stochastic games. In such games, the players can observe only the initial probability distribution on the states and the previous actions. One can define the value v_λ in the same way as above. One can also consider what happens with v_λ if λ tends to 0. And for this model, it is possible that even for finite games v_λ diverges when λ tends to 0. See [Zil16] for an example of such a game.

An analogous model in continuous time is **continuous-time Markov games**, in which players are allowed to choose actions at any moment of time. Players' actions at time t may depend only on the current state and on t (with some technical measurability conditions). Such games were introduced in [Zac64], and later studied in many other papers, for example, in [GHL03] and [GHL05].

In a continuous-time Markov game, player 1 tries to maximize the λ-discounted total payoff $E(\int_0^{+\infty} \lambda e^{-\lambda t} g_t dt)$, where g_t is instantaneous payoff at time t. Player 2 tries to minimize it. Just as in discrete-time games, under some weak conditions maxmin and minmax coincide, and the resulting quantity is called the value. One can also consider a more general total payoff $E(\int_0^{+\infty} k(t) g_t dt)$, where k is a nonincreasing continuous function.

The article [Sor17] considers discretizations of continuous-time Markov games. In a discretization, players can act only at times $t_1, t_2, t_3, \ldots$ (and not at any moment of time, like in a usual continuous-time Markov game), and the state variable still changes as in continuous-time model. [Sor17] considers

both the case of perfect observation of the state and the state-blind case. For each of these cases it is proved that if $\sup\{t_{i+1}-t_i\}$ tends to 0, then the value of a discretization of a continuous-time Markov game converges to a unique viscosity solution of a differential equation.

Zero-sum stochastic games with perfect observation of the state and with stage duration were introduced in [Ney13]. Starting from a stochastic game Γ (with perfect observation of the state) with stage duration 1, Neyman considers a family Γ_h of stochastic games in which players act at time $h, 2h, 3h, \dots$, and the payoffs and leaving probabilities are normalized at each stage, that is, they are proportional to h. This gives the value $v_{h,\lambda}$ depending both on the stage duration h and the discount factor λ. A particular case of interest is when h is small, which approximates a game played in continuous time with λ-discounted payoff $\int_0^{+\infty} \lambda e^{-\lambda t} g_t dt$.

In [Ney13], Neyman considered the asymptotics of $v_{h,\lambda}$ when either h or λ tends to 0. Among other things, it was proved that when h tends to 0, the value of a finite game with stage duration tends to a unique solution of a functional equation. A corollary of this result states that the limit value (when the discount factor λ tends to 0) of the game with stage duration 1 always coincide with the limit value of the game with vanishing stage duration (and one exists if and only if another exists). Afterwards, the article [SV16] generalized some of the results from [Ney13] and obtained new results; it considers the case in which the state and action spaces may be compact, and stage durations h_n may depend on the stage number n. In our paper, we also assume that stage duration may depend on the stage number.

We introduce a more general total payoff for a game with stage duration, which, when $\sup h_i$ is small, approximates the continuous-time game with total payoff $\int_0^{+\infty} k(t) g_t dt$, where k is a nonincreasing continuous function. Our **Theorem 1** states that the value of such a stochastic game with stage duration converges uniformly, as $\sup h_i$ tends to 0, to a unique viscosity solution of a differential equation. This theorem generalizes an already known result for discounted games ([Ney13, Theorem 1], [SV16, Corollary 7.1]). The proof is based on a result from [Sor17].

The study of **zero-sum state-blind stochastic games with stage duration** is the main goal of this article. In §4.2, we give a natural definition of such games. In this setting, we show that many asymptotic results from [Ney13, SV16] do not hold anymore. In §4.3, we consider Example 2 of a state-blind stochastic game, showing in particular that the limit value (when the discount factor λ tends to 0) of the game with stage duration 1 *does not* coincide with the limit value of the game with vanishing stage duration. As was said before, such a situation is impossible in the case of stochastic games with perfect observation of the state.

Our main result, **Theorem 2**, states that when $\sup h_i$ tends to 0, the value of a state-blind stochastic game with stage duration converges uniformly to a unique viscosity solution of a partial differential equation. Thus Theorem 2 is an analogue of Theorem 1 for state-blind stochastic games. In the particular case of discounted games the equation is autonomous (does not depend on t), see Corollary 3. The proof of Theorem 2 has the same architecture as a similar result in [Sor17].

As in the case of stochastic games with perfect observation of the state, there is a connection between continuous-time Markov games and state-blind stochastic games with small stage duration.

2 Organization of the paper

In §3.1, we give all the necessary information about stochastic games with perfect observation of the state. In §3.2, we give the Shapley operator of a game, which gives a functional equation that allows one to find a value of a stochastic game. In §3.3, we recall the definition of stochastic games with perfect observation of the state and with stage duration, discuss it in details, and prove some technical results. In §3.4, we state some known results about such games. In §3.5, we state Theorem 1. The proof of this result is given in §7.

In §4.1, we introduce and discuss the model of state-blind stochastic games. In §4.2, we introduce

state-blind stochastic games with stage duration. In §4.3, we consider an example of a state-blind stochastic game with stage duration. In §4.4, we state Theorem 2, which proof is given in §5.

In §6, we discuss the connection between games with stage duration and continuous-time games. In §6.1, we discuss it in the case of perfect observation of the state, and in §6.2, we discuss it in the state-blind case.

In §8, we give some final comments.

3 Games with stage duration (perfect observation of the state)

3.1 Zero-sum stochastic games with perfect observation of the state

In this section, we introduce all the necessary notions from the theory of zero-sum stochastic games with perfect observation of the state. This section partially based on books [MSZ15], [LRS19], and [Sor02].

(Minor generalization of a construction from [Sha53]). A *zero-sum stochastic game with perfect observation of the state* is a 5-tuple $(\Omega, I, J, \{g_m\}_{m\in\mathbb{N}^*}, \{P_m\}_{m\in\mathbb{N}^*})$, where:

- Ω is a finite non-empty set of states;
- I is a finite non-empty set of actions of player 1;
- J is a finite non-empty set of actions of player 2;
- $g_k : I \times J \times \Omega \to \mathbb{R}$ is the k-th stage payoff function of player 1;
- $P_k : I \times J \to \{\text{row-stochastic matrices } |\Omega| \times |\Omega|\}$ is a transition probability function at the k-th stage.

Recall that a matrix $A = (a_{ij})$ is called row-stochastic if $a_{ij} \geq 0$ for all i, j, and $\sum_j a_{ij} = 1$ for any fixed i.

In this section, we are going to call such games "stochastic games", without specifying that there is a perfect observation of the state. Starting from the next section, we are going to specify it. We denote by $P_k(i,j)(\omega_a, \omega_b)$ the (ω_a, ω_b)-th element of the matrix $P_k(i,j)$.
A stochastic game $(\Omega, I, J, \{g_m\}_{m\in\mathbb{N}^*}, \{P_m\}_{m\in\mathbb{N}^*})$ proceeds in stages as follows. The initial state ω_1 is known to the players. At each stage $n \in \mathbb{N}^*$:

1. The players observe the current state ω_n;
2. Players choose mixed actions. Player 1 chooses $x_n \in \Delta(I)$ and player 2 chooses $y_n \in \Delta(J)$;
3. An action $i_n \in I$ of player 1 (respectively $j_n \in J$ of player 2) is chosen according to the probability measure $x_n \in \Delta(I)$ (respectivey $y_n \in \Delta(J)$);
4. Player 1 obtains a payoff $g_n = g_n(i_n, j_n, \omega_n)$, while player 2 obtains payoff $-g_n$. The new state w_{n+1} is chosen according to the probability law $P_n = P_n(i_n, j_n)(\omega_n, \cdot)$.

The above description of the game is known to the players.

0-players stochastic games also are known as Markov chains, and 1-player stochastic games are known as Markov decision processes.

Remark 1 (Games with infinite state space)**.** In the construction above, we assumed that the state space Ω is finite. However, this is not always the case in this paper. If Ω is not finite, then instead of transition probability functions $P_k : I \times J \to \{\text{row-stochastic matrices } |\Omega| \times |\Omega|\}$ we are going to consider $\widetilde{P}_k : I \times J \to \Omega \times \Omega$ such that:

1. For each $i \in I, j \in J, \omega_1, \omega_2 \in \Omega$ we have $\widetilde{P}_k(i,j)(\omega_1, \omega_2) \geq 0$;
2. For each fixed i, j, ω_1 there is only finite number of ω_2 such that $\widetilde{P}_k(i,j)(\omega_1, \omega_2) > 0$;
3. For each fixed i, j, ω_1 we have $\sum_{\omega_2\in\Omega} \widetilde{P}_k(i,j)(\omega_1, \omega_2) = 1$.

$\widetilde{P}_k(i,j)(\omega_1,\omega_2)$ is still the probability to change the state from ω_1 to ω_2, if at the current stage the players' action profile is (i,j). So, we will still call $\widetilde{P}_k(i,j)$ the transition probability functions. Note that if the state at the current stage is ω_1, there is only finite number of possible states at the next stage. In the general model of stochastic games, this is not always the case, but we will need only this particular case.

Remark 2. In the original construction given by Lloyd Shapley, the payoff functions g_m and the transition probability functions P_m do not depend on the stage number. However, we need the dependency on the stage number.

A history of length $t \in \mathbb{N}$ for the stochastic game $(\Omega, I, J, \{g_m\}_{m\in\mathbb{N}^*}, \{P_m\}_{m\in\mathbb{N}^*})$ is $(\omega_1, i_1, j_1, \omega_2, i_2, j_2, \ldots, \omega_{t-1}, i_{t-1}, j_{t-1}, \omega_t)$. The set of all histories of length t is $H_t := \Omega\times(I\times J\times\Omega)^{t-1}$. A *(behavior) strategy* of player 1 (respectively player 2) is a function $\sigma : \bigcup_{t\geq 1} H_t \to \Delta(I)$ (respectively $\tau : \bigcup_{t\geq 1} H_t \to \Delta(J)$). Players' strategies induce probability distribution on the set $\Omega\times(I\times J\times\Omega)^{\mathbb{N}^*}$. (Indeed, strategies induce a probability distribution on the set H_1, then on the set H_2, etc. By Kolmogorov extension theorem, this probability can be extended in a unique way to the set $\Omega\times(I\times J\times\Omega)^{\mathbb{N}^*}$). In particular, given starting state $\omega\in\Omega$, strategies $\sigma : \bigcup_{t\geq 1} H_t \to \Delta(I), \tau : \bigcup_{t\geq 1} H_t \to \Delta(J)$, and induced by them probability distribution $P^{\omega}_{\sigma,\tau}$ on $\Omega\times(I\times J\times\Omega)^{\mathbb{N}^*}$, we can consider an expectation $E^{\omega}_{\sigma,\tau}$ of a random variable on $\bigcup_{t\geq 1} H_t$.

Now, we need to choose how to compute a total payoff function.

Definition 1 (Stochastic game $\Gamma(\{b_m\}_{m\in\mathbb{N}^*})$ with payoff $\sum_{m=1}^{\infty} b_m g_m$)**.** Fix a sequence $\{b_m\}_{m\in\mathbb{N}^*}$ with $b_m \geq 0$ $(m\in\mathbb{N}^*)$ and with $0 < \sum_{m=1}^{\infty} b_m < \infty$. Fix also a stochastic game $\Gamma = (\Omega, I, J, \{g_m\}_{m\in\mathbb{N}^*}, \{P_m\}_{m\in\mathbb{N}^*})$. The un-normalized payoff of stochastic game $\Gamma(\{b_m\}_{m\in\mathbb{N}^*})$ with payoff $\sum_{m=1}^{\infty} b_m g_m$ depends on a strategy profile (σ,τ) and initial state ω, and is equal to

$$E^{\omega}_{\sigma,\tau}\left(\sum\nolimits_{i=1}^{\infty} b_i g_i\right).$$

The normalized payoff of stochastic game $\Gamma(\{b_m\}_{m\in\mathbb{N}^*})$ with payoff $\sum_{m=1}^{\infty} b_m g_m$ is equal to

$$E^{\omega}_{\sigma,\tau}\left(\frac{1}{\sum\limits_{m=1}^{\infty} b_m}\sum_{i=1}^{\infty} b_i g_i\right).$$

Some particular cases of the above definitions have its own name. For a fixed $T\in\mathbb{N}^*$, the *repeated* T *times game* Γ^T is obtained if we set in the above definition $b_m = 1$ for $m = 1, 2, \ldots, T$, and $b_m = 0$ for $m \geq T+1$. For a fixed $\lambda\in(0,1)$, the λ-*discounted*[1] *game* Γ^{λ} is obtained if we set in the above definition $b_m = (1-\lambda)^{m-1}$ for all $m\in\mathbb{N}^*$.

Analogously to single-shot zero-sum games, we may define the value and the (ε-)optimal strategies of stochastic games.

Definition 2. Fix a stochastic game $\Gamma = (\Omega, I, J, \{g_m\}_{m\in\mathbb{N}^*}, \{P_m\}_{m\in\mathbb{N}^*})$. The stochastic game $\Gamma(\{b_m\}_{m\in\mathbb{N}^*})$ with payoff $\sum_{m=1}^{\infty} b_m g_m$ is said to have an un-normalized value $V : \Omega\to\mathbb{R}$ if for all $\omega\in\Omega$ we have

$$V(\omega) = \sup_{\sigma}\inf_{\tau} E^{\omega}_{\sigma,\tau}\left(\sum\nolimits_{i=1}^{\infty} b_i g_i\right) = \inf_{\tau}\sup_{\sigma} E^{\omega}_{\sigma,\tau}\left(\sum\nolimits_{i=1}^{\infty} b_i g_i\right).$$

[1]Note that in the given definition the weight of i-th stage is $(1-\lambda)^{i-1}$. In the economic literature, the weight of i-th stage is often λ^{i-1}.

Given the un-normalized value $V(\omega)$, the normalized value is $v(\omega) := \left(\sum_{m=1}^{\infty} b_m\right)^{-1} V(\omega)$.

For any $\varepsilon \geq 0$, a pair of strategies (σ^*, τ^*) is called *ε-optimal* if

$$v(\omega) + \varepsilon \geq E^{\omega}_{\sigma^*,\tau}\left(\left(\sum_{m=1}^{\infty} b_m\right)^{-1} \sum\nolimits_{i=1}^{\infty} b_i g_i\right) \text{ for any strategy } \tau \text{ of player 2;}$$

$$v(\omega) - \varepsilon \leq E^{\omega}_{\sigma,\tau^*}\left(\left(\sum_{m=1}^{\infty} b_m\right)^{-1} \sum\nolimits_{i=1}^{\infty} b_i g_i\right) \text{ for any strategy } \sigma \text{ of player 1.}$$

0-optimal strategies are called *optimal*.

Sometimes instead of the transition probability function P, we consider the *kernel* $q : I \times J \to \{\text{matrices } |\Omega| \times |\Omega|\}$ defined by the expression

$$q(i,j)(\omega,\omega') = \begin{cases} P(i,j)(\omega,\omega') & \text{if } \omega \neq \omega'; \\ P(i,j)(\omega,\omega) - 1 & \text{if } \omega = \omega'. \end{cases}$$

In particular, we are sometimes going to define a stochastic game by using kernels instead of transition probability functions. Note that for any $\omega \in \Omega$ we have $\sum_{\omega' \in \Omega} q(i,j)(\omega,\omega') = 0$.

3.2 Shapley operator

Shapley operator is a useful tool which uses the recursive structure of a stochastic game to find its value.

Fix an initial time $n \in \mathbb{N}^*$. We consider a stochastic game $G_n = (\Omega, I, J, \{g_m\}_{m\in\mathbb{N}^*}, \{P_m\}_{m\in\mathbb{N}^*})$, where $\sum_{k=1}^{\infty} \|g_k\|_\infty < \infty$, with total payoff (depending on a strategy profile (σ,τ) and initial state ω)

$$E^{\omega}_{\sigma,\tau}\left(\sum\nolimits_{i=n}^{\infty} g_i\right).$$

Denote by $v_n(\omega)$ the value of such a game, and suppose that the following assumption holds.

Assumption 1.

1. I and J are finite;
2. Ω is a compact metric space;
3. The function g is bounded;
4. For each fixed $n \in \mathbb{N}^*$ and $i \in I$, the function $(j,\omega) \mapsto \int_\Omega f(\omega) P_n(i,j)(\omega,\omega') d\omega'$ is continuous for any bounded continuous function f on Ω, and for each fixed $n \in \mathbb{N}^*$ and $j \in J$, the function $(i,\omega) \mapsto \int_\Omega f(\omega) P_n(i,j)(\omega,\omega') d\omega'$ is continuous for any bounded continuous function f on Ω.

For a sequence of maps $S_1, S_2, \ldots$ from a Banach space C to itself, for any $z \in C$, we denote $\prod_{i=1}^{\infty} S_i(z) := \lim_{n\to\infty}(S_1 \circ S_2 \circ \cdots \circ S_n(z))$.

Proposition 1 (Shapley operator and its properties)**.** Consider the described above game G_n. Denote by T_n ($n \in \mathbb{N}^*$) the operator

$$\begin{aligned} &\psi_n : \{\text{Continuous functions on } \Omega\} \to \{\text{Continuous functions on } \Omega\}, \\ &f(\omega) \mapsto \mathtt{Val}_{I\times J}[g_n(i,j,\omega) + \langle P_n(i,j)(\omega,\cdot), f(\cdot)\rangle]. \end{aligned}$$

Then:

1. for each $n \in \mathbb{N}^*$ the operator ψ_n is nonexpansive, i.e. $\|\psi_n(f-g)\|_\infty \leq \|f-g\|_\infty$;
2. If $\overline{\psi}_n$ is the operator

$$\overline{\psi}_n : \{\text{Continuous functions on } \Omega\} \to \{\text{Continuous functions on } \Omega\},$$
$$f(\omega) \mapsto \mathtt{Val}_{I\times J}\left[\overline{g}_n(i,j,\omega) + \langle \overline{P}_n(i,j)(\omega,\cdot), f(\cdot)\rangle\right],$$

then

$$\left\|\psi_n(f) - \overline{\psi}_n(f)\right\|_\infty \leq \|g_n - \overline{g}_n\|_\infty + \left\|P_n - \overline{P}_n\right\|_1 \cdot \|f\|_\infty;$$

3. The value $v_n(\omega)$ of the game G_n exists, is unique and is equal to $\prod_{i=n}^{\infty} \psi_i(0)$. (In particular, such a limit is well-defined);
4. We have for any $n \in \mathbb{N}^*$ and any $\omega \in \Omega$

$$v_n(\omega) = (\psi_n v_{n+1})(\omega);$$

5. For any $n \in \mathbb{N}^*$, there exists an optimal Markov strategy in the game G_n.

An analogue of this proposition for λ-discounted games and finitely number of times repeated games is well-known, see, for example, [MSZ15, Theorem IV.3.2, Proposition IV.3.3]. The more general case given here can be proved analogously.

3.3 Zero-sum stochastic games with perfect observation of the state and with stage duration

In this subsection, we are going to introduce games with stage duration. §3.3.1 is a minor generalization of [Ney13, pp. 237–238]. §§3.3.2-3.3.3 generalize [Ney13, p. 240] and give some comments. §3.3.4 contains new information.

Let us fix the notation that will be used during the entire section. Let T be either a positive number or $+\infty$. Let T_∞ be a partition of $[0,T)$; in other words, T_∞ is a strictly increasing sequence $\{t_n\}_{n\in\mathbb{N}^*}$ such that $t_1 = 0$ and $t_n \xrightarrow{n\to\infty} T$. For each given partition T_∞, denote $h_n = t_{n+1} - t_n$ for each $n \in \mathbb{N}^*$. Figure 1 illustrates these notions.

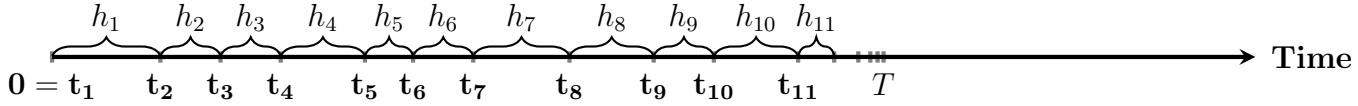


Figure 1: Partition $T_\infty = \{t_n\}_{n\in\mathbb{N}^*}$ of $[0,T)$ for the case when T is finite

Throughout this entire subsection, we fix a stochastic game (Ω, I, J, g, q), where q is the kernel of the game. (Note that payoff function g and kernel q do not depend on the stage number).

3.3.1 Definition

In a stochastic game with n-th stage duration h_n, state space Ω and strategy spaces I and J of player 1 and player 2 are independent of the partition T_∞. Payoff function and kernel depend on T_∞. For $n \in \mathbb{N}^*$, the n-th stage payoff function is $g_n = h_n g$, and the n-th stage kernel function is $q_n = h_n q$. The following definition summarizes this.

Definition 3. Given a stochastic game (Ω, I, J, g, q) and a partition T_∞ of $[0,T)$, the *stochastic game with perfect observation of the state and with n-th stage duration* h_n is the stochastic game $(\Omega, I, J, \{h_n g\}_{n\in\mathbb{N}^*}, \{h_n q\}_{n\in\mathbb{N}^*})$.

We are interested in behavior of games when the duration of each stage is vanishing, i.e., we want to know what happens when $\sup_{i\in\mathbb{N}^*} h_i \to 0$.

Remark 3 (Comparison between stochastic games with stage duration and "usual" stochastic games)**.** We consider a "usual" stochastic game as being a stochastic game with stage duration, in which duration of each stage is 1 (i.e., $T = \infty$ and $t_n = n - 1$ $(\forall n \in \mathbb{N}^*)$).
In a stochastic game with n-th stage duration h_n, the payoff g_n and kernel q_n of n-th stage are proportional to h_n. Informally, we may consider that in such a game players act at times $0, t_1, t_2, \ldots$ See Figures 2 and 3.

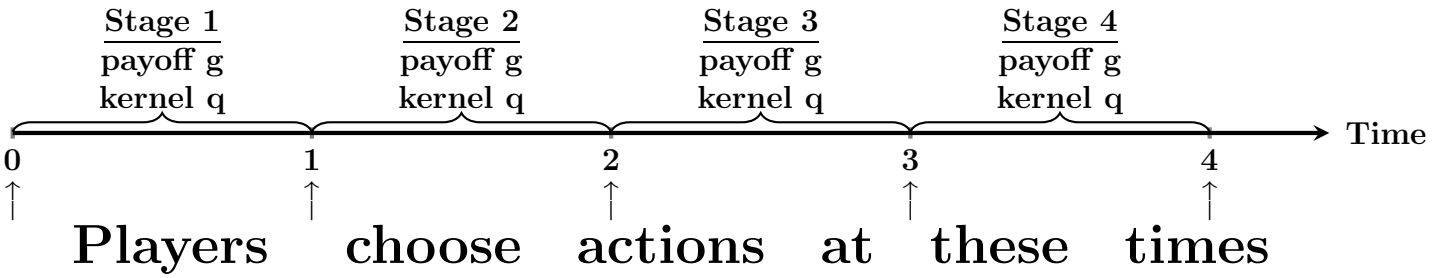


Figure 2: "Usual" stochastic game: duration of each stage is 1

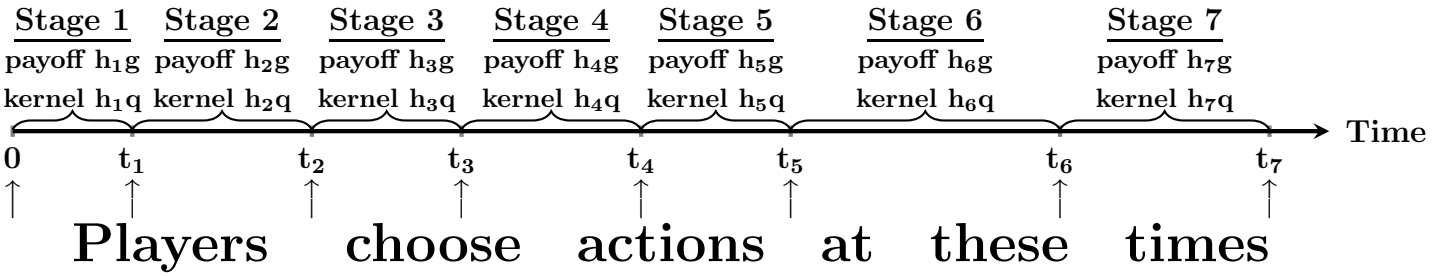


Figure 3: Stochastic game with n-th stage duration h_n: n-th stage payoff and kernel are proportional to h_n

3.3.2 Discounted stochastic games with stage duration

When talking about discounted stochastic games with stage duration, we always assume that $T = +\infty$, so that T_∞ is a partition of $\mathbb{R}_+$.

For each $h \in (0, 1]$, consider a discount factor $\alpha_h \in [0, 1)$. We want to impose some natural condition on α_h, which will allow us to study the value when $\sup_i h_i$ tends to 0. Such a condition is given by the following definition.

Definition 4. A family of h-dependent discount factors α_h is called *admissible* if $\lim\limits_{h \to 0+} \dfrac{\alpha_h}{h}$ exists. The limit is called the *asymptotic discount rate*.

Remark 4 (Why such a condition?)**.** Let us give two reasons why we chose such a family α_h.

1. Consider a stochastic game in which the duration of each stage is h, with payoff $g_h(i, j, \omega) = h$ for all $(i, j, \omega) \in I \times J \times \Omega$. Then the valuation of g_h with discount factor α_h is

$$\sum_{k=1}^{\infty} (1 - \alpha_h)^{k-1} h = \frac{h}{\alpha_h}.$$

   Since we are interested in the behavior of games when h tends to 0, we require the existence of $\lim_{h \to 0+} \frac{\alpha_h}{h}$.

2. Consider again a stochastic game in which the duration of each stage is h, but with any payoff function g. If $\lim_{h\to 0+} \frac{\alpha_h}{h} = \lambda$, then $\alpha_h \underset{h\to 0+}{=} \lambda h + o(h)$. Given a sequence $g_1, g_2, \dots$ of numbers in $\mathbb{R}$, the normalized valuation of this sequence with discount factor α_h is $\approx \sum_{i=1}^{\infty} \lambda h(1-\lambda h)^{i-1} g_i$ (we multiplied the sum by λ because $\sum_{i=1}^{\infty} h(1-\lambda h)^{i-1} = 1/\lambda$). One can see it as a normalized λh-discounted total payoff in the game with n-th stage payoff g_n, so that the discount factor λh is proportional to h, in the same way as the payoff and the kernel functions are proportional to h (in a stochastic game in which each stage has duration h).

Example 1 (Families of admissible discount factors)**.**

1. $\alpha_h = 1 - e^{-\lambda h}$ (Family with asymptotic discount rate $\lambda \in (0, +\infty)$);
2. $\alpha_h = \lambda h$ (Family with asymptotic discount rate $\lambda \in (0, 1]$);
3. $\alpha_h = \begin{cases} 0, & \text{if } h > 1/\lambda; \\ \lambda h, & \text{if } h \leq 1/\lambda. \end{cases}$ (Family with asymptotic discount rate $\lambda \in (0, +\infty)$);
4. $\alpha_h = h^2$ (Family with asymptotic discount rate 0);
5. $\alpha_h = \sqrt{h}$ (Family with asymptotic discount rate $+\infty$).

In what follows, we do not consider families of discount factors with asymptotic discount rate 0 or $+\infty$, since these cases are degenerate and are not interesting. Note that for each fixed asymptotic discount rate $\lambda \in (0, +\infty)$, if h is small enough, then the examples 1.2 and 1.3 above are equivalent.

Now, we define discounted games with stage duration.

Definition 5 (λ-discounted stochastic game with stage duration)**.** Let α_h be an admissible family of discount factors with asymptotic discount rate λ.
λ-discounted stochastic game with n-th stage duration h_n is a stochastic game with payoff

$$\sum_{i=1}^{\infty} \left(\prod_{j=1}^{i-1} (1 - \alpha_{h_i}) \right) h_i g_i.$$

Denote by $V^{\alpha}_{T_\infty,\lambda}$ the un-normalized value of such a game. Denote by $v^{\alpha}_{T_\infty,\lambda} = V^{\alpha}_{T_\infty,\lambda} \Big/ \sum_{i=1}^{\infty} \left(\prod_{j=1}^{i-1} (1 - \alpha_{h_i}) \right) h_i$ the normalized value of such a game.

Note that both $V^{\alpha}_{T_\infty,\lambda}$ and $v^{\alpha}_{T_\infty,\lambda}$ are functions of the initial state.

Remark 5. In the above definition:

1. If $\alpha_h = 1 - e^{-\lambda h}$, then $\sum_{i=1}^{\infty} \left(\prod_{j=1}^{i-1} (1 - \alpha_{h_i}) \right) h_i g_i = \sum_{i=1}^{\infty} \left(\prod_{j=1}^{i-1} \exp\{-\lambda h_j\} \right) h_i g_i = \sum_{i=1}^{\infty} \exp\{-\lambda t_i\} h_i g_i$;
2. If $\alpha_h = \lambda h$, then $\sum_{i=1}^{\infty} \left(\prod_{j=1}^{i-1} (1 - \alpha_{h_i}) \right) h_i g_i = \sum_{i=1}^{\infty} \left(\prod_{j=1}^{i-1} (1 - \lambda h_j) \right) h_i g_i$, and $v_{T_\infty,\lambda} = \lambda V_{T_\infty,\lambda}$ (it follows from the fact that $\sum_{i=1}^{\infty} \left(\prod_{j=1}^{i-1} (1 - \lambda h_j) \right) h_i = \frac{1}{\lambda}$, see Lemma 2(2) for a proof).

The definition of λ-discounted stochastic games with stage duration depends on the family α_h. However, if $\sup_{i\in\mathbb{N}} h_i$ tends to 0, then a choice of α_h does not matter anymore. To prove this we need the following lemma.

Lemma 1. Fix $\lambda \in (0, +\infty)$. Consider two admissible families α_h and β_h of discount factors with asymptotic discount rate λ. Then for all families (parametrized by partitions T_∞ of $\mathbb{R}_+$) of streams $x(T_\infty) = (g_1, g_2, \ldots)$ with $|g_i| \leq Ch_i$ $(i \in \mathbb{N}^*)$, the difference

$$\sum_{i=1}^{\infty} \left(\prod_{j=1}^{i-1} (1 - \alpha_j) \right) g_i - \sum_{i=1}^{\infty} \left(\prod_{j=1}^{i-1} (1 - \beta_j) \right) g_i$$

tends to 0 as $\sup_{i \in \mathbb{N}} h_i$ tends to 0.

Before the proof of Lemma 1, let us first state and prove the following simple lemma.

Lemma 2. Fix $\lambda \in (0, 1)$ and a sequence $H_\infty = \{h_i\}_{i \in \mathbb{N}^*}$ with $h_i \in (0, 1]$. We have:

1. $\prod_{i=1}^{\infty} (1 - \lambda h_i) = 0 \iff \sum_{i=1}^{\infty} h_i = +\infty$;
2. If $\sum_{i=1}^{\infty} h_i = +\infty$, then $\sum_{i=1}^{\infty} \left(\prod_{j=1}^{i-1} (1 - \lambda h_j) \right) h_i = \frac{1}{\lambda}$;
3. For all $n \in \mathbb{N}^*$ we have $\prod_{i=1}^{n} (1 - \lambda h_i) \geq 1 - \lambda \sum_{i=1}^{n} h_i$.

Proof. The first assertion is a standard result from elementary analysis. We prove the second assertion. For each $i \in \mathbb{N}$ denote $k_i = 1 - \lambda h_i$. We have

$$\sum_{i=1}^{\infty} \left(\prod_{j=1}^{i-1} (1 - \lambda h_j) \right) h_i = \left(\frac{1 - k_1}{\lambda} + \frac{k_1 - k_1 k_2}{\lambda} + \frac{k_1 k_2 - k_1 k_2 k_3}{\lambda} + \ldots \right) = \frac{1}{\lambda},$$

where the last equality holds because $\prod_{i=1}^{\infty} k_i = 0$ by the first assertion of the lemma.

We prove the third assertion by induction on $n \in \mathbb{N}^*$. The case $n = 1$ is straightforward to verify. Assume that for $n = k$ the assertion holds, i.e. we have $\prod_{i=1}^{k} (1 - \lambda h_i) \geq 1 - \lambda \sum_{i=1}^{k} h_i$. For $n = k + 1$, we have

$$\prod_{i=1}^{k+1} (1 - \lambda h_i) \geq \left(1 - \lambda \sum_{i=1}^{k} h_i \right) (1 - \lambda h_{k+1}) \geq 1 - \lambda \sum_{i=1}^{k+1} h_i. \qquad \square$$

We are ready to prove Lemma 1.

The proof of Lemma 1. Fix a family of streams $x(H_\infty) = (g_1, g_2, \ldots)$ with $|g_i| \leq Ch_i$ $(i \in \mathbb{N}^*)$, and assume that α_h and β_h are two families of discount factors with asymptotic discount rate λ. In that case we have $1 - \alpha_h = 1 - \lambda h + m(h)$ and $1 - \beta_h = 1 - \lambda h + n(h)$, where $m(h)/h \to 0$ and $n(h)/h \to 0$ as $h \to 0$. Consider the sum

$$\sum_{i=1}^{\infty} \left(\prod_{j=1}^{i-1} (1 - \lambda h_j + k(h_j)) \right) h_i.$$

where $k(h_j)$ is either $m(h_j)$ or $n(h_j)$. Now fix h' such that for all h with $0 < h \leq h'$ we have $0 < \lambda \pm \frac{|m(h)|}{h} < 1$ and $0 < \lambda \pm \frac{|n(h)|}{h} < 1$. For any sequence H_∞ with $0 < h_i \leq h'$ we have by the second assertion of Lemma 2

$$\sum_{i=1}^{\infty} \left(\prod_{j=1}^{i-1} (1 - \lambda h_j + k(h_j)) \right) h_i \leq \sum_{i=1}^{\infty} \left(\prod_{j=1}^{i-1} \left[1 - \left(\lambda - \sup_{j \in \mathbb{N}} \left\{ \frac{|m(h_j)|}{h_j}, \frac{|n(h_j)|}{h_j} \right\} \right) h_j \right] \right) h_i = \frac{1}{\lambda - \sup\limits_{j \in \mathbb{N}} \frac{|k(h_j)|}{h_j}};$$

$$\sum_{i=1}^{\infty} \left(\prod_{j=1}^{i-1} (1 - \lambda h_j + k(h_j)) \right) h_i \geq \sum_{i=1}^{\infty} \left(\prod_{j=1}^{i-1} \left[1 - \left(\lambda + \sup_{j \in \mathbb{N}} \left\{ \frac{|m(h_j)|}{h_j}, \frac{|n(h_j)|}{h_j} \right\} \right) h_j \right] \right) h_i = \frac{1}{\lambda + \sup\limits_{j \in \mathbb{N}} \frac{|k(h_j)|}{h_j}}.$$

Hence $\sum_{i=1}^{\infty}\left(\prod_{j=1}^{i-1}(1-\lambda h_j+k(h_j))\right)h_i \to \frac{1}{\lambda}$ as $\sup_{j\in\mathbb{N}} h_j \to 0$. Now we have

$$\left|\sum_{i=1}^{\infty}\left(\prod_{j=1}^{i-1}\alpha_j\right)g_i-\sum_{i=1}^{\infty}\left(\prod_{j=1}^{i-1}\beta_j\right)g_i\right| \leq C\left(\sum_{i=1}^{\infty}\prod_{j=1}^{i-1}(1-\lambda h_j+m(h_j))h_i-\sum_{i=1}^{\infty}\prod_{j=1}^{i-1}(1-\lambda h_j+n(h_j))h_i\right).$$

By the above discussion we know that

$$\sum_{i=1}^{\infty}\prod_{j=1}^{i-1}(1-\lambda h_j+m(h_j))h_i-\sum_{i=1}^{\infty}\prod_{j=1}^{i-1}(1-\lambda h_j+n(h_j))h_i \xrightarrow{\sup_{j\in\mathbb{N}} h_j\to 0} \frac{1}{\lambda}-\frac{1}{\lambda}=0. \qquad \square$$

Corollary 1. If α_h and β_h are two families of discount factors with asymptotic discount rate λ, then we have:

$$\lim_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=+\infty}} V^{\alpha}_{T_\infty,\lambda} = \lim_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=+\infty}} V^{\beta}_{T_\infty,\lambda}; \tag{1}$$

$$\lim_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=+\infty}} \lambda V^{\alpha}_{T_\infty,\lambda} = \lim_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=+\infty}} v^{\alpha}_{T_\infty,\lambda} = \lim_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=+\infty}} v^{\beta}_{T_\infty,\lambda} = \lim_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=+\infty}} \lambda V^{\beta}_{T_\infty,\lambda}. \tag{2}$$

Proof of Corollary 1. (1) follows from Lemma 1, and (2) follows from (1) and Remark 5. $\square$

The analogue of the above lemma was stated in [Ney13] for the games in which each stage has the same duration.

All of our results will be connected with the behavior of games when $\sup_{i\in\mathbb{N}^*} h_i$ tends to 0, so we are going to write $V_{T_\infty,\lambda}$ instead of $V^{\alpha}_{T_\infty,\lambda}$. Analogously, we write $v_{T_\infty,\lambda}$ instead of $v^{\alpha}_{T_\infty,\lambda}$.

In [Ney13, SV16], only the family $\alpha_h = \lambda h$ was used. In this paper we will use both the family $\alpha_h = \lambda h$ and the family $\alpha_h = 1 - e^{-\lambda h}$.

3.3.3 Why such a definition of discounted stochastic games with stage duration?

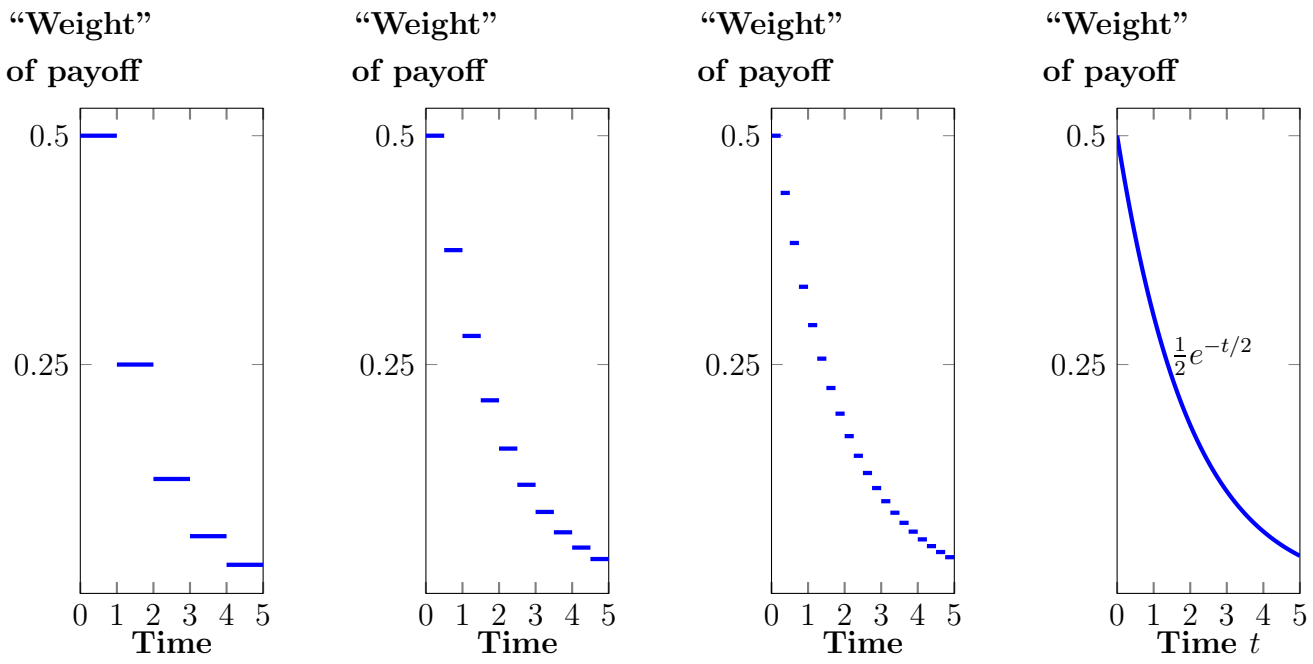
In this subsection, we informally explain why we chose such a definition of discounted games with stage duration. Recall that there is a partition $\{t_i\}_{i\in\mathbb{N}^*}$ of $\mathbb{R}_+$, with $t_{i+1}-t_i=h_i$.

We have chosen α_h in such a way that when all h_i are small, the value of the λ-discounted stochastic game with stage duration is close to the value of the analogous λ-discounted continuous-time game. In a continuous-time game, players can choose actions at any time (under some conditions), and at each time t they receive instantaneous payoff g_t. The total payoff is $\int_0^{\infty}\lambda e^{-\lambda t}g_t dt$. If g_t is continuous in t, then during a period of time $[t_n, t_{n+1}]$, the received payoff in the continuous-time game is

$$\int_{t_n}^{t_{n+1}}\lambda e^{-\lambda t}g_t dt \underset{h_i\to 0}{\approx} g_{t_n}e^{-\lambda t_n}\int_0^{h_n}\lambda e^{-\lambda t}dt = g_{t_n}e^{-\lambda t_n}(1-e^{-\lambda h_n}) \underset{h_i\to 0}{\approx} e^{-\lambda t_n}\lambda h_n g_{t_n} = \lambda h_n g_{t_n}\prod_{i=1}^{n}e^{-h_i}$$

$$\underset{h_i\to 0}{\approx} \lambda h_n g_{t_n}\prod_{i=1}^{n}(1-\lambda h_i).$$

Now, if we assume that $h_n g_{t_n}$ is the n-th stage payoff in a stochastic game with n-th stage duration h_n, then $\lambda h_n g_{t_n}\prod_{i=1}^{n}(1-\lambda h_i)$ is n-th stage payoff in the λ-discounted game. Thus the total payoff of a continuous-time game is close to the total payoff of analogous discrete-time game, and one can prove that indeed, at least in the case of the perfect observation of the state, the values of λ-discounted games with n-th stage duration h_n tend to the value of analogous continuous-time λ-discounted game when $\sup_{i\in\mathbb{N}^*} h_i \to 0$. We discuss it in more details in §6. See Figure 4.



(a) $h = 1$, $\lambda = 1/2$ (b) $h = 1/2$, $\lambda = 1/2$ (c) $h = 1/4$, $\lambda = 1/2$ (d) $h \to 0$, $\lambda = 1/2$

Figure 4: A demonstration of how the weight of each stage in a stochastic game in which each stage has duration h tends to the weight function $\frac{1}{2}e^{-t/2}$ of a continuous-time game. (In (a),(b),(c), the weight $\lambda(1-\lambda h)^{n-1}$ of n-th stage is shown during the time $t \in [(n-1)h, nh]$).

.

3.3.4 Generalized total payoff

Here, we give a total payoff which generalizes the discounted total payoff given above.

Fix a nonincreasing continuous function $k : [0, T] \to \mathbb{R}_+$ with $\int_0^T k(t)dt = 1$.

Definition 6. *Stochastic game (Ω, I, J, g, q) with perfect observation of the state and with n-th stage duration h_n, weight function $k(t)$, initial time $t_n \in T_\infty$, initial state ω, and strategies σ, τ of players*, is a stochastic game $(\Omega, I, J, \{h_n g\}_{n\in\mathbb{N}^*}, \{h_n q\}_{n\in\mathbb{N}^*})$ with total payoff

$$G_{T_\infty,k}(t_n, \omega) := E^{\omega,t_n}_{\sigma,\tau}\left(\sum_{i=n}^{\infty} h_i k(t_i) g_i\right).$$

Denote by $v_{T_\infty,k}(t_n, \omega)$ the value of a game with such a total payoff.

If $t \in [t_n, t_{n+1}]$ and $t = \alpha t_n + (1-\alpha)t_{n+1}$, then we define

$$v_{T_\infty,k}(t, \omega) = \alpha v_{T_\infty,k}(t_n, \omega) + (1-\alpha) v_{T_\infty,k}(t_{n+1}, \omega).$$

Remark 6 (Discounted games)**.** If starting time is 0, $T = +\infty$, and $k(t) = \lambda \exp\{-\lambda t\}$, we obtain the normalized λ-discounted stochastic game with stage duration with $\alpha_h = e^{-\lambda h}$.

Remark 7 (Repeated finitely number of times games)**.** If we take $T \in \mathbb{N}^*$, initial time 0, $k(t) = 1/T$ for $t \in [0, T]$, $h_n = 1$ for $n = 1, \ldots, T$, and $h_n = 0$ for $n > T$, then we receive the case of usual repeated T times games with total payoff $\frac{1}{T}\sum_{i=1}^{T} g_i$. Note that formally such a choice of h_n does not satisfy to the definition, because h_n should always be strictly positive. To make it satisfy to the definition, we may consider any partition T_∞ such that $t_n \to T$, with initial time and $k(t)$ being the same as before. Such a case was considered and studied in [SV16].

Remark 8. Note that $\sum_{i=1}^{\infty} h_i k(t_i)$ is a (left) Riemann sum for the integral $\int_0^T k(t)dt$. Hence for any $\varepsilon > 0$ there is $\delta > 0$ such that for any partition T_∞ with $\sup_{i\in\mathbb{N}^*} h_i < \delta$, we have

$$\left\|\sum_{i=1}^{\infty} h_i k(t_i) g_i\right\|_\infty \leq \left|\sum_{i=1}^{\infty} h_i k(t_i)\right| \|g\|_\infty \leq \left(\int_0^T k(t)dt + \varepsilon\right) \|g\|_\infty = (1+\varepsilon)\|g\|_\infty .$$

Remark 9 (Why such a definition)**.** As in the case with the discounted games, if all h_i are small, then the value of the stochastic game with n-th stage duration h_n, weight function $k(t)$, and initial time $s \in \mathbb{R}_+$, is close to the value of the continuous-time game with total payoff $\int_s^\infty k(t) g_t dt$. See §6.1.

3.4 Known results

We denote by $v_{1,\lambda}$ the value of the λ-discounted stochastic game in which each stage has duration 1. The following two propositions are about discounted games. The following proposition is [SV16, Proposition 7.4], it is a partial generalization of [Ney13, Theorem 1].

Proposition 2. For a fixed stochastic game (Ω, I, J, g, q), there is $K > 0$ such that for any $\lambda \in (0, +\infty)$ and partition T_∞ of $\mathbb{R}_+$, we have

$$\left\| v_{T_\infty,\lambda} - v_{1,\frac{\lambda}{1+\lambda}} \right\| \leq K \sup_{i\in\mathbb{N}^*} h_i.$$

The following proposition is [SV16, Proposition 7.6].

Proposition 3. For a fixed stochastic game (Ω, I, J, g, q) there is $K > 0$ such that for any partition $T_\infty = \{t_n\}_{n\in\mathbb{N}^*}$ of $\mathbb{R}_+$ we have

$$\|v_{T_\infty,\lambda} - v_{1,\lambda}\| \leq K\lambda.$$

Recall that if X is a finite set, and $f, g : X \to \mathbb{R}$ are two functions, then $\langle f(\cdot), g(\cdot)\rangle := \sum_{x\in X} f(x)g(x)$.

Corollary 2. Fix a stochastic game (Ω, I, J, g, q). We have

1. Uniform limit $\lim\limits_{\substack{\sup_{i\in\mathbb{N}^*} h_i \to 0 \\ h_1+h_2+\ldots=+\infty}} v_{T_\infty,\lambda}$ exists and is equal to $v_{1,\frac{\lambda}{1+\lambda}}$;
2. Uniform limit $\lim\limits_{\substack{\sup_{i\in\mathbb{N}^*} h_i \to 0 \\ h_1+h_2+\ldots=+\infty}} v_{T_\infty,\lambda}$ exists and is a unique solution of the equation (in $v(\omega)$)
$$\lambda v(\omega) = \mathtt{Val}_{I\times J}\left[\lambda g(i,j,\omega) + \langle q(i,j)(\omega,\cdot), v(\cdot)\rangle\right]; \tag{3}$$
3. For any partition T_∞ of $\mathbb{R}_+$, the uniform limits $\lim_{\lambda\to 0} v_{T_\infty,\lambda}$ and $\lim_{\lambda\to 0} v_{1,\lambda}$ exist or do not exist simultaneously. In the case of existence, for any two partitions T^1_∞, T^2_∞ of $\mathbb{R}_+$ we have $\lim_{\lambda\to 0} v_{T^1_\infty,\lambda} = \lim_{\lambda\to 0} v_{T^2_\infty,\lambda}$;
4. If the uniform limit $\lim_{\lambda\to 0} v_{1,\lambda}$ exists, then
$$\lim_{\substack{\sup_{i\in\mathbb{N}^*} h_i \to 0 \\ h_1+h_2+\ldots=+\infty}} \left(\lim_{\lambda\to 0} v_{T_\infty,\lambda}\right) = \lim_{\lambda\to 0}\left(\lim_{\substack{\sup_{i\in\mathbb{N}^*} h_i \to 0 \\ h_1+h_2+\ldots=+\infty}} v_{T_\infty,\lambda}\right) = \lim_{\lambda\to 0} v_{1,\lambda}.$$

   In particular, all the above limits exist. If the uniform limit $\lim_{\lambda\to 0} v_{1,\lambda}$ does not exist, then no of the above uniform limits exist.

Proof of Corollary 2. Assertions 1, 3, 4 follow directly from Propositions 2 and 3. Assertion 2 follows from assertion 1 by Proposition 1. □

Afterwards, we are going to see that many of statements of the above corollary do not hold if there is no perfect observation of the state.

Remark 10. When defining stochastic games, we assumed that action sets I, J and a set of states Ω are finite. In this case it is proven in [BK76] that the limit value $\lim_{\lambda\to 0} v_{1,\lambda}$ always exists. However, we can also consider games in which I, J, or Ω are not finite. In this case, the existence of even the pointwise limit is not guaranteed. See [Zil16] or [RZ20] for an example of a game with finite action spaces and compact state space without such a limit, and see [Vig13] or [Zil16, §4.3] for an example of a game with compact action spaces and finite state space without such a limit. See also [SV15] for a construction which can build games without a limit value.

The above remark combined with Corollary 2(4) provides us with examples of games in which the pointwise limit $\lim\limits_{\lambda\to 0}\lim\limits_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=+\infty}} v_{T_\infty,\lambda}$ does not exist.

3.5 A new result

First, we give a definition of a viscosity solution.

Definition 7. A function $u : [0,T)\times\Omega\to\mathbb{R}$ is called a *viscosity solution* of the differential equation

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

if:

1. for any $\mathcal{C}^1$ function $\psi : [0,T)\times\Omega\to\mathbb{R}$ with $u-\psi$ having a strict local maximum at $(\overline{t},\overline{\omega})\in[0,T)\times\Omega$ we have $0\le\frac{d}{dt}\psi(\overline{t},\overline{\omega})+\mathtt{val}_{I\times J}[k(\overline{t})g(i,j,\overline{\omega})+\langle q(i,j)(\overline{\omega},\cdot)\,,\psi(\overline{t},\cdot)\rangle]$;
2. for any $\mathcal{C}^1$ function $\psi : [0,T)\times\Omega\to\mathbb{R}$ with $u-\psi$ having a strict local minimum at $(\overline{t},\overline{\omega})\in[0,T)\times\Omega$ we have $0\ge\frac{d}{dt}\psi(\overline{t},\overline{\omega})+\mathtt{val}_{I\times J}[k(\overline{t})g(i,j,\overline{\omega})+\langle q(i,j)(\overline{\omega},\cdot)\,,\psi(\overline{t},\cdot)\rangle]$.

Remark 11. If $u(t,\omega)$ is a classical solution of the above differential equation, then $u(t,\omega)$ is a viscosity solution of this equation. This is true because if $u-\psi$ has a strict local maximum at $\left(\overline{t},\overline{\omega}\right)$, then we have $\frac{d}{dt}\psi(\overline{t},\overline{\omega})=\frac{d}{dt}u(\overline{t},\overline{\omega})$ and $\nabla\psi(\overline{t},\overline{\omega})=\nabla u(\overline{t},\overline{\omega})$. The same thing is true if $u-\psi$ has a strict local minimum at $\left(\overline{t},\overline{\omega}\right)$.

Now, we give one new result which generalizes Corollary 2(2).

Theorem 1. If (Ω,I,J,g,q) is a stochastic game, then $\lim\limits_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=T}} v_{T_\infty,k}$ exists and is a unique viscosity solution of the differential equation (in $v(t,\omega)$)

$$0=\frac{d}{dt}v(t,\omega)+\mathtt{val}_{I\times J}[k(t)g(i,j,\omega)+\langle q(i,j)(\omega,\cdot)\,,v(t,\cdot)\rangle]. \tag{4}$$

The proof of Theorem 1 is closely related to a model given in [Sor17], so we will give it in §7, after we discuss that model.

Remark 12. For the discounted case $k(t)=\lambda e^{-\lambda t}$, (4) transforms into (3) (make a substitution $v(t,\omega)\mapsto e^{-\lambda t}v(\omega)$).

4 State-blind stochastic games with stage duration

In this section, we consider the games, in which players cannot observe the current state. In §4.1, we give a definition of state-blind stochastic games, and show that each state-blind stochastic game is equivalent to a stochastic game with perfect observation of the state. In §4.2, we give a natural definition of state-blind stochastic games with stage duration. In 4.3, we provide an example that shows the difference between games with and without observation of the state. In §4.4, we give a new result showing that in the case of state-blind stochastic games, the uniform limit $\lim_{\sup_{i\in\mathbb{N}^*} h_i\to 0} v_{T_\infty,k}$ exists and is a unique viscosity solution of a partial differential equation.

4.1 The model of zero-sum state-blind stochastic games

A zero-sum state-blind stochastic game is played in the same way as a stochastic game with perfect observation of the state, but players cannot observe the current state. At the beginning of each stage, players can only observe the actions of players at the previous stages and the initial probability distribution on the states.

We can define strategies, λ-discounted and repeated finitely times games, total payoffs, values, in the same way as in §3.1. In particular, a strategy of player 1 is an indexed by $n \in \mathbb{N}^*$ collection of functions $(i_1, j_1, i_2, j_2, \dots, i_{n-1}, j_{n-1}) \mapsto \Delta(I)$, where $i_m \in I$ and $j_m \in J$. The value is a function of a probability law p_0, according to which the initial state is chosen.

Any state-blind stochastic game is equivalent to a stochastic game with perfect observation of the state. Consider the following construction from [Zil16, §1.3].

Given a zero-sum state-blind stochastic game $G = (\Omega, I, J, \{g_m\}_{m\in\mathbb{N}^*}, \{P_n\}_{n\in\mathbb{N}^*})$, we define the stochastic game $\Gamma(G) = (\Delta(\Omega), I, J, \{g_m^\gamma\}_{m\in\mathbb{N}^*}, \{P_n^\gamma\}_{n\in\mathbb{N}^*})$ with perfect observation of the state. The function $g_n^\gamma : I \times J \times \Delta(\Omega) \to \mathbb{R}$ is defined by $g_n^\gamma(i, j, p) = \sum_{\omega\in\Omega} p(\omega) g_n(i, j, \omega)$.

Now let us define $\{P_n^\gamma\}_{n\in\mathbb{N}^*}$. For each $n \in \mathbb{N}^*, \omega' \in \Omega, p \in \Delta(\Omega), i \in I, j \in J$ denote

$$P_n(i, j)(p, \omega') := \sum\nolimits_{\omega\in\Omega} p(\omega) P_n(i, j)(\omega, \omega').$$

If the current stage is n and players have belief $p \in \Delta(\Omega)$ about the current state, then after playing $(i, j) \in I \times J$ their posterior belief that the current state is ω is equal to

$$P_n(i, j, \alpha)(p, \omega) := \begin{cases} P_n(i, j)(p, \omega) / \left(\sum_{\omega\in\Omega} P_n(i, j)(p, \omega)\right) & \text{, if } \sum_{\omega\in\Omega} P_n(i, j)(p, \omega) \neq 0; \\ 0 & \text{, otherwise.} \end{cases}$$

The function $P_n^\gamma : I \times J \to \Delta(\Omega) \times \Delta(\Omega)$ is defined by

$$P_n^\gamma(i, j)(p, p') = \begin{cases} 1 & \text{, if } p'(\omega) = P_n(i, j, \alpha)(p, \omega) \text{ for all } \omega \in \Omega; \\ 0 & \text{, otherwise.} \end{cases}$$

Similarly, each strategy s in G has an analogous strategy $\Gamma(s)$ in $\Gamma(G)$.

Definition 8. A strategy in a stochastic game with perfect observation of the state is said to be *Markov* if in this strategy the players' mixed actions at each stage depend only on the current stage number and on the current state.
A strategy s in a state-blind stochastic game G is said to be *Markov* if the strategy $\Gamma(s)$ in the stochastic game $\Gamma(G)$ with perfect observation of the state is Markov.

4.2 Zero-sum state-blind stochastic games with stage duration

As before, T is either a positive number or $+\infty$. T_∞ is a partition of $[0, T)$; in other words, T_∞ is a strictly increasing sequence $\{t_n\}_{n\in\mathbb{N}^*}$ such that $t_1 = 0$ and $t_n \xrightarrow{n\to\infty} T$. For each given partition T_∞, denote $h_n = t_{n+1} - t_n$ for each $n \in \mathbb{N}^*$.

Definition 9. Fix a zero-sum state-blind stochastic game (Ω, I, J, g, q), where q is the kernel. The *state-blind stochastic game with n-th stage duration h_n* is the state-blind stochastic game

$$(\Omega, I, J, \{h_m g\}_{m\in\mathbb{N}^*}, \{h_m q\}_{m\in\mathbb{N}^*}).$$

Note that payoff function g and kernel q in a fixed game do not depend on the stage number.

We can define λ-discounted games as in §3.3.2, and we can give a total payoff for games with stage duration as in §3.3.4.

Note that Lemma 1 still holds, so that two families of discounts $\alpha_h = 1 - \lambda h$ and $\beta_h = e^{-\lambda h}$ are still equivalent when $\sup_{i \in \mathbb{N}^*} h_i \to 0$.

Remark 13 (Why state-blind case is more difficult than the case of perfect observation of the states?)**.** If we denote by T_h the Shapley operator of the game in which each stage has duration h, then in the case of perfect observation of the state we have $T_h = hT_1 + (1-h)Id$. This makes studying of such games relatively easy. In the case of state-blind stochastic games, such an equality is not true anymore.

4.3 Example

Now, let us give an example of the state-blind stochastic game with stage duration, which shows that the behavior of state-blind stochastic games is different from the behavior of games with perfect observation of the state. The example is from Guillaume Vigeral (private communication).

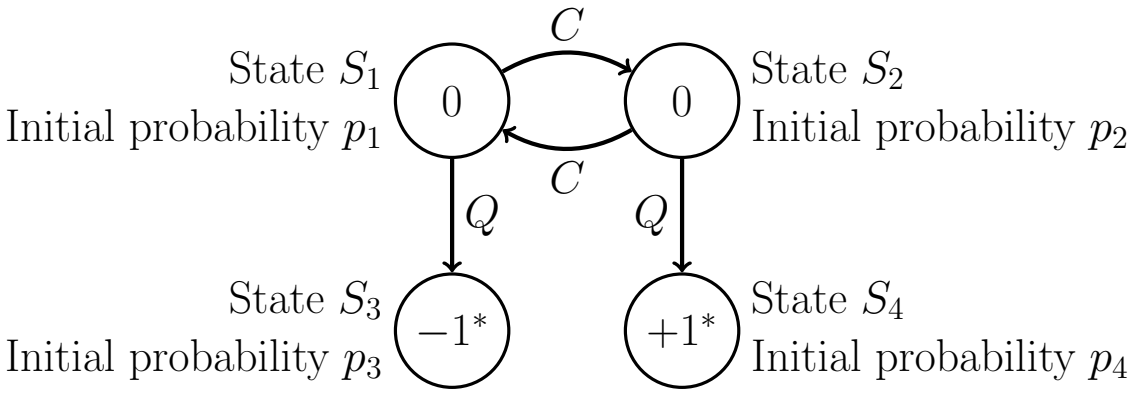


Figure 5: 1-player game in which each stage has duration 1

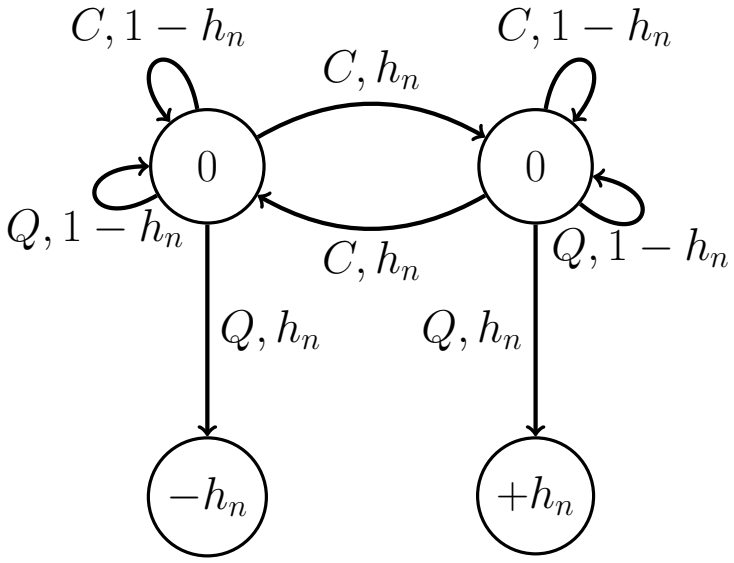


Figure 6: 1-player game in which n-th stage has duration h_n; transitions and payoff during n-th stage

Example 2. Consider the following one-player state-blind stochastic game. (See Figure 5). There are 4 states S_i $(i = 1, 2, 3, 4)$, states S_3 and S_4 are absorbing. The payoff in state S_3 is -1, and the payoff

in state S_4 is 1. The payoff in states S_1, S_2 is always 0, for any actions of the players. Player 2 has only one action and cannot influence the game. Player 1 has two actions, C and Q. Figure 5 shows the transitions between states. For examples, if the current state is S_1 and player 1 plays C, then the next state is S_2. Now, we consider the state-blind stochastic game with n-th stage duration h_n. The state and action spaces are still the same, but the transition probabilities have changed. Figure 6 shows transition probabilities during n-th stage. The arrow from state s_1 to state s_2 with caption (X, pr) means that the probability to go from state s_1 to state s_2 while playing action X is pr. Denote the initial probability distribution on the states by $p = (p_1, p_2, p_3, p_4)$, where p_i is the probability that starting state is S_i.

Proposition 4. In Example 2, the uniform limit $\lim\limits_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=+\infty}} v_{T_\infty,\lambda}(p)$ exists, and we have

$$\lim_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=+\infty}} v_{T_\infty,\lambda}(p) = p_4 - p_3 + \frac{1}{1+\lambda}\max\{0, p_2 - p_1\},$$

whereas

$$v_{1,\lambda}(p) = p_4 - p_3 + \max\{(p_2 - p_1)(1-\lambda), (p_1 - p_2)(1-\lambda)^2\}.$$

Remark 14 (Difference between stochastic games with stage duration in the case of perfect observation of the state and in the case of no observation of the state)**.** In the case of perfect observation of the state, there is an equality $\lim\limits_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=+\infty}} v_{T_\infty,\lambda}(p) = v_{1,\frac{\lambda}{1+\lambda}}(p)$, which clearly does not hold in this game. Moreover, $\lim\limits_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=+\infty}} v_{T_\infty,\lambda}(p) \neq v_{1,\lambda'}(p)$ for any $\lambda' \in (0,1)$. The equality $\lim\limits_{\lambda\to 0}\lim\limits_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=+\infty}} v_{T_\infty,\lambda}(p) = \lim\limits_{\lambda\to 0} v_{1,\lambda}(p)$ also does not hold here, but holds for the case of perfect observation of the state.

Remark 15. Absence of the equality $\lim\limits_{\lambda\to 0}\lim\limits_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=+\infty}} v_{T_\infty,\lambda} = \lim_{\lambda\to 0} v_{1,\lambda}$ in the above example suggests that there may be games, in which the limit value $\lim\limits_{\lambda\to 0} v_{1,\lambda}$ does not exist, but the limit value of the game with vanishing stage duration ($\lim_{\lambda\to 0}\lim\limits_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=+\infty}} v_{T_\infty,\lambda}$) exists. And such a game indeed exists, Theorem 1 from [Nov24] provides an example. (See also Remark 28 below).

We give an elementary proof of this proposition. Later in Remark 17 we give a possible shorter (but less elementary) proof.

Proof of Proposition 4. Denote

$$A_n := \{\text{The current state is } S_1, \text{ after action } C \text{ is played } n \text{ times}\};$$
$$B_n := \{\text{The current state is } S_2, \text{ after action } C \text{ is played } n \text{ times}\}.$$

We consider several cases.

Case 1: $p_3 = p_4 = 0, p_2 \geq p_1$.

It is easy to see that for all $n \in \mathbb{N}^*$ we have

$$P(A_{n+1}) = (1-h_n)P(A_n) + h_n P(B_n) \text{ and } P(B_{n+1}) = (1-h_n)P(B_n) + h_n P(A_n). \tag{5}$$

Assume that $P(B_n) \geq P(A_n)$ for some n. In this case by (5) we have $P(B_{n+1}) \leq P(B_n)$. Since $P(B_0) = p_2 \geq p_1 = P(A_0)$, we proved that $P(B_n)$ is maximal when $n = 0$.

Thus an optimal strategy is to always play Q.

Now, consider a 0-players stochastic game G with perfect observation of the state, with states s_1, s_2. Payoff is 0 in state s_1 and $p_2 - p_1$ in state s_2, the starting state is s_1. The probability to go from state s_1 to state s_2 is 1, and the probability to go from state s_2 to state s_2 is 1 too. See Figure 7. Denote by $\widetilde{v}_{1,\lambda}$ the value of λ-discounted game G_λ. It is easy to see that $\widetilde{v}_{1,\lambda} = (p_2 - p_1)(1-\lambda)$. Now, consider the game $G_{T_\infty,\lambda}$ with n-th stage duration h_n. Denote by $\widetilde{v}_{T_\infty,\lambda}$ the value of λ-discounted game $G_{T_\infty,\lambda}$ with n-th stage duration h_n. By Corollary 2(1) we have

$$\lim_{\substack{\sup_{i\in\mathbb{N}^*} h_i \to 0 \\ h_1+h_2+\ldots=+\infty}} \widetilde{v}_{T_\infty,\lambda} = \widetilde{v}_{1,\frac{\lambda}{1+\lambda}} = \frac{p_2 - p_1}{1+\lambda}.$$

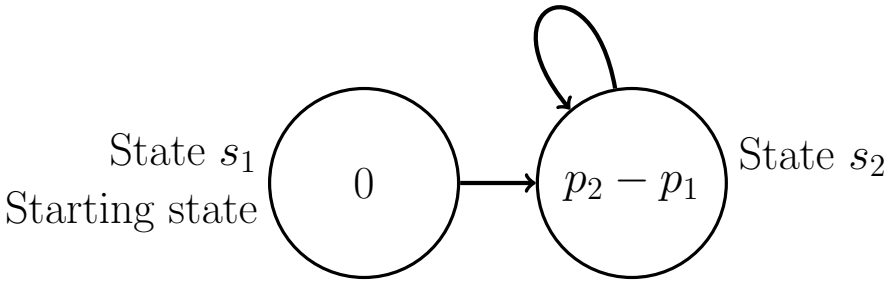


Figure 7: Game G

Now note that by construction of the game G, we have

$$\lim_{\substack{\sup_{i\in\mathbb{N}^*} h_i \to 0 \\ h_1+h_2+\ldots=+\infty}} v_{T_\infty,\lambda}(p) = \lim_{\substack{\sup_{i\in\mathbb{N}^*} h_i \to 0 \\ h_1+h_2+\ldots=+\infty}} \widetilde{v}_{T_\infty,\lambda} = \frac{p_2 - p_1}{1+\lambda},$$

where $p = (p_1, p_2, 0, 0)$ with $p_2 \geq p_1$.

Case 2: $p_3 = p_4 = 0, p_1 > p_2$.

Assume that $P(A_n) > P(B_n)$ for some n and $\sup_{i\in\mathbb{N}^*} h_i \leq 1/2$. In this case by (5) we have

$$\begin{aligned} P(A_n) > P(B_n) &\iff (1-2h_n)P(A_n) > (1-2h_n)P(B_n) \iff \\ \iff (1-h_n)P(A_n) + h_n P(B_n) &> (1-h_n)P(B_n) + h_n P(A_n) \iff P(A_{n+1}) > P(B_{n+1}). \end{aligned}$$

Since $P(A_0) = p_1 > p_2 = P(B_0)$, This means that $P(A_n) > P(B_n)$ for all $n \in \mathbb{N}^*$. Thus the only optimal strategy is to always play C, hence we have

$$\lim_{\substack{\sup_{i\in\mathbb{N}^*} h_i \to 0 \\ h_1+h_2+\ldots=+\infty}} v_{T_\infty,\lambda}(p) = 0.$$

Case 3: Any $p = (p_1, p_2, p_3, p_4)$ with $p_1 + p_2 + p_3 + p_4 = 1$.

From cases 1 and 2 it follows that

$$\lim_{\substack{\sup_{i\in\mathbb{N}^*} h_i \to 0 \\ h_1+h_2+\ldots=+\infty}} v_{T_\infty,\lambda}(p) = p_4 - p_3 + \frac{1}{1+\lambda}\max\{0, p_2 - p_1\}.$$

The fact that $v_{1,\lambda}(p) = p_4 - p_3 + \max\{(p_2-p_1)(1-\lambda), (p_1-p_2)(1-\lambda)^2\}$ is a simple computation. □

4.4 New result: the uniform limit $\lim_{\sup_{i\in\mathbb{N}^*} h_i\to 0} v_{T_\infty,k}$ is a viscosity solution of a partial differential equation

In this subsection, we present our main result which says that for state-blind games, $\lim\limits_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=+\infty}} v_{T_\infty,\lambda}$ exists and is a unique viscosity solution of a partial differential equation.

Recall that if X is a finite set, $\zeta\in\Delta(X)$, and μ is a $|X|\times|X|$ matrix, then for any $x\in X$

$$(\zeta*\mu)(x):=\sum_{x'\in X}\zeta(x')\cdot\mu_{x'x}.$$

Definition 10. A function $u:[0,T)\times\Delta(\Omega)\to\mathbb{R}$ is called a *viscosity solution* of the partial differential equation

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

if:

1. for any $\mathcal{C}^1$ function $\psi:[0,T)\times\Delta(\Omega)\to\mathbb{R}$ with $u-\psi$ having a strict local maximum at $(\overline{t},\overline{\omega})\in[0,T)\times\Omega$ we have $0\le\frac{d}{dt}\psi(\overline{t},\overline{p})+\mathtt{val}_{I\times J}[k(\overline{t})g(i,j,\overline{p})+\langle\overline{p}*q(i,j),\nabla\psi(\overline{t},\overline{p})\rangle]$;
2. for any $\mathcal{C}^1$ function $\psi:[0,T)\times\Delta(\Omega)\to\mathbb{R}$ with $u-\psi$ having a strict local minimum at $(\overline{t},\overline{\omega})\in[0,T)\times\Omega$ we have $0\ge\frac{d}{dt}\psi(\overline{t},\overline{p})+\mathtt{val}_{I\times J}[k(\overline{t})g(i,j,\overline{p})+\langle\overline{p}*q(i,j),\nabla\psi(\overline{t},\overline{p})\rangle]$.

By a construction in §4.1, any state-blind stochastic game with state space Ω is equivalent to a stochastic game with state space $\Delta(\Omega)$. The following theorem considers the value of this equivalent game.

Theorem 2. If (Ω,I,J,g,q) is a state-blind stochastic game, then uniform limit $\lim\limits_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=T}} v_{T_\infty,k}(t,p)$ exists and is a unique viscosity solution of a partial differential equation (in $v(t,p)$)

$$0=\frac{d}{dt}v(t,p)+\mathtt{val}_{I\times J}[k(t)g(i,j,p)+\langle p*q(i,j),\nabla v(t,p)\rangle]. \tag{6}$$

For the discounted case the above equation is more simple:

Corollary 3. If (Ω,I,J,g,q) is a state-blind stochastic game, then $\lim\limits_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=+\infty}} v_{T_\infty,\lambda}(p)$ exists and is a unique viscosity solution of a partial differential equation (in $v(p)$)

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

Proof. In (6), make a substitution $v(t,p)\mapsto e^{-\lambda t}v(p)$. □

Remark 16. Theorem 1 and Theorem 2 are two "twin" theorems, the first one is for stochastic games with perfect observation of the state, and the second one is for state-blind stochastic games.

Remark 17 (Use of Corollary 3)**.** Corollary 3 has some practical use for computing the values of games with vanishing stage duration.
First, one can check that at least if p is such that $p_1\neq p_2$, then $\lim\limits_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=+\infty}} v_{T_\infty,\lambda}(p)$ from Proposition 4 satisfies to the partial differential equation given in Corollary 3.
Second, with the help of Corollary 3 one can prove that there exists a game in which the limit value $\lim\limits_{\lambda\to 0} v_{1,\lambda}$ does not exist, but the limit value of the game with vanishing stage duration ($\lim\limits_{\lambda\to 0}\lim\limits_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=+\infty}} v_{T_\infty,\lambda}$) exists. See [Nov24] for a paper on the matter. Note that such a situation is impossible in the case of games with perfect observation of the state (by Corollary 2(3)), so one needs to consider the case of games with imperfect observation of the state. (See also Remark 28 below).

5 The proof of Theorem 2

The proof consists of two parts. First, we prove Lemma 3. Second, we use this lemma to prove Theorem 2.

Lemma 3 is an analogue of [Sor17, Proposition 3.11], while Theorem 2 is an analogue of [Sor17, Proposition 3.12]. The proof of Lemma 3 is different from the proof of its analogue, while the proof of Theorem 2 is almost identical to [Sor17, proof of Proposition 3.12].

Recall that if X is a finite set, $\zeta \in \Delta(X)$, and μ is a $|X| \times |X|$ matrix, then for any $x \in X$

$$(\zeta * \mu)(x) = \sum_{x' \in X} \zeta(x') \cdot \mu_{x'x}.$$

We denote for $x \in \Delta(I), y \in \Delta(J), \omega \in \Omega$

$$q(x,y) := \sum_{i \in I, j \in J} x(i)y(j)q(i,j);$$
$$g(x,y,\omega) := \sum_{i \in I, j \in J} x(i)y(j)g(i,j,\omega).$$

Lemma 3. The family $\{v_{T_\infty,k}(t,p)\}_{T_\infty}$ is equilipschitz-continuous and equibounded for all partitions T_∞ with $\sup_{i \in \mathbb{N}^*} h_i$ small enough, i.e., there are positive constants C_1, C_2, C_3 and there is $\delta \in (0,1]$ such that for any $t^1, t^2 \in [0,T), p_1, p_2 \in \Delta(\Omega)$ and for any partition $T_\infty = \{t_n\}_{n \in \mathbb{N}^*}$ with $t_{n+1} - t_n \leq \delta$, we have

$$|v_{T_\infty,k}(t^1,p_1) - v_{T_\infty,k}(t^2,p_2)| \leq C_1\|p_1 - p_2\|_1 + C_2|t^1 - t^2|;$$
$$|v_{T_\infty,k}(t^1,p_1)| \leq C_3.$$

Proof. First, we prove equiboundedness. By Remark 8 (with $\varepsilon = 1$) we have for $\sup_{i \in \mathbb{N}^*} h_i$ small enough

$$|v_{T_\infty,k}(t^1,p_1)| \leq |v_{T_\infty,k}(0,p_1)| \leq \|g\|_\infty \sum_{j=1}^{\infty} h_j k(t_j) \leq 2\|g\|_\infty,$$

The rest of the proof is devoted to equilipschitz-continuity.

For $x \in \Delta(I), y \in \Delta(J)$, denote

$$P^h(x,y) := Id + hq(x,y).$$

Denote for all $p \in \Delta(\Omega), \omega \in \Omega, x \in X, y \in Y$

$$\widehat{p}^h(x,y)(\omega) := (p * P^h(x,y))(\omega).$$

It is easy to see that if

- For some $n \in \mathbb{N}^*$ we have $t_{n+1} - t_n = h$;
- At the n-th stage players play a (mixed) action profile (x,y);
- The distribution of states at the start of the n-th stage is p,

then $\widehat{p}^h(x,y)$ is the distribution of states at the start of $(n+1)$-th stage.

We have for all $h \in (0,1], x \in X, y \in Y$

$$\left\|\widehat{p}_1^h(x,y) - \widehat{p}_2^h(x,y)\right\|_1 \leq \left\|(p_1 - p_2) * P^h(x,y)\right\|_1 = \left\|(P^h(x,y))^T \cdot (p_1 - p_2)\right\|_1,$$

where $p_1 - p_2$ is a vector column, and we have

$$\left\|(P^h(x,y))^T \cdot (p_1 - p_2)\right\|_1 \leq \|(P^h(x,y))^T\|_{\mathtt{op}}\|p_1 - p_2\|_1,$$

where $\|(P^h(x,y))^T\|_{\mathtt{op}}$ is the operator norm of the operator $z \mapsto (P^h(x,y))^T \cdot z$, i.e.

$$\left\|(P^h(x,y))^T\right\|_{\mathtt{op}} = \sup_{z:\|z\|_1=1} \left\|(P^h(x,y))^T \cdot z\right\|_1.$$

It is easy to see that

$$\left\|(P^h(x,y))^T\right\|_{\mathtt{op}} = 1,$$

thus we have

$$\left\|\widehat{p}_1^h(x,y) - \widehat{p}_2^h(x,y)\right\|_1 \leq \|p_1 - p_2\|_1. \tag{7}$$

Fix $t_n \in T_\infty$, and $p_1, p_2 \in \Delta(\Omega)$. By Proposition 1(5) there exists a profile of optimal Markov strategies (σ_1, τ_1) (respectively there exists a profile of optimal Markov strategies (σ_2, τ_2)), if starting time is t_n and initial distribution of the states is p_1 (respectively p_2). For a strategy (σ_i, τ_i) $(i = 1, 2)$, denote by (x_i^j, y_i^j) the profile of mixed actions played at the j-th stage (it depends on the j-th stage distribution of states p_i^j). We have

$$|v_{T_\infty,k}(t_n, p_1) - v_{T_\infty,k}(t_n, p_2)| \leq \sum_{j=n}^{\infty} h_j k(t_j) \left|g(x_1^j, y_1^j, p_1^j) - g(x_2^j, y_2^j, p_2^j)\right|, \tag{8}$$

where $p_i^1 = p_i$, and $p_i^{j+1} = \widehat{p_i^j}^{h_j}(x_i^j, y_i^j)$ (for $i = 1, 2$ and $j \geq 1$). Now, we have for any j (assuming without the loss of generality assume that $g(x_1^j, y_1^j, p_1^j) \geq g(x_2^j, y_2^j, p_2^j)$)

$$\begin{aligned}
\left|g(x_1^j, y_1^j, p_1^j) - g(x_2^j, y_2^j, p_2^j)\right| &\leq \left|g(x_1^j, y_2^j, p_1^j) - g(x_1^j, y_2^j, p_2^j)\right| \\
&= \left|\sum_{k\in\Omega} p_1^j(k) g(x_1^j, y_2^j, \omega_k) - \sum_{k\in\Omega} p_2^j(k) g(x_1^j, y_2^j, \omega_k)\right| \\
&\leq \|g\|_\infty \sum_{k\in\Omega} \left|p_1^j(k) - p_2^j(k)\right| = \|g\|_\infty \left\|p_1^j - p_2^j\right\|_1 \leq \|g\|_\infty \|p_1 - p_2\|_1,
\end{aligned} \tag{9}$$

where the last inequality follows from (7). By Remark 8 (with $\varepsilon = 1$) and by combining (8) and (9), we obtain

$$|v_{T_\infty,k}(t_n, p_1) - v_{T_\infty,k}(t_n, p_2)| \leq \|g\|_\infty \|p_1 - p_2\|_1 \sum_{j=n}^{\infty} h_j k(t_j) \leq 2 \|g\|_\infty \|p_1 - p_2\|_1, \tag{10}$$

where the second inequality holds if $\sup_{n\in\mathbb{N}^*} h_n$ is small enough.

Fix $p \in \Delta(\Omega)$, and $t_n, t_m \in T_\infty$ with $t_m \geq t_n$. Let (σ_1, τ_1) (respectively (σ_2, τ_2)) be a profile of optimal Markov strategies, if starting time is t_n (respectively t_m) and initial distribution of the states is p. For a strategy (x_i, y_i), denote by (x_i^j, y_i^j) the profile of mixed actions played at the j-th stage (it depends on the current distribution p_i^j). Without the loss of generality assume that $g(x_1^j, y_1^j, p_1^j) \geq g(x_2^j, y_2^j, p_2^j)$. We have

$$\begin{aligned}
|v_{T_\infty,k}(t_n, p) - v_{T_\infty,k}(t_m, p)| &\leq \sum_{j=n}^{m-1} h_j k(t_j) \left|g(x_1^j, y_1^j, p_1^j)\right| + \sum_{j=m}^{\infty} h_j k(t_j) \left|g(x_1^j, y_1^j, p_1^j) - g(x_2^j, y_2^j, p_2^j)\right| \\
&\leq \|k\|_\infty \|g\|_\infty |t_n - t_m| + \sum_{j=m}^{\infty} h_j k(t_j) \left|g(x_1^j, y_1^j, p_1^j) - g(x_2^j, y_2^j, p_2^j)\right|.
\end{aligned} \tag{11}$$

where $p_1^n = p_2^m = p$ and $p_i^{j+1} = \widehat{p_i^j}^{h_j}(x_i^j, y_i^j)$ (for $i = 1, j \geq n$ or $i = 2, j \geq m$). There exists $\overline{p} \in \Delta(\Omega)$ such that

$$\|p_2^m - p_1^m\|_1 = \left\| p - \left(\prod_{j=n}^{m-1}(1-h_j)p + \left(1 - \prod_{j=n}^{m-1}(1-h_j)\right)\overline{p} \right) \right\|_1 . \tag{12}$$

By Lemma 2(3) we have

$$\begin{aligned}
\left\| p - \left(\prod_{j=n}^{m-1}(1-h_j)p + \left(1 - \prod_{j=n}^{m-1}(1-h_j)\right)\overline{p} \right) \right\|_1 &\leq \left\| \left(1 - \prod_{j=n}^{m-1}(1-h_j)\right) p \right\|_1 + \left\| \left(1 - \prod_{j=n}^{m-1}(1-h_j)\right)\overline{p} \right\|_1 \\
&\leq \left\| \left(1 - \left(1 - \sum_{j=n}^{m-1} h_j\right)\right) p \right\|_1 + \left\| \left(1 - \left(1 - \sum_{j=n}^{m-1} h_j\right)\right)\overline{p} \right\|_1 \\
&\leq |t_n - t_m| \left(\|p\|_1 + \|\overline{p}\|_1 \right) = 2\,|t_n - t_m| .
\end{aligned} \tag{13}$$

By combining (12) and (13), we have

$$\|p_2^m - p_1^m\|_1 \leq 2\,|t_n - t_m| . \tag{14}$$

By Remark 8 (with $\varepsilon = 1$) and by combining (7), (11), and (14), we have

$$\begin{aligned}
|v_{T_\infty,k}(t_n, p) - v_{T_\infty,k}(t_m, p)| &\leq \|k\|_\infty \|g\|_\infty |t_n - t_m| + 2\,|t_n - t_m| \sum_{j=n}^{\infty} h_j k(t_j) \\
&\leq \|k\|_\infty \|g\|_\infty |t_n - t_m| + 4|t_n - t_m| = (\|k\|_\infty \|g\|_\infty + 4)|t_n - t_m|.
\end{aligned} \tag{15}$$

By combining (10) and (15), we have for any $p_1, p_2 \in \Delta(\Omega)$, $t^1 = t_n \in T_\infty$, and $t^2 = t_m \in T_\infty$

$$\begin{aligned}
|v_{T_\infty,k}(t^1, p_1) - v_{T_\infty,k}(t^2, p_2)| &\leq |v_{T_\infty,k}(t^1, p_1) - v_{T_\infty,k}(t^1, p_2)| + |v_{T_\infty,k}(t^1, p_2) - v_{T_\infty,k}(t^2, p_2)| \\
&\leq 2\,\|g\|_\infty \,\|p_1 - p_2\|_1 + (\|k\|_\infty \|g\|_\infty + 4)|t^1 - t^2|.
\end{aligned} \tag{16}$$

Now, we prove that this inequality holds for any $t^1, t^2 \in [0, T)$. Denote $C_1 := 2\,\|g\|_\infty$ and $C_2 := \|k\|_\infty \|g\|_\infty + 4$. Without the loss of generality, assume that $t^1 \geq t^2$. By the definition of $v_{T_\infty,\lambda}$ there exist integers n, m and numbers $\alpha \in [0, 1], \beta \in [0, 1]$ such that for any $p \in \Delta(\Omega)$ we have

$$v_{T_\infty,k}(t^1, p) = \alpha v_{T_\infty,k}(t_n, p) + (1-\alpha) v_{T_\infty,k}(t_{n+1}, p) \;\text{ and }\; v_{T_\infty,k}(t^2, p) = \beta v_{T_\infty,k}(t_m, p) + (1-\beta) v_{T_\infty,k}(t_{m+1}, p).$$

We have

$$\begin{aligned}
|v_{T_\infty,k}(t^1, p_1) - v_{T_\infty,k}(t^1, p_2)| &\leq \alpha |v_{T_\infty,k}(t_n, p_1) - v_{T_\infty,k}(t_n, p_2)| + (1-\alpha)|v_{T_\infty,k}(t_{n+1}, p_1) - v_{T_\infty,k}(t_{n+1}, p_2)|, \\
&\leq C_1 \|p_1 - p_2\|_1.
\end{aligned}$$

If $\beta \geq \alpha$, then $t_{m+1} > t_m \geq t_{n+1} > t_n$, and we have

$$\begin{aligned}
|v_{T_\infty,k}(t^1, p_2) - v_{T_\infty,k}(t^2, p_2)| &\leq \alpha |v_{T_\infty,k}(t_n, p_1) - v_{T_\infty,k}(t_m, p_2)| + (\beta - \alpha)|v_{T_\infty,k}(t_{n+1}, p_1) - v_{T_\infty,k}(t_m, p_2)|, \\
&\quad + (1-\beta)|v_{T_\infty,k}(t_{n+1}, p_1) - v_{T_\infty,k}(t_{m+1}, p_2)| \\
&\leq \alpha(C_1 \|p_1 - p_2\|_1 + C_2|t_m - t_n|) + (\beta - \alpha)(C_1 \|p_1 - p_2\|_1 + C_2|t_m - t_{n+1}|) \\
&\quad + (1-\beta)(C_1 \|p_1 - p_2\|_1 + C_2|t_{m+1} - t_{n+1}|) \\
&= C_1 \|p_1 - p_2\|_1 + C_2|\alpha(t_m - t_n) + (\beta - \alpha)(t_m - t_{n+1}) + \beta(t_{m+1} - t_{n+1})| \\
&= C_1 \|p_1 - p_2\|_1 + C_2|\beta t_m + (1-\beta)t_{m+1} - (\alpha t_n + (1-\alpha)t_{n+1})| \\
&= C_1 \|p_1 - p_2\|_1 + C_2|t^2 - t^1|.
\end{aligned}$$

If $\alpha \geq \beta$, then $t_{m+1} \geq t_{n+1}$ and $t_m > t_n$, and a computation similar to the above one shows that

$$|v_{T_\infty,k}(t^1, p_2) - v_{T_\infty,k}(t^2, p_2)| \leq C_1 \|p_1 - p_2\|_1 + C_2 |t^2 - t^1|.$$

Thus we proved (16) for any $t^1, t^2 \in [0, T)$, and hence we proved the lemma. □

Denote $U := \left\{ \text{accumulation points of uniform (in both } p \text{ and } t) \text{ limit } \lim_{\substack{\sup_{i\in\mathbb{N}^*} h_i \to 0 \\ h_1+h_2+\ldots=T}} v_{T_\infty,k}(t,p) \right\}$.

Corollary 4. $U \neq \varnothing$.

Proof. Follows directly from the previous lemma and the Arzelà–Ascoli theorem. □

Proof of Theorem 2. Fix a partition T_∞. By Proposition 1(4) we have for any $t_n \in T_\infty$

$$v_{T_\infty,k}(t_n, p) = \mathtt{Val}_{I\times J} \left[h_n k(t_n) g(i,j,p) + v_{T_\infty,k}(t_{n+1}, \widetilde{p}(i,j)) \right], \tag{17}$$

where $\widetilde{p}(i_n, j_n) = p * (Id + h_n q(i,j))$.

By Corollary 4 we have $U \neq \varnothing$. We are going to prove that any $\mathcal{U} \in U$ is a viscosity solution of the partial differential equation

$$0 = \frac{d}{dt} V(t,p) + \mathtt{val}_{I\times J}[k(t) g(i,j,p) + \langle p * q(i,j), \nabla V(t,p) \rangle].$$

Let $\psi(t,p)$ be a $\mathcal{C}^1$ function such that $\mathcal{U} - \psi$ has strict local maximum at $(\overline{t}, \overline{p}) \in [0,T) \times \Delta(\Omega)$. Consider a sequence of partitions $\{T_\infty(m)\}_{m\in\mathbb{N}^*}$ such that in the partition $T_\infty(m)$ we have $\sup_{i\in\mathbb{N}^*} h_i \to 0$ as $m \to \infty$, and such that a sequence $W_m = v_{T_\infty(m),\lambda}$ converging uniformly to $\mathcal{U}$ as $m \to \infty$, and let $(t^*(m), p(m))$ be a maximizing (locally near $(\overline{t}, \overline{p})$) sequence for $(W_m - \psi)(t,p)$, where $t^*(m) \in T_\infty(m)$. In particular, $(t^*(m), p(m))$ converges to $(\overline{t}, \overline{p})$ as $m \to \infty$. Given an optimal in (17) mixed strategy $x^*(m) \in \Delta(I)$, one has with $t^*(m) = t_n \in T_\infty(m)$

$$W_m(t_n, p(m)) \leq E_{x^*(m),y} \left[h_n k(t_n) g(i,j,p(m)) + W_m(t_{n+1}, \widetilde{p}(i,j)) \right], \ \forall y \in \Delta(J).$$

For m large enough, the choice of $(t^*(m), p(m))$ implies

$$\psi(t_n, p(m)) - W_m(t_n, p(m)) \leq \psi(t_{n+1}, \widetilde{p}(i,j)) - W_m(t_{n+1}, \widetilde{p}(i,j)), \ \forall i \in I, j \in J.$$

By using the continuity of k and ψ being $\mathcal{C}^1$, and the Taylor's theorem for $\psi(t_{n+1}, \cdot)$, one obtains for all $y \in \Delta(J)$

$$\begin{aligned}
\psi(t_n, p(m)) &\leq E_{x^*(m),y} \left[h_n k(t_n) g(i,j,p(m)) + \psi(t_{n+1}, \widetilde{p}(i,j)) \right] \\
&= E_{x^*(m),y} \left[h_n k(t_n) g(i,j,p(m)) + \psi(t_{n+1}, p(m) * (Id + h_n q(i,j))) \right] \\
&\leq h_n k(t_n) g(x^*(m), y, p(m)) + \psi(t_{n+1}, p(m)) \\
&\quad + h_n E_{x^*(m),y} \langle p(m) * q(i,j), \nabla \psi(t_{n+1}, p(m)) \rangle + o(h_n),
\end{aligned}$$

where $o(h_n)/h_n \to 0$ when $h_n \to 0$. This gives for all $y \in \Delta(J)$

$$\begin{aligned}
0 \leq h_n \frac{\psi(t_{n+1}, p(m)) - \psi(t_n, p(m))}{h_n} + h_n k(t_n) g(x^*(m), y, p(m)) + \\
h_n E_{x^*(m),y} \langle p(m) * q(i,j), \nabla \psi(t_{n+1}, p(m)) \rangle + o(h_n).
\end{aligned}$$

Hence by dividing by h_n and taking the limit as $m \to \infty$, one obtains, for some accumulation point $x^* \in \Delta(I)$ (we use again the continuity of k and ψ being $\mathcal{C}^1$)

$$0 \leq \frac{d}{dt}\psi(\overline{t},\overline{p}) + k(\overline{t})g(x^*,y,\overline{p}) + E_{x^*,y}\langle \overline{p} * q(i,j), \nabla\psi(\overline{t},\overline{p})\rangle + o(h_n) \quad \forall y \in \Delta(J).$$

Analogously one can prove that if $\psi(t,p)$ is a $\mathcal{C}^1$ function such that $\mathcal{U} - \psi$ has strict local minimum at $(\overline{t},\overline{p}) \in [0,T) \times \Delta(\Omega)$, and $y^* \in \Delta(J)$ is optimal in (17), then

$$0 \geq \frac{d}{dt}\psi(\overline{t},\overline{p}) + k(\overline{t})g(x,y^*,\overline{p}) + E_{x,y^*}\langle \overline{p} * q(i,j), \nabla\psi(\overline{t},\overline{p})\rangle + o(h_n) \quad \forall x \in \Delta(I).$$

Thus $\mathcal{U}$ is a viscosity solution of

$$0 = \frac{d}{dt}V(t,p) + \mathtt{val}_{I\times J}[k(t)g(i,j,p) + \langle p * q(i,j), \nabla V(t,p)\rangle].$$

The uniqueness follows from [Sor17, Proposition 3.9]. □

6 Connection between games with stage duration and continuous-time Markov games

Throughout this entire subsection, T is either a positive number or $+\infty$. T_∞ is a partition of $[0,T)$; in other words, T_∞ a sequence $\{t_n\}_{n\in\mathbb{N}^*}$ such that $t_1 = 0$ and $t_n \xrightarrow{n\to\infty} T$. For each given partition T_∞, denote $h_n = t_{n+1} - t_n$ for each $n \in \mathbb{N}^*$.

6.1 Perfect observation of the state

In this section, we are going to briefly introduce the continuous-time Markov games (§6.1.1) and its discretization (§6.1.2). In §6.1.3, we consider the connection between continuous-time Markov games, its discretization, and stochastic games with stage duration.

6.1.1 Zero-sum continuous-time Markov games

This subsection briefly describes zero-sum continuous-time Markov games. In such games, two players jointly control a continuous-time Markov chain.

This subsection assumes that the reader know main facts about continuous-time Markov chains. See [And91] for a book about continuous-time Markov chains. See also [GHL09, Appendices B and C].

This subsection is based on [GHL03] and [GHL05].

A *zero-sum continuous-time Markov game* is a 5-tuple (Ω, I, J, g, q), where:

- Ω is the set of states;
- I is the non-empty set of actions of player 1;
- J is the non-empty set of actions of player 2;
- $g : I \times J \times \Omega \to \mathbb{R}$ is the instantaneous payoff function of player 1;
- $q : I \times J \to \{\text{matrices } |\Omega| \times |\Omega| \text{ satisfying property } *\}$ is a function which is called an *infinitesimal generator* of the game. 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.

In what follows, we assume that the sets I, J, Ω are finite. We denote by $q(i,j)(\omega_a,\omega_b)$ the (ω_a,ω_b)-th element of the matrix $q(i,j)$.

The game is played as follows. The initial state $\omega_0 \in \Omega$ is known to the players. Both players observe continuously the current state. Whenever the system is at state $i \in \Omega$ at time $t > 0$, players choose actions $i_t \in I, j_t \in J$ according to some mixed action, $\sigma_t : \Omega \to \Delta(I)$ for player 1 or $\tau_t : \Omega \to \Delta(J)$

for player 2. A *strategy* σ (or τ) is an indexed by $t \in \mathbb{R}_+$ collection of σ_t (or τ_t). There are some conditions on strategies, but we are not going to talk about it here, see [GHL03] or [GHL05] for details. We denote by Π_i $(i = 1, 2)$ the set of strategies of player i.

Given two strategies σ and τ, denote

$$q(\sigma, \tau, t)(\omega_a, \omega_b) := \sum_{i \in I, j \in J} q(i, j)(\omega_a, \omega_b) \cdot (\sigma_t(\omega_a))(i) \cdot (\tau_t(\omega_a))(j);$$

$$g(\sigma, \tau, \omega, t) := \sum_{i \in I, j \in J} g(i, j, \omega) \cdot (\sigma_t(\omega))(i) \cdot (\tau_t(\omega))(j).$$

Now, two strategies σ and τ generate a nonhomogeneous infinitesimal generator $\{q(t)\}_{t \in \mathbb{R}_+}$ in which (ω_a, ω_b)-th element of the matrix $q(t)$ is $q(\sigma, \tau, t)(\omega_a, \omega_b)$.

Under some conditions on strategies σ, τ, there exists a unique transition probability function

$$p : \mathbb{R}_+ \times \Omega \times \mathbb{R}_+ \times \Omega \times \Pi_1 \times \Pi_2 \to [0, 1]$$

such that

$$q(\sigma, \tau, s)(\omega_a, \omega_b) = \lim_{t \to s^+} \frac{p(s, \omega_a, t, \omega_b, \sigma, \tau) - \delta_{\omega_a \omega_b}}{t - s}.$$

$p(s, \omega_a, t, \omega_b, \sigma, \tau)$ is the probability that the state is ω_b at time t, given that the strategies of players are σ, τ and at time s the state was ω_a.

Definition 11. Fix a continuous-time Markov game (Ω, I, J, g, q). Given:

- two strategies σ, τ of players;
- an initial state ω_0;
- an initial time $s \in \mathbb{R}_+$;

the *total payoff* of the game (Ω, I, J, g, q) is

$$G_\lambda^{\mathtt{cont1}}(\omega_0, s, \sigma, \tau) := E_{\sigma,\tau}^{\omega_0, s} \left(\int_s^T \lambda e^{-\lambda(t-s)} g(i_t, j_t, \omega_t) dt \right).$$

The game (Ω, I, J, g, q) is said to have a *value* $v_\lambda^{\mathtt{cont1}}(\omega_0, s)$ if

$$v_\lambda^{\mathtt{cont1}}(\omega_0, s) = \sup_\sigma \inf_\tau G_\lambda^{\mathtt{cont1}}(\omega_0, s, \sigma, \tau) = \inf_\tau \sup_\sigma G_\lambda^{\mathtt{cont1}}(\omega_0, s, \sigma, \tau).$$

Denote $v_\lambda^{\mathtt{cont1}}(\omega_0) = v_\lambda^{\mathtt{cont1}}(\omega_0, 0)$. Games with such a total payoff (up to a multiplication by λ) were studied in [GHL05]. An article [GHL03] considers an alternative total payoff.

There is the following result.

Proposition 5. ([GHL05, Theorem 5.1]). $v_\lambda^{\mathtt{cont1}}(\omega_0)$ exists and is a unique solution of the equation (in $v(\omega)$)

$$\lambda v(\omega) = \mathtt{Val}_{I \times J} \left[\lambda g(i, j, \omega) + \langle q(i, j)(\omega, \cdot), v(\cdot) \rangle \right].$$

Remark 18. The above proposition holds not only for finite Ω, I, J, but in a more general case too.

Remark 19. An article [GHL05] considers a non-normalized total payoff, i.e., there is no multiplication by λ. Because of that, the equation in [GHL05, Theorem 5.1] has a slightly different form than the equation given here.

0-players games are precisely homogeneous continuous-time Markov chains. 1-player games are continuous-time Markov decision processes. See [GHL09] for a book about continuous-time Markov decision processes.

Remark 20. Note that this is not the only model of continuous-time games. The main difference between models is the definition of strategy spaces for the players. See [Ney17] for an alternative model of continuous-time games. See also [Isa99] for a book about differential games, a specific class of continuous-time games, in which a state is a point in $\mathbb{R}^n$ that changes continuously.

6.1.2 Discretization of zero-sum continuous-time Markov games

Now, we consider the model from [Sor17]. As before, T is either a positive number or $+\infty$. T_∞ is a partition of $[0,T)$; in other words, T_∞ a strictly increasing sequence $\{t_n\}_{n\in\mathbb{N}^*}$ such that $t_1=0$ and $t_n \xrightarrow{n\to\infty} T$. For each given partition T_∞, denote $h_n = t_{n+1}-t_n$ for each $n\in\mathbb{N}^*$.

Loosely speaking, we want to consider continuous-time Markov game, in which players only allowed to act at times $t_1, t_2, t_3, \dots$ At time $t\in[t_i,t_{i+1})$ players should act according to their decision at time t_i. Let us give a formal definition.

Fix a continuous-time Markov game (Ω, I, J, g, q) and fix also a nonincreasing continuous function $k:[0,T]\to\mathbb{R}_+$ with $\int_0^T k(t)dt = 1$.

A *discretization of a continuous-time Markov game, in which n-th action is taken at time* t_n proceeds in stages as follows. The initial state ω_1 is known to the players. At each stage $n\in\mathbb{N}^*$:

1. The players observe the current state ω_n;
2. Players choose mixed actions. Player 1 chooses $x_n\in\Delta(I)$ and player 2 chooses $y_n\in\Delta(J)$;
3. An action $i_n\in I$ of player 1 (respectively $j_n\in J$ of player 2) is chosen according to the probability measure $x_n\in\Delta(I)$ (respectivey $y_n\in\Delta(J)$);
4. Player 1 obtains a payoff $g_n = \int_{t_n}^{t_{n+1}} k(t)g(i_n,j_n,\omega_n)dt$, while player 2 obtains payoff $-g_n$. The new state ω_{n+1} is chosen according to the probability law $P_n = (\exp\{h_n q(i_n,j_n)\})(\omega_n,\cdot)$ (i.e., we consider ω_n-th row of the matrix $\exp\{h_n q(i_n,j_n)\}$, and this matrix's row generates a probability distribution on Ω).

The above description of the game is known to the players.

Remark 21. Note that above we just described a stochastic game $(\Omega, I, J, \{g_m\}_{m\in\mathbb{N}^*}, \{P_m\}_{m\in\mathbb{N}^*})$, where P_m is m-th stage transition probability function, and

$$g_n(i,j,\omega) = \int_{t_n}^{t_{n+1}} k(t)g(i,j,\omega)dt \quad \text{and} \quad P_n(i,j)(\omega,\omega') = (\exp\{h_n q(i,j)\})(\omega,\omega').$$

By the above remark, we can define strategies in the same way as in §3.1. Now let us define the total payoff function.

For each $t\in[h_n,h_{n+1})$, we denote by (x_t,y_t) (respectively (i_t,j_t)) a mixed (respectively pure) action which was chosen at time h_n.

Definition 12. Fix a discretization of a continuous-time Markov game (Ω, I, J, g, q), in which n-th action is taken at time t_n. Given:

- two strategies σ,τ of players;
- an initial state ω;
- an initial time $t_n\in T_\infty$;
- a nonincreasing continuous function $k:[0,T]\to\mathbb{R}_+$ with $\int_0^T k(t)dt = 1$,

the *total payoff* is

$$G^{\texttt{cont2}}_{T_\infty,k}(t_n,\omega) := E^{\omega,t_n}_{\sigma,\tau}\left(\int_{t_n}^T k(t)g(i_t,j_t,\omega_t)dt\right).$$

The game is said to have a *value* $v^{\texttt{cont2}}_{T_\infty,k}(t_n,\omega)$ if

$$v^{\texttt{cont2}}_{T_\infty,k}(t_n,\omega) = \sup_\sigma \inf_\tau G^{\texttt{cont2}}_{T_\infty,k}(t_n,\omega) = \inf_\tau \sup_\sigma G^{\texttt{cont2}}_{T_\infty,k}(t_n,\omega).$$

Note that the payoff depends, among other things, on the initial time $t_n\in T_\infty$.

We can define $G^{\texttt{cont2}}_{T_\infty,k}(t,\omega)$ and $v^{\texttt{cont2}}_{T_\infty,k}(t,\omega)$ for any $t\in[0,T)$ via linearity.

Some particular cases of the above definition have its own name. For finite T, the *game* $\Gamma^T(t)$ has a total payoff which is obtained if we set in the above definition $k(t)=1/T$ for any $t\in[0,T]$. For

$T = +\infty$ and fixed $\lambda \in (0,1)$, the λ-*discounted game* $\Gamma^\lambda(t)$ has a total payoff which is obtained if we set in the above definition $k(t) = \lambda e^{-\lambda t}$.

We are interested in the behavior of the value when $\sup_{i\in\mathbb{N}^*} h_i$ is vanishing, i.e. we are interested in the studying of

$$\lim_{\substack{\sup_{i\in\mathbb{N}^*} h_i \to 0 \\ h_1+h_2+\ldots=T}} v^{\mathtt{cont2}}_{T_\infty,k}(t,\omega).$$

There is the following result.

Proposition 6. ([Sor17, §4]). If (Ω, I, J, g, q) is a discretization of a continuous-time Markov game, then uniform limit $\lim\limits_{\substack{\sup_{i\in\mathbb{N}^*} h_i \to 0 \\ h_1+h_2+\ldots=T}} v^{\mathtt{cont2}}_{T_\infty,k}$ exists and is a unique viscosity solution of the differential equation (in $v(t,\omega)$)

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

Corollary 5. If (Ω, I, J, g, q) is a discretization of a continuous-time Markov game and starting time is 0, then uniform limit $\lim\limits_{\substack{\sup_{i\in\mathbb{N}^*} h_i \to 0 \\ h_1+h_2+\ldots=+\infty}} v^{\mathtt{cont2}}_{T_\infty,\lambda}$ exists and is a unique solution of the equation (in $v(\omega)$)

$$\lambda v(\omega) = \mathtt{Val}_{I\times J}\left[\lambda g(i,j,\omega) + \langle q(i,j)(\omega,\cdot)\,, v(\cdot)\rangle\right].$$

Remark 22 (Semi-Markov games)**.** Let us give a one more model of continuous-time games which is similar to the model here. Zero-sum semi-Markov games is a model in which, informally speaking, players play as in a continuous-time Markov game, but their set of strategies is very limited: players choose action in the start, and they can take another action only at the moment of the state change. After that, the players again cannot change the state until the state changes again, and so on. There are quite a lot of papers about semi-Markov games, for example, [LS92].

6.1.3 Comparison between models

First, we discuss a connection between 3 models of continuous-time games which were defined above.

Remark 23 (Comparison of our model with other models)**.**

1. In the continuous-time Markov games, players are generally allowed to choose actions at any time (respecting some measurability conditions); see Figure 8;
2. In the semi-Markov games, players are allowed to choose actions only at the time of the state change; see Figure 9;
3. In the discretizations of continuous-time Markov games, players are allowed to choose actions at some fixed times, depending on T_∞; see Figure 10;

So, given that $\sup_{i\in\mathbb{N}^*} h_i$ is small enough, we can consider that our model is, loosely speaking, situated between the model of continuous-time Markov games and the model of semi-Markov games.
However, when $\sup_{i\in\mathbb{N}^*} h_i \to 0$, our model more and more replicates the model of continuous-time Markov games, since the players can choose actions more and more often. This is confirmed by Proposition 7 below.

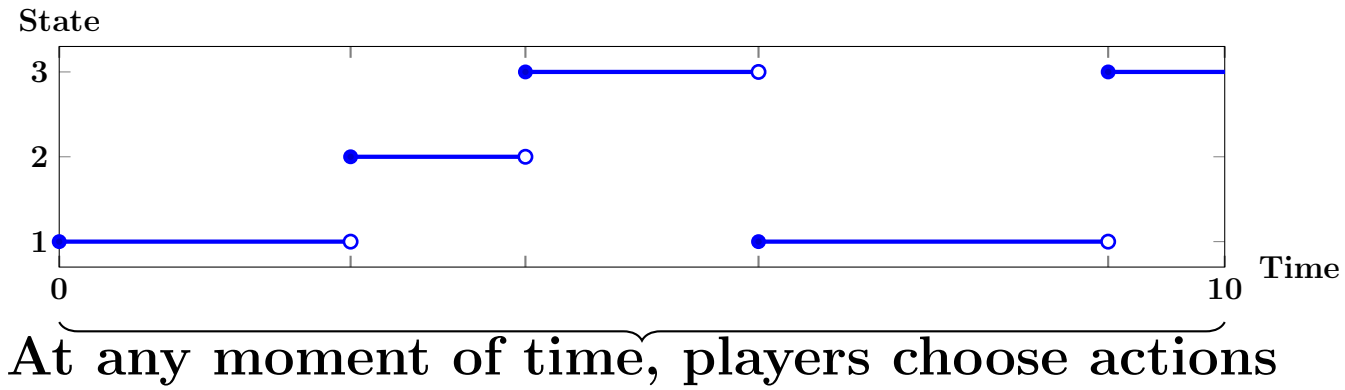


Figure 8: Continuous-time Markov game

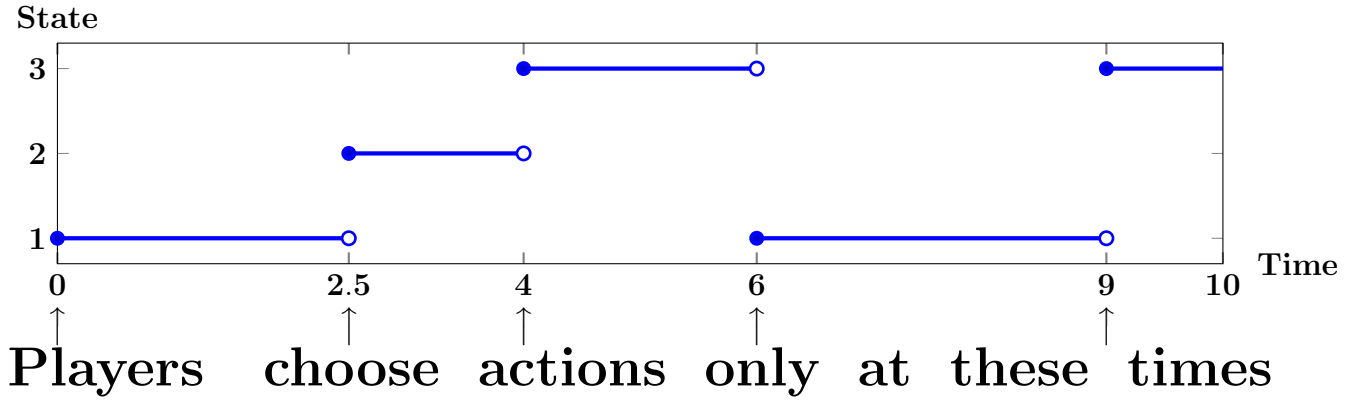


Figure 9: Semi-Markov game

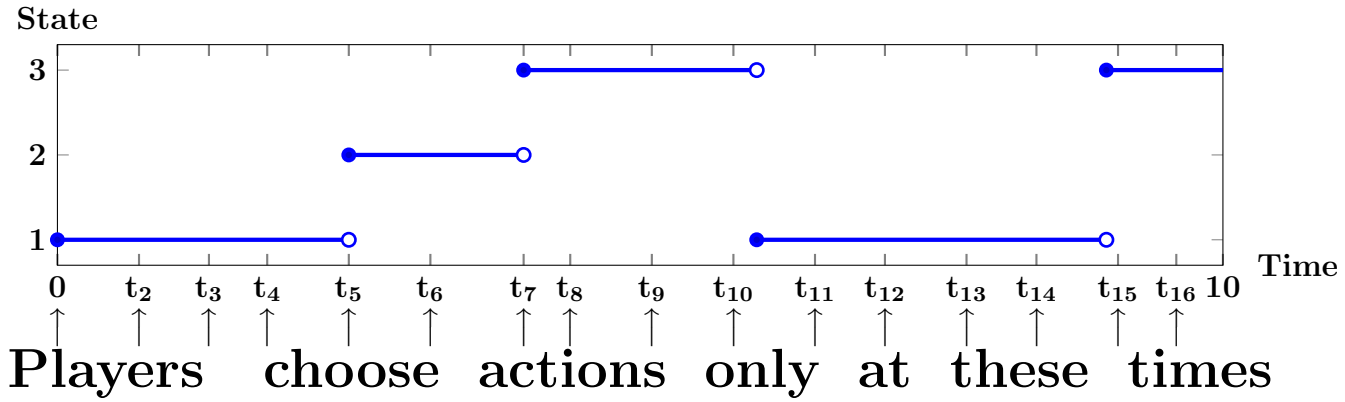


Figure 10: Continuous-time Markov game with stage duration

Proposition 7.

1. Fix a continuous-time Markov game (Ω, I, J, g, q_1), where q_1 is an infinitesimal operator, and fix a stochastic game (Ω, I, J, g, q_2), where q_2 is the kernel. If $q_1 = q_2 = q$, then both uniform limits $\lim_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0 \\ h_1+h_2+\ldots=T}} v^{\mathtt{cont2}}_{T_\infty,k}$ and $\lim_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0 \\ h_1+h_2+\ldots=T}} v_{T_\infty,k}$ exist and each of these limit function is a unique viscosity solution of the differential equation (in $v(t,\omega)$)

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

2. Fix a continuous-time Markov game (Ω, I, J, g, q_1), where q_1 is its infinitesimal operator, and fix a stochastic game (Ω, I, J, g, q_2), where q_2 is its kernel. If $q_1 = q_2 = q$, then both uniform limits $\lim\limits_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=+\infty}} v^{\texttt{cont2}}_{T_\infty,\lambda}$ and $\lim\limits_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=+\infty}} v_{T_\infty,\lambda}$ exist, and we have

$$v^{\texttt{cont1}}_{T_\infty,\lambda} = \lim_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=+\infty}} v^{\texttt{cont2}}_{T_\infty,\lambda} = \lim_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=+\infty}} v_{T_\infty,\lambda},$$

and each of these functions is a unique solution of the equation (in $v(t,\omega)$)

$$\lambda v(\omega) = \texttt{Val}_{I\times J}\left[\lambda g(i,j,\omega) + \langle q(i,j)(\omega,\cdot)\,, v(\cdot)\rangle\right].$$

Proof. Assertion 1. follows directly from Theorem 1 and Proposition 6. Assertion 2. follows directly from Corollaries 3, 5, and Proposition 5. □

Remark 24. The first assertion of the above proposition shows that if $\sup_{i\in\mathbb{N}^*} h_i$ is small, then the value of the stochastic game (Ω, I, J, g, q) with n-th stage duration h_n, and the value of the discretization of the continuous-time Markov game (Ω, I, J, g, q), in which n-th action is taken at time t_n, are close to each other.
The second assertion shows that in the discounted case, the above two values are in addition close to the value of the continuous-time Markov game (Ω, I, J, g, q). So, continuous-time Markov game can be considered as some kind of the "limit" of both the stochastic game with stage duration and the discretization of the continuous-time Markov game.
The author assumes that the second assertion holds not only for the discounted case, but for the general case too. However, it seems like such case were not studied in the literature about continuous-time Markov games.

6.2 No observation of the state

In this section, we give the above notions for the state-blind case.

6.2.1 Discretization of zero-sum continuous-time state-blind Markov games

It is straightforward to define discretization of continuous-time state-blind Markov games (unlike the continuous-time state-blind Markov games).

Recall that T is either a positive number or $+\infty$; $T_\infty = \{t_n\}_{n\in\mathbb{N}^*}$ is a partition of $[0,T)$; $h_n = t_{n+1} - t_n$ for each $n \in \mathbb{N}^*$.

Definition 13. Fix a continuous-time game (Ω, I, J, g, q). A *discretization of a continuous-time Markov game with public signals, in which n-th action is taken at time* t_n is the state-blind stochastic game $(\Omega, I, J, \{g_m\}_{m\in\mathbb{N}^*}, \{P_m\}_{m\in\mathbb{N}^*})$. Here

$$g_m(i,j,\omega) := \int_{t_m}^{t_{m+1}} k(t)g(i,j,\omega)dt \quad \text{and} \quad P_m(i,j)(\omega,\omega') := (\exp\{h_m q(i,j)\})(\omega,\omega')$$

is the m-th stage transition probability function, and $k : [0,T] \to \mathbb{R}_+$ is a nonincreasing continuous function with $\int_0^T k(t)dt = 1$.

For each initial time t and initial probability distribution on the states p, we can define the value $v^{\texttt{cont2}}_{T_\infty,k}(t,p)$ in the same way as in §6.1.2.

Such games were studied in [Sor17], in which the following result was proved.

Proposition 8. [Sor17, Proposition 5.3]. If (Ω, I, J, g, q) is a discretization of a continuous-time state-blind Markov game, then uniform limit $\lim\limits_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=T}} v^{\texttt{cont2}}_{T_\infty,k}(t,p)$ exists and is a unique viscosity solution of the partial differential equation (in $v(t,p)$)

$$0 = \frac{d}{dt}v(t,p) + \texttt{val}_{I\times J}[k(t)g(i,j,p) + \langle p * q(i,j), \nabla v(t,p)\rangle].$$

6.2.2 Comparison between models

The following proposition is a partial generalization of Proposition 7 for the state-blind stochastic games.

Proposition 9. Fix a discretization of a continuous-time state-blind Markov game (Ω, I, J, g, q_1), where q_1 is its infinitesimal operator, and fix a state-blind stochastic game (Ω, I, J, g, q_2), where q_2 is its kernel. If $q_1 = q_2 = q$, then the uniform limits $\lim\limits_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=+\infty}} v^{\texttt{cont2}}_{T_\infty,\lambda}$ and $\lim\limits_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=+\infty}} v_{T_\infty,\lambda}$ exist, and we have

$$\lim_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=+\infty}} v^{\texttt{cont2}}_{T_\infty,\lambda} = \lim_{\substack{\sup_{i\in\mathbb{N}^*} h_i\to 0\\ h_1+h_2+\ldots=+\infty}} v_{T_\infty,\lambda},$$

and each of these limit functions is a unique viscosity solution of the partial differential equation (in $v(t,p)$)

$$0 = \frac{d}{dt}v(t,p) + \texttt{val}_{I\times J}[k(t)g(i,j,p) + \langle p * q(i,j), \nabla v(t,p)\rangle].$$

Proof. Follows directly from Theorem 2 and Proposition 8. □

7 The proof of Theorem 1

The proof given below is a generalization of [SV16, proofs of Lemma 8.1 and Proposition 8.1]. Recall that $t_n \to T \in (0, +\infty]$ as $n \to \infty$, and for any $n \in \mathbb{N}^*$ we have $h_n = t_{n+1} - t_n$.

Proof. Let $v^{\texttt{cont2}}_{T_\infty,k}$ be the value of the discretization of the continuous-time game (Ω, I, J, g, q), where q is an infinitesimal generator. Define for $n \in \mathbb{N}^*$

$$\begin{aligned}
&\psi^h_n : \{\text{Continuous functions on } \Omega\} \to \{\text{Continuous functions on } \Omega\},\\
&f(\omega) \mapsto \texttt{Val}_{I\times J}[k(t_n)h_n g(i,j,\omega) + \langle (Id + h_n q(i,j))(\omega,\cdot)\,, f(\cdot)\rangle];\\
&\overline{\psi}^h_n : \{\text{Continuous functions on } \Omega\} \to \{\text{Continuous functions on } \Omega\},\\
&f(\omega) \mapsto \texttt{Val}_{I\times J}\left[\int_{t_n}^{t_{n+1}} k(t)g(i,j,\omega)dt + \langle \exp\{h_n q(i,j)\}(\omega,\cdot)\,, f(\cdot)\rangle\right].
\end{aligned}$$

First, note that by Proposition 1(2) we have for any continuous f and any $n \in \mathbb{N}^*$

$$\left\|\psi^h_n(f) - \overline{\psi}^h_n(f)\right\|_\infty \le \left|\left(k(t_n)h_n - \int_{t_n}^{t_{n+1}} k(t)dt\right)\right| \|g\|_\infty + \|(Id + h_n q) - \exp\{h_n q\}\|_1 \cdot \|f\|_\infty. \quad (18)$$

By the mean value theorem for integrals there exists $c \in (t_n, t_{n+1})$ such that for any $n \in \mathbb{N}^*$ we have

$$\frac{1}{h_n}\left|\left(k(t_n)h_n - \int_{t_n}^{t_{n+1}} k(t)dt\right)\right| = |k(t_n) - k(c)| \le k(t_n) - k(t_{n+1}). \quad (19)$$

There exists $C > 0$ such that for any $n \in \mathbb{N}^*$ we have

$$\frac{1}{h_n} \|(Id + h_n q) - \exp\{h_n q\}\|_1 = \frac{1}{h_n} \left\| (Id + h_n q) - \sum_{k=0}^{\infty} \frac{(h_n q)^k}{k!} \right\|_1 \leq \frac{1}{h_n} C h_n^2 = C h_n. \tag{20}$$

By combining (18)-(20), we obtain that there exist $C_1 > 0, C_2 > 0$ such that for any continuous f and any $n \in \mathbb{N}^*$ we have

$$\frac{1}{h_n} \left\| \psi_n^h(f) - \overline{\psi}_n^h(f) \right\|_\infty \leq C_1 |k(t_n) - k(t_{n+1})| + C_2 h_n \|f\|_\infty. \tag{21}$$

By (21) and Proposition 1(1,3) we have

$$\begin{aligned}
\left\| v_{T_\infty,k} - v_{T_\infty,k}^{\texttt{cont2}} \right\|_\infty &= \left\| \prod_{i=1}^{\infty} \psi_i^h(0) - \prod_{i=1}^{\infty} \overline{\psi}_i^h(0) \right\|_\infty \\
&\leq \left\| \psi_1^h \left(\prod_{i=2}^{\infty} \psi_i^h(0) \right) - \overline{\psi}_1^h \left(\prod_{i=2}^{\infty} \psi_i^h(0) \right) \right\|_\infty + \left\| \overline{\psi}_1^h \left(\prod_{i=2}^{\infty} \psi_i^h(0) \right) - \overline{\psi}_1^h \left(\prod_{i=2}^{\infty} \overline{\psi}_i^h(0) \right) \right\|_\infty \\
&\leq \left\| \psi_1^h \left(\prod_{i=2}^{\infty} \psi_i^h(0) \right) - \overline{\psi}_1^h \left(\prod_{i=2}^{\infty} \psi_i^h(0) \right) \right\|_\infty + \left\| \prod_{i=2}^{\infty} \psi_i^h(0) - \prod_{i=2}^{\infty} \overline{\psi}_i^h(0) \right\|_\infty .
\end{aligned}$$

By induction we obtain for any $N \in \mathbb{N}^*$

$$\left\| v_{T_\infty,k} - v_{T_\infty,k}^{\texttt{cont2}} \right\|_\infty \leq \sum_{m=1}^{N} \left(\left\| \psi_m^h \left(\prod_{i=m+1}^{\infty} \psi_i^h(0) \right) - \overline{\psi}_m^h \left(\prod_{i=m+1}^{\infty} \psi_i^h(0) \right) \right\|_\infty \right) + \left\| \prod_{i=N+1}^{\infty} \psi_i^h(0) - \prod_{i=N+1}^{\infty} \overline{\psi}_i^h(0) \right\|_\infty . \tag{22}$$

Note that by Proposition 1(3) we have for any $m \in \mathbb{N}^*$ and any initial time $t_m \in T_\infty$

$$v_{T_\infty,k}(t_m, \cdot) = \prod_{i=m}^{\infty} \psi_i^h(0) \text{ and } v_{T_\infty,k}^{\texttt{cont2}}(t_m, \cdot) = \prod_{i=m}^{\infty} \overline{\psi}_i^h(0). \tag{23}$$

Fix $\varepsilon > 0$. There is $S \in (0, T)$ such that for any $t_m \geq S$ and any $\sup_{i \in \mathbb{N}^*} h_i$ small enough we have (we use a computation similar to the one in Remark 8)

$$\|v_{T_\infty,k}(t_m, \cdot)\|_\infty \leq \|g\|_\infty \left(\int_S^T k(t) dt + \varepsilon \right) \leq 2\varepsilon \|g\|_\infty \quad \text{and} \quad \|v_{T_\infty,k}^{\texttt{cont2}}(t_m, \cdot)\|_\infty \leq 2\varepsilon \|g\|_\infty. \tag{24}$$

Also, if $\sup_{i \in \mathbb{N}^*} h_i$ is small enough, then by Remark 8 we have for any t_n

$$\|v_{T_\infty,k}(t_m, \cdot)\|_\infty \leq (1 + \varepsilon) \|g\|_\infty. \tag{25}$$

And analogously $\|v_{T_\infty,k}^{\texttt{cont2}}(t_m, \cdot)\|_\infty \leq (1 + \varepsilon) \|g\|_\infty$.

Now, let $N(T_\infty) \in \mathbb{N}^*$ be such that $t_{N(T_\infty)-1} < S$ and $t_{N(T_\infty)} \geq S$. Note that it depends on the partition T_∞ of $[0, T)$. By (21)–(25), there exist $C_1 > 0, C_2 > 0$ such that

$$\begin{aligned}
\left\| v_{T_\infty,k} - v_{T_\infty,k}^{\texttt{cont2}} \right\|_\infty &\leq \sum_{m=1}^{N(T_\infty)} \left(\left\| \psi_m^h \left(\prod_{i=m+1}^{\infty} \psi_i^h(0) \right) - \overline{\psi}_m^h \left(\prod_{i=m+1}^{\infty} \psi_i^h(0) \right) \right\|_\infty \right) + 4\varepsilon \|g\|_\infty \\
&\leq \sup_{i \in \mathbb{N}^*} h_i \cdot \sum_{m=1}^{N(T_\infty)} \left(C_1 (k(t_m) - k(t_{m+1})) + (1 + \varepsilon) C_2 h_m \|g\|_\infty \right) + 4\varepsilon \|g\|_\infty \\
&\leq \sup_{i \in \mathbb{N}^*} h_i \cdot \left(C_1 k(0) + (1 + \varepsilon) C_2 t_{N(T_\infty)} \|g\|_\infty \right) + 4\varepsilon \|g\|_\infty \xrightarrow{\sup_{i \in \mathbb{N}^*} h_i \to 0} 4\varepsilon \|g\|_\infty \xrightarrow{\varepsilon \to 0} 0.
\end{aligned}$$

Now, the statement of the theorem follows directly from Proposition 6. □

Remark 25. In the above proof, we used the Shapley equation to prove that if $\sup h_i$ is small, then the value $v_{T_\infty,k}$ of a stochastic game with stage duration is close to the value $v^{\texttt{cont2}}_{T_\infty,k}$ of a discretization of a continuous-time Markov game, which is known to converge when $\sup h_i \to 0$. The analogous is not done in the proof of Theorem 2, because for state-blind stochastic games the Shapley equation has a much more complicated structure. Namely, in this case instead of estimating the difference

$$\langle (Id + h_n q(i,j))(\omega,\cdot)\,, f(\cdot)\rangle - \langle (\exp\{h_n q(i,j)\})(\omega,\cdot)\,, f(\cdot)\rangle,$$

we will need to estimate the difference

$$f(t_{n+1}, \widetilde{p}_1(i,j)) - f(t_{n+1}, \widetilde{p}_2(i,j)),$$

where $\widetilde{p}_1(i,j) = p * (Id + h_n q(i,j))$ and $\widetilde{p}_2(i,j) = p * \exp\{h_n q(i,j)\}$. It is not clear how to find an appropriate estimate (which is small in comparison with $\sup h_i$).

Remark 26 (The case of $T \neq +\infty$)**.** The above proof can be simplified if we assume that $T \neq +\infty$. In this case by (21), (22), (23), (25) we have

$$\begin{aligned}
\left\|v_{T_\infty,k} - v^{\texttt{cont2}}_{T_\infty,k}\right\|_\infty &\leq \sum_{m=1}^{+\infty} \left(\left\| \psi^h_m \left(\prod_{i=m+1}^{\infty} \psi^h_i(0) \right) - \overline{\psi}^h_m \left(\prod_{i=m+1}^{\infty} \psi^h_i(0) \right) \right\|_\infty \right) \\
&\leq \sup_{i\in\mathbb{N}^*} h_i \cdot \sum_{m=1}^{+\infty} \left(C_1(k(t_m) - k(t_{m+1})) + (1+\varepsilon) C_2 h_m \|g\|_\infty \right) \\
&\leq \sup_{i\in\mathbb{N}^*} h_i \cdot \left(C_1 k(0) + (1+\varepsilon) C_2 T \|g\|_\infty \right) \xrightarrow{\sup_{i\in\mathbb{N}^*} h_i \to 0} 0.
\end{aligned}$$

8 Final comments

Remark 27 (The case of non-finite stochastic games with stage duration)**.** In this article, we assumed that the state and action spaces are finite. However, it is not necessary, and we may assume that the state space and/or the action spaces are compact metric spaces. Indeed, it is straightforward to define games with stage duration for this more general case. Under some weak conditions on the payoff and transition probability functions, there is a value for the game with stage duration, for any fixed partition T_∞.

Note that for the model of the discretization of continuous-time Markov games this is *not* the case, because it is not clear how to define the transition probabilities in the case when the state space is an (infinite) compact metric space.

Theorem 2 still holds in this more general setting, because the proof of Theorem 2 uses only the Shapley equation, and as long as the Shapley equation holds, the proof of Theorem 2 holds too.

But the proof of Theorem 1 presented above does not work, because it uses a result on the discretization of continuous-time Markov games. However, the author assumes that one may prove the generalization of Theorem 1 by following the proof of its analogue from [Sor17].

Remark 28 (Stochastic games with public signals)**.** In this article, we considered two types of games: stochastic games with perfect observation of the state and state-blind stochastic games. We may also consider an intermediate case of stochastic games with public signals. In such games players are given a public signal that depends on the current state, but they may not observe the state itself.

It is possible to give a natural definition of games with stage duration and public signals, which is done in [Nov24].

9 Acknowledgements

The author is grateful to his research advisor Guillaume Vigeral for constant attention to this work. The author is also grateful to Sylvain Sorin for useful discussions.

10 References

[And91] William J. Anderson. *Continuous-Time Markov Chains: An Applications-Oriented Approach*. Springer Series in Statistics. Springer New York, NY, 1 edition, 1991.

[BK76] Truman Bewley and Elon Kohlberg. The asymptotic theory of stochastic games. *Mathematics of Operations Research*, 1(3):197–208, 1976.

[GHL03] Xianping Guo and Onésimo Hernández-Lerma. Zero-sum games for continuous-time markov chains with unbounded transition and average payoff rates. *Journal of Applied Probability*, 40(2):327–345, 2003.

[GHL05] Xianping Guo and Onésimo Hernández-Lerma. Zero-sum continuous-time markov games with unbounded transition and discounted payoff rates. *Bernoulli*, 11(6):1009–1029, 2005.

[GHL09] Xianping Guo and Onésimo Hernández-Lerma. *Continuous-Time Markov Decision Processes: Theory and Applications*. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 1 edition, 2009.

[Isa99] Rufus Isaacs. *Differential Games: A Mathematical Theory with Applications to Warfare and Pursuit, Control and Optimization*. Dover Publications, 1999.

[LRS19] Rida Laraki, Jérôme Renault, and Sylvain Sorin. *Mathematical Foundations of Game Theory*. Universitext. Springer International Publishing, 2019.

[LS92] Arbind Kumar Lal and Sagnik Sinha. Zero-sum two-person semi-markov games. *Journal of Applied Probability*, 29(1):56–72, 1992.

[MSZ15] Jean-François Mertens, Sylvain Sorin, and Shmuel Zamir. *Repeated Games*. Econometric Society Monographs. Cambridge University Press, 2015.

[Ney13] Abraham Neyman. Stochastic games with short-stage duration. *Dynamic Games and Applications*, 3(2):236–278, 2013.

[Ney17] Abraham Neyman. Continuous-time stochastic games. *Games and Economic Behavior*, 104:92 – 130, 2017.

[Nov24] Ivan Novikov. Zero-sum stochastic games with vanishing stage duration and public signals. preprint, 2024.

[RZ20] Jérôme Renault and Bruno Ziliotto. Hidden stochastic games and limit equilibrium payoffs. *Games and Economic Behavior*, 124:122–139, 2020.

[Sha53] Lloyd Stowell Shapley. Stochastic games. *Proceedings of the National Academy of Sciences*, 39(10):1095–1100, 1953.

[Sor02] Sylvain Sorin. *A First Course on Zero-Sum Repeated Games*, volume 37 of *Mathématiques et Applications*. Springer-Verlag Berlin Heidelberg, first edition, 2002.

[Sor17] Sylvain Sorin. Limit value of dynamic zero-sum games with vanishing stage duration. *Mathematics of Operations Research*, 43(1):51–63, 2017.

[SV15] Sylvain Sorin and Guillaume Vigeral. Reversibility and oscillations in zero-sum discounted stochastic games. *Journal of Dynamics and Games*, 2(1):103–115, 2015.

[SV16] Sylvain Sorin and Guillaume Vigeral. Operator approach to values of stochastic games with varying stage duration. *International Journal of Game Theory*, 45(1):389–410, 2016.

[Vig13] Guillaume Vigeral. A zero-sum stochastic game with compact action sets and no asymptotic value. *Dynamic Games and Applications*, 3(2):172–186, 2013.

[Zac64] Lars Erik Zachrisson. *Advances in Game Theory. (AM-52), Volume 52*, pages 211–254. Princeton University Press, 1964.

[Zil16] Bruno Ziliotto. Zero-sum repeated games: Counterexamples to the existence of the asymptotic value and the conjecture maxmin $= \lim v_n$. *The Annals of Probability*, 44(2):1107 – 1133, 2016.