



**HAL**  
open science

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

Ivan Novikov

► **To cite this version:**

Ivan Novikov. Zero-Sum State-Blind Stochastic Games with Vanishing Stage Duration. 2024. hal-04401887v2

**HAL Id: hal-04401887**

**<https://hal.science/hal-04401887v2>**

Preprint submitted on 18 Jul 2024 (v2), last revised 24 Sep 2024 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Limit Value in Zero-Sum Stochastic Games with Vanishing Stage Duration and Public Signals

Ivan Novikov, Université Paris Dauphine (CEREMADE)

## Abstract

In stochastic games with stage duration  $h$ , players act at times  $0, h, 2h$ , and so on. The payoff and leaving probabilities are proportional to  $h$ . As  $h$  approaches 0, such discrete-time games approximate games played 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 examine the same question for the case of state-blind stochastic games. Our main finding is that, as  $h$  approaches 0, the value of any state-blind stochastic game with stage duration  $h$  converges to the 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

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Organization of the paper</b>	<b>3</b>
<b>3</b>	<b>Games with stage duration (perfect observation of the state)</b>	<b>4</b>
3.1	Zero-sum stochastic games with perfect observation of the state . . . . .	4
3.2	Zero-sum stochastic games with perfect observation of the state and with stage duration . . . . .	5
3.3	Theorem 1: In the case of perfect observation of the state, the uniform limit $\lim_{\sup_{i \in \mathbb{N}^*} h_i \rightarrow 0} v_{T_\infty, k}$ is the unique viscosity solution of some differential equation . . . . .	6
3.4	The proof of Theorem 1 . . . . .	7
3.4.1	Preliminaries: Shapley operator . . . . .	7
3.4.2	Preliminaries: The discretization of a zero-sum continuous-time Markov game . . . . .	7
3.4.3	The proof of Theorem 1 . . . . .	8
<b>4</b>	<b>State-blind stochastic games with stage duration</b>	<b>9</b>
4.1	The model of zero-sum state-blind stochastic games . . . . .	10
4.2	Zero-sum state-blind stochastic games with stage duration . . . . .	10
4.3	Theorem 2: in the state-blind case, the uniform limit $\lim_{\sup_{i \in \mathbb{N}^*} h_i \rightarrow 0} v_{T_\infty, k}$ is the unique viscosity solution of some differential equation . . . . .	11
4.4	The proof of Theorem 2 . . . . .	11
<b>5</b>	<b>Discounted stochastic games with stage duration</b>	<b>15</b>
<b>6</b>	<b>Final comments</b>	<b>18</b>
<b>7</b>	<b>Acknowledgements</b>	<b>19</b>
<b>8</b>	<b>References</b>	<b>19</b>

**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  $f, g : X \rightarrow \mathbb{R}$  are two functions, then

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

If  $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in X$ , then  $\langle x, y \rangle := \sum_{i=1}^n x_i y_i$ ;

- If  $X$  is a finite set,  $\zeta \in \Delta(X)$ , and  $\mu$  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 finite sets and  $g : I \times J \rightarrow \mathbb{R}$  is a function, then

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

i.e.  $\text{Val}_{I \times J}[g(i, j)]$  is the value of the one-shot zero-sum game with action spaces  $I, J$  and with payoff function  $g$ .

# 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 beginning of each stage, player 1 and player 2 observe the current state and remember the actions taken in previous stages. They then 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 state is chosen according to some probability law, depending on players' actions and the current state. For a fixed discount factor  $\lambda \in (0, 1]$ , player 1 aims to maximize the  $\lambda$ -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; conversely, player 2 aims to minimize it. Under standard assumptions, maxmin and minmax coincide, and the resulting quantity is called the value, denoted by  $v_\lambda$ .

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  $v_\lambda$  in the same way as above.

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  $\lambda$ -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 standard assumptions 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 [Sor18] considers discretizations of continuous-time Markov games. In a discretization, players can act only at specific times  $t_1, t_2, t_3, \dots$ , rather than at any moment of time, as in a standard continuous-time Markov game. However, the state variable continues to change as it would in the continuous-time model. The article [Sor18] 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 the discretization of a continuous-time Markov game converges to the 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]. They propose another way to approximate continuous-time games. Starting from a stochastic game  $\Gamma$  (with perfect observation of the state) with stage duration 1, Neyman considers a family  $\Gamma_h$  of stochastic games in which players act at time 0,  $h$ ,  $2h$ , and so on. In these games, both the payoffs and leaving probabilities are normalized at each stage, making them proportional to  $h$ . This gives the value  $v_{h,\lambda}$  depending both on the stage duration  $h$  and the discount factor  $\lambda$ . A particular case of interest is when  $h$  is small, which approximates a game played in continuous time with  $\lambda$ -discounted payoff  $\int_0^{+\infty} \lambda e^{-\lambda t} g_t dt$ .

In [Ney13], Neyman considered the asymptotics of  $v_{h,\lambda}$  when  $h$  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 the unique solution of a functional equation. Subsequently, the article [SV16] generalized some of the results from [Ney13] and obtained new findings; 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.

In this framework of games with stage duration [Ney13], we introduce more general total payoffs for a game with stage duration. When  $\sup h_i$  is small, these payoffs approximate the continuous-time game with the 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 the 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 [Sor18].

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. 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 the 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 1. The proof of Theorem 2 has the same structure as a similar result in [Sor18].

## 2 Organization of the paper

In §3.1, we provide all the necessary information about stochastic games with perfect observation of the state. In §3.2, we define stochastic games with perfect observation of the state and with stage duration, which generalize the discounted case considered in literature before. In §3.3, we state Theorem 1, which proof is given in §3.4.

In §4.1, we recall 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 state Theorem 2, which proof is given in §4.4.

In §5, we give some results about the discounted case, in both the framework of perfect observation of the state and the state-blind framework. In §6, 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 is partially based on books [MSZ15], [LRS19], and [Sor02].

Now, we give a 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:

- $\Omega$  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;
- $g_k : I \times J \times \Omega \rightarrow \mathbb{R}$  is the  $k$ -th stage payoff function of player 1;
- $P_k : I \times J \rightarrow \{\text{row-stochastic matrices } |\Omega| \times |\Omega|\}$  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  $i$ . Note that in our model we allow  $g_k$  and  $P_k$  to be dependent on the stage number  $k$ , which is not the case in the original model from [Sha53].

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  $(\omega_a, \omega_b)$ -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  $\omega_1$  is known to the players. At each stage  $n \in \mathbb{N}^*$ :

1. Players observe the current state  $\omega_n$  and remember actions of each other at the previous stage.
2. Players simultaneously 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  $x_n \in \Delta(I)$  (respectively  $y_n \in \Delta(J)$ ). They are observed by both players.
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  $\omega_{n+1}$  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 assumed to be common knowledge.

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  $(\omega_1, i_1, j_1, \omega_2, i_2, j_2, \dots, \omega_{t-1}, i_{t-1}, j_{t-1}, \omega_t)$ . 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  $\sigma : \bigcup_{t \geq 1} H_t \rightarrow \Delta(I)$  (respectively  $\tau : \bigcup_{t \geq 1} H_t \rightarrow \Delta(J)$ ). 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 an initial state  $\omega \in \Omega$ , strategies  $\sigma : \bigcup_{t \geq 1} H_t \rightarrow \Delta(I), \tau : \bigcup_{t \geq 1} H_t \rightarrow \Delta(J)$ , and the induced probability distribution  $P_{\sigma, \tau}^\omega$  on  $\Omega \times (I \times J \times \Omega)^{\mathbb{N}^*}$ , we can consider the expectation  $E_{\sigma, \tau}^\omega$  of a random variable on  $\bigcup_{t \geq 1} H_t$ .

Now, we need to choose how to compute a total payoff function. Given a stochastic game  $\Gamma = (\Omega, I, J, \{g_m\}_{m \in \mathbb{N}^*}, \{P_m\}_{m \in \mathbb{N}^*})$  and a sequence  $\{b_m\}_{m \in \mathbb{N}^*}$  with  $b_m \geq 0$  ( $m \in \mathbb{N}^*$ ) and with  $0 < \sum_{m=1}^\infty b_m < \infty$ , we consider the stochastic game  $\Gamma(\{b_m\}_{m \in \mathbb{N}^*})$  with total payoff  $E_{\sigma, \tau}^\omega(\sum_{i=1}^\infty b_i g_i)$ . (It depends on a strategy profile  $(\sigma, \tau)$  and an initial state  $\omega$ ).

Some particular cases of the above definitions have its own name. For a fixed  $T \in \mathbb{N}^*$ , the

(normalized) repeated  $T$  times game  $\Gamma^T$  is obtained if we set in the above definition  $b_m = 1/T$  for  $m = 1, 2, \dots, T$ , and  $b_m = 0$  for  $m \geq T + 1$ . For a fixed  $\lambda \in (0, 1)$ , the (normalized)  $\lambda$ -discounted<sup>1</sup> game  $\Gamma^\lambda$  is obtained if we set in the above definition  $b_m = \lambda(1 - \lambda)^{m-1}$  for all  $m \in \mathbb{N}^*$ .

Analogously to single-shot zero-sum games, we may define the value and the ( $\varepsilon$ -)optimal strategies of stochastic games. Given 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 a value  $V : \Omega \rightarrow \mathbb{R}$  if for all  $\omega \in \Omega$  we have  $V(\omega) = \sup_{\sigma} \inf_{\tau} E_{\sigma, \tau}^{\omega}(\sum_{i=1}^{\infty} b_i g_i) = \inf_{\tau} \sup_{\sigma} E_{\sigma, \tau}^{\omega}(\sum_{i=1}^{\infty} b_i g_i)$ .

Sometimes instead of the transition probability function  $P$ , we consider the *kernel*  $q : I \times J \rightarrow \{\text{matrices } |\Omega| \times |\Omega|\}$  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 Zero-sum stochastic games with perfect observation of the state and with stage duration

In this subsection, we introduce games with stage duration.

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

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

The state space  $\Omega$  and the action spaces  $I, J$  of both players are independent of the partition  $T_\infty$  in a stochastic game with  $n$ -th stage duration  $h_n$ . The payoff function and the kernel depend on  $T_\infty$ . 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 1.** Given a stochastic game  $(\Omega, I, J, g, q)$  and a partition  $T_\infty$  of  $[0, T)$ , the *stochastic game with perfect observation of the state and with  $n$ -th stage duration  $h_n$*  is the stochastic game  $(\Omega, I, J, \{h_n g\}_{n \in \mathbb{N}^*}, \{h_n q\}_{n \in \mathbb{N}^*})$ .

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

**Definition 2.** The *stochastic game with perfect observation of the state and with  $n$ -th stage duration  $h_n$ , weight function  $k(t)$ , initial time  $t_n \in T_\infty$ , and initial state  $\omega$* , is the stochastic game  $(\Omega, I, J, \{h_n g\}_{n \in \mathbb{N}^*}, \{h_n q\}_{n \in \mathbb{N}^*})$  with total payoff  $G_{T_\infty, k}(t_n, \omega) := E_{\sigma, \tau}^{\omega, t_n} \left( \sum_{i=n}^{\infty} h_i k(t_i) g_i \right)$ . Denote by  $v_{T_\infty, k}(t_n, \omega)$  the value of the game with such a total payoff.

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

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

We are interested in the behavior of such games when the duration of each stage is vanishing, i.e. we want to know what happens when  $\sup h_i \rightarrow 0$ .

**Remark 1** (Why such a definition). If all  $h_i$  are small, then the value of a stochastic game with

---

<sup>1</sup>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  $\lambda^{i-1}$ .

$n$ -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 Propositions 2 and 3 in §6.

**Example 1** (Discounted games). If initial time is 0,  $T = +\infty$ , and  $k(t) = \lambda \exp\{-\lambda t\}$ , we obtain the normalized  $\lambda$ -discounted stochastic game with stage duration. We discuss it in more details in §5.

**Example 2** (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  $n = 1, \dots, T$ , and  $h_n = 0$  for  $n > T$ , then we recover the case of normalized repeated  $T$  times games with total payoff  $\frac{1}{T} \sum_{i=1}^T g_i$ .

**Remark 2.** 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  $T_\infty$  with  $\sup_{i \in \mathbb{N}^*} h_i < \delta$ , we have

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

### 3.3 Theorem 1: In the case of perfect observation of the state, the uniform limit $\lim_{\sup_{i \in \mathbb{N}^*} h_i \rightarrow 0} v_{T_\infty, k}$ is the unique viscosity solution of some differential equation

First, we recall the definition of a viscosity solution.

**Definition 3.** Given a function  $H : [0, T) \times \Omega \times \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ , a function  $u : [0, T) \times \Omega \rightarrow \mathbb{R}$  is called a *viscosity solution* of the differential equation

$$0 = \frac{d}{dt} u(t, \omega) + H(t, \omega, u(t, \omega), \nabla u(t, \omega))$$

if:

1. for any  $\mathcal{C}^1$  function  $\psi : [0, T) \times \Omega \rightarrow \mathbb{R}$  with  $u - \psi$  having a strict local maximum at  $(\bar{t}, \bar{\omega}) \in [0, T) \times \Omega$  we have  $0 \leq \frac{d}{dt} \psi(\bar{t}, \bar{\omega}) + H(\bar{t}, \bar{\omega}, \psi(\bar{t}, \bar{\omega}), \nabla \psi(\bar{t}, \bar{\omega}))$ ;
2. for any  $\mathcal{C}^1$  function  $\psi : [0, T) \times \Omega \rightarrow \mathbb{R}$  with  $u - \psi$  having a strict local minimum at  $(\bar{t}, \bar{\omega}) \in [0, T) \times \Omega$  we have  $0 \geq \frac{d}{dt} \psi(\bar{t}, \bar{\omega}) + H(\bar{t}, \bar{\omega}, \psi(\bar{t}, \bar{\omega}), \nabla \psi(\bar{t}, \bar{\omega}))$ .

**Theorem 1.** If  $(\Omega, I, J, g, q)$  is a finite stochastic game with perfect observation of the state, then the uniform limit  $\lim_{\substack{\sup_{i \in \mathbb{N}^*} h_i \rightarrow 0 \\ h_1 + h_2 + \dots = T}} v_{T_\infty, k}$  exists and is the unique viscosity solution of the differential equation (in

$v(t, \omega)$ )

$$0 = \frac{d}{dt} v(t, \omega) + \mathbf{Val}_{I \times J} [k(t)g(i, j, \omega) + \langle q(i, j)(\omega, \cdot), v(t, \cdot) \rangle]. \quad (1)$$

**Remark 3.** For the discounted case  $k(t) = \lambda e^{-\lambda t}$ , (1) transforms into

$$\lambda v(\omega) = \mathbf{Val}_{I \times J} [\lambda g(i, j, \omega) + \langle q(i, j)(\omega, \cdot), v(\cdot) \rangle]. \quad (2)$$

(make a substitution  $v(t, \omega) \mapsto e^{-\lambda t} v(\omega)$ ). For this particular case, Theorem 1 is already known. The result [Ney13, Theorem 1] proves it for the uniform case in which  $h_n = h$  for all  $n$ , while the result [SV16, Corollary 7.1] proves it for the general case.

**Remark 4.** The above theorem is similar to [Sor18, Proposition 4.3], which concerns a similar result in another model of games with stage duration.

## 3.4 The proof of Theorem 1

### 3.4.1 Preliminaries: Shapley operator

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

Fix an initial time  $n \in \mathbb{N}^*$ . We consider a finite stochastic game  $G_n = (\Omega, I, J, \{g_m\}_{m \in \mathbb{N}^*}, \{P_m\}_{m \in \mathbb{N}^*})$  with perfect observation of the state, where  $\{P_m\}_{m \in \mathbb{N}^*}$  are transitional probability functions and  $\sum_{k=1}^{\infty} \|g_k\|_{\infty} < \infty$ , and with total payoff  $E_{\sigma, \tau}^{\omega}(\sum_{i=n}^{\infty} g_i)$ . Denote by  $v_n(\omega)$  the value of such a game.

Denote by  $C(\Omega, \mathbb{R})$  the set of continuous functions from  $\Omega$  to  $\mathbb{R}$ . For a sequence of maps  $S_1, S_2, \dots$  from a Banach space  $C$  to itself, for any  $z \in C$ , we denote  $\prod_{i=1}^{\infty} S_i(z) := \lim_{n \rightarrow \infty} (S_1 \circ S_2 \circ \dots \circ S_n(z))$ .

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

$$\psi_n : C(\Omega, \mathbb{R}) \rightarrow C(\Omega, \mathbb{R}), \quad f(\omega) \mapsto \mathbf{val}_{I \times J} [g_n(i, j, \omega) + \langle P_n(i, j)(\omega, \cdot), f(\cdot) \rangle].$$

Then:

1. For each  $n \in \mathbb{N}^*$  the operator  $\psi_n$  is nonexpansive, i.e.  $\|\psi_n(f - g)\|_{\infty} \leq \|f - g\|_{\infty}$ ;
2. If  $\bar{G}_n = (\Omega, I, J, \{\bar{g}_m\}_{m \in \mathbb{N}^*}, \{\bar{P}_m\}_{m \in \mathbb{N}^*})$  is a finite stochastic game which has the same properties as the game  $G_n$ , and  $\bar{\psi}_n$  is the operator

$$\bar{\psi}_n : C(\Omega, \mathbb{R}) \rightarrow C(\Omega, \mathbb{R}), \quad f(\omega) \mapsto \mathbf{val}_{I \times J} [\bar{g}_n(i, j, \omega) + \langle \bar{P}_n(i, j)(\omega, \cdot), f(\cdot) \rangle],$$

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

3. The value  $v_n(\omega)$  of the game  $G_n$  exists and is equal to  $\prod_{i=n}^{\infty} \psi_i(0)$ . (In particular, such a product is well-defined);
4. For any  $n \in \mathbb{N}^*$  and any  $\omega \in \Omega$  we have  $v_n(\omega) = (\psi_n v_{n+1})(\omega)$ ;
5. For any  $n \in \mathbb{N}^*$ , there exists an optimal Markov strategy in the game  $G_n$ .

An analogue of this proposition for  $\lambda$ -discounted games and finitely repeated games is well-known, see, for example, [MSZ15, Theorem IV.3.2, Proposition IV.3.3]. The more general case given here (for a more general total payoff  $\sum_{i=n}^{\infty} g_i$ ) can be proved analogously.

**Remark 5.** (cf. [MSZ15]). We considered the case of a finite game  $G$ , but analogous proposition holds also in a more general setting, for example, if:

1.  $I$  and  $J$  are finite;
2.  $\Omega$  is a compact metric space;
3.  $g$  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  $\Omega$ , 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  $\Omega$ .

### 3.4.2 Preliminaries: The discretization of a zero-sum continuous-time Markov game

Now, we briefly recall the model from [Sor18]. As before,  $T$  is either a positive number or  $+\infty$ , and  $T_{\infty}$  is a partition of  $[0, T]$ ;  $h_n = t_{n+1} - t_n$  for each  $n \in \mathbb{N}^*$ .

We consider a continuous-time Markov game  $(\Omega, I, J, g, q)$ , in which players are only allowed to act at times  $t_1, t_2, t_3, \dots$ . At any time  $t \in [t_i, t_{i+1})$  players act according to their decision at time  $t_i$ . In the above game  $\Omega$  is a finite state space,  $I$  and  $J$  are finite sets of actions of players,  $g : I \times J \times \Omega \rightarrow \mathbb{R}$  is an instantaneous payoff function, and  $q : I \times J \rightarrow \{|\Omega| \times |\Omega| \text{ matrices } A \text{ such that } a_{ij} \geq 0 \forall i \neq j, a_{ii} \leq 0 \forall i, \text{ and } \sum_{j=1}^{|\Omega|} a_{ij} = 0\}$  is an infinitesimal generator.

We fix a nonincreasing continuous function  $k : [0, T] \rightarrow \mathbb{R}_+$  with  $\int_0^T k(t) dt = 1$ . The *discretization*



of a continuous-time Markov game  $(\Omega, I, J, g, q)$ , in which player acts at times  $t_1, t_2, \dots$  is the stochastic game  $(\Omega, I, J, \{g_m\}_{m \in \mathbb{N}^*}, \{P_m\}_{m \in \mathbb{N}^*})$  with perfect observation of the state, where  $P_m$  is the  $m$ -th stage transition probability function, and

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

Given two strategies  $\sigma, \tau$  of players, an initial state  $\omega$ , an initial time  $t_n \in T_\infty$ , the *total payoff* is  $G_{T_\infty, k}^{\text{cont}}(t_n, \omega) := E_{\sigma, \tau}^{\omega, t_n} \left( \int_{t_n}^T k(t)g(i_t, j_t, \omega_t)dt \right)$ . The game is said to have a *value*  $v_{T_\infty, k}^{\text{cont}}(t_n, \omega)$  if  $v_{T_\infty, k}^{\text{cont}}(t_n, \omega) = \sup_\sigma \inf_\tau G_{T_\infty, k}^{\text{cont}}(t_n, \omega) = \inf_\tau \sup_\sigma G_{T_\infty, k}^{\text{cont}}(t_n, \omega)$ . We can define  $G_{T_\infty, k}^{\text{cont}}(t, \omega)$  and  $v_{T_\infty, k}^{\text{cont}}(t, \omega)$  for any  $t \in [0, T)$  via linearity.

The uniform limit  $\lim_{\substack{\sup_{i \in \mathbb{N}^*} h_i \rightarrow 0 \\ h_1 + h_2 + \dots = T}} v_{T_\infty, k}^{\text{cont}}$  is the unique viscosity solution of the differential equation  $0 = \frac{d}{dt}v(t, \omega) + \mathbf{Val}_{I \times J}[k(t)g(i, j, \omega) + \langle q(i, j)(\omega, \cdot), v(t, \cdot) \rangle]$ , see [Sor18, Proposition 4.3].

### 3.4.3 The proof of Theorem 1

The proof given below is a generalization of [SV16, proofs of Lemma 8.1 and Proposition 8.1].

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

$$\begin{aligned} \psi_n^h &: C(\Omega, \mathbb{R}) \rightarrow C(\Omega, \mathbb{R}), \\ f(\omega) &\mapsto \mathbf{Val}_{I \times J}[k(t_n)h_n g(i, j, \omega) + \langle (Id + h_n q(i, j))(\omega, \cdot), f(\cdot) \rangle]; \\ \bar{\psi}_n^h &: C(\Omega, \mathbb{R}) \rightarrow C(\Omega, \mathbb{R}), \\ f(\omega) &\mapsto \mathbf{Val}_{I \times J} \left[ \int_{t_n}^{t_{n+1}} k(t)g(i, j, \omega)dt + \langle \exp\{h_n q(i, j)\}(\omega, \cdot), f(\cdot) \rangle \right]. \end{aligned}$$

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

$$\left\| \psi_n^h(f) - \bar{\psi}_n^h(f) \right\|_\infty \leq \left| k(t_n)h_n - \int_{t_n}^{t_{n+1}} k(t)dt \right| \|g\|_\infty + \|(Id + h_n q) - \exp\{h_n q\}\|_1 \cdot \|f\|_\infty. \quad (3)$$

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

$$\frac{1}{h_n} \left| k(t_n)h_n - \int_{t_n}^{t_{n+1}} k(t)dt \right| = |k(t_n) - k(c)| \leq k(t_n) - k(t_{n+1}). \quad (4)$$

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

$$\frac{1}{h_n} \|(Id + h_n q) - \exp\{h_n q\}\|_1 = \frac{1}{h_n} \left\| (Id + h_n q) - \sum_{k=0}^{\infty} \frac{(h_n q)^k}{k!} \right\|_1 \leq \frac{1}{h_n} C h_n^2 = C h_n. \quad (5)$$

By combining (3)-(5), 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

$$\frac{1}{h_n} \left\| \psi_n^h(f) - \bar{\psi}_n^h(f) \right\|_\infty \leq C_1 (k(t_n) - k(t_{n+1})) + C_2 h_n \|f\|_\infty. \quad (6)$$

By (6) and Proposition 1(1,3) we have

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

$$\leq \left\| \psi_1^h \left( \prod_{i=2}^{\infty} \psi_i^h(0) \right) - \bar{\psi}_1^h \left( \prod_{i=2}^{\infty} \psi_i^h(0) \right) \right\|_{\infty} + \left\| \prod_{i=2}^{\infty} \psi_i^h(0) - \prod_{i=2}^{\infty} \bar{\psi}_i^h(0) \right\|_{\infty}.$$

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

$$\|v_{T_{\infty},k} - v_{T_{\infty},k}^{\text{cont}}\|_{\infty} \leq \sum_{m=1}^N \left( \left\| \psi_m^h \left( \prod_{i=m+1}^{\infty} \psi_i^h(0) \right) - \bar{\psi}_m^h \left( \prod_{i=m+1}^{\infty} \psi_i^h(0) \right) \right\|_{\infty} \right) + \left\| \prod_{i=N+1}^{\infty} \psi_i^h(0) - \prod_{i=N+1}^{\infty} \bar{\psi}_i^h(0) \right\|_{\infty}. \quad (7)$$

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

$$v_{T_{\infty},k}(t_m, \cdot) = \prod_{i=m}^{\infty} \psi_i^h(0) \text{ and } v_{T_{\infty},k}^{\text{cont}}(t_m, \cdot) = \prod_{i=m}^{\infty} \bar{\psi}_i^h(0). \quad (8)$$

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 2)

$$\|v_{T_{\infty},k}(t_m, \cdot)\|_{\infty} \leq \|g\|_{\infty} \left( \int_S^T k(t) dt + \varepsilon \right) \leq 2\varepsilon \|g\|_{\infty} \quad \text{and} \quad \|v_{T_{\infty},k}^{\text{cont}}(t_m, \cdot)\|_{\infty} \leq 2\varepsilon \|g\|_{\infty}. \quad (9)$$

Also, if  $\sup_{i \in \mathbb{N}^*} h_i$  is small enough, then by Remark 2 we have for any  $t_m$

$$\|v_{T_{\infty},k}(t_m, \cdot)\|_{\infty} \leq (1 + \varepsilon) \|g\|_{\infty}. \quad (10)$$

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

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  $T_{\infty}$  of  $[0, T)$ . By (6)–(10), there exist  $C_1 > 0, C_2 > 0$  such that

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

Now, the statement of the theorem follows directly from [Sor18, Proposition 4.3].  $\square$

**Remark 6** (The case of  $T \neq +\infty$ ). The above proof can be simplified if we assume that  $T \neq +\infty$ . In this case by (6), (7), (8), (10) we have

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

## 4 State-blind stochastic games with stage duration

In this section, we consider stochastic games in which players cannot observe the current state. In §4.1, we recall the 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 games with stage duration in this framework. In §4.3, we give a new result showing that in

the case of state-blind stochastic games, the uniform limit  $\lim_{\sup_{i \in \mathbb{N}^*} h_i \rightarrow 0} v_{T_\infty, k}$  exists and is the unique viscosity solution of some partial differential equation. The proof of this result is given in §4.4.

## 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 only remember the actions of players at the previous stage and the initial probability distribution on the states.

We still assume that the state space  $\Omega$  and the action spaces  $I, J$  are finite.

We can define the strategies,  $\lambda$ -discounted and finitely repeated games, total payoffs, values, in the same way as in §3.1. In particular, a strategy of player 1 is a 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. It is known that the value exists if  $\Omega, I, J$  are finite, see [Sor02].

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

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

Now let us define  $\{P_n^\gamma\}_{n \in \mathbb{N}^*}$ . 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  $\omega'$  is equal to  $P_n(i, j)(p, \omega') := \sum_{\omega \in \Omega} p(\omega) P_n(i, j)(\omega, \omega')$ . The function  $P_n^\gamma : I \times J \rightarrow \Delta(\Omega) \times \Delta(\Omega)$  is defined by

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

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

**Definition 4.** 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  $+\infty$ , and  $T_\infty$  is a partition of  $[0, T)$ ;  $h_n = t_{n+1} - t_n$  for each  $n \in \mathbb{N}^*$ .

**Definition 5.** Fix a zero-sum state-blind stochastic game  $(\Omega, I, J, g, q)$ , where  $q$  is the kernel. The *state-blind stochastic game with  $n$ -th stage duration  $h_n$*  is the state-blind stochastic game

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

We can define a total payoff for games with stage duration as in §3.2.

**Remark 7** (Why state-blind case is more difficult than the case of perfect observation of the state?). 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 no longer holds.

### 4.3 Theorem 2: in the state-blind case, the uniform limit $\lim_{\sup_{i \in \mathbb{N}^*} h_i \rightarrow 0} v_{T_\infty, k}$ is the unique viscosity solution of some differential equation

The following theorem is an analogue of Theorem 1 for state-blind stochastic games.

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

$$0 = \frac{d}{dt} v(t, p) + \text{Val}_{I \times J} [k(t)g(i, j, p) + \langle p * q(i, j), \nabla v(t, p) \rangle]. \quad (11)$$

For the discounted case the above equation is more simple:

**Corollary 1.** If  $(\Omega, I, J, g, q)$  is a finite state-blind stochastic game, then the uniform limit  $\lim_{\substack{\sup_{i \in \mathbb{N}^*} h_i \rightarrow 0 \\ h_1 + h_2 + \dots = +\infty}} v_{T_\infty, \lambda}(p)$  exists and is the unique viscosity solution of the partial differential equation

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

*Proof.* In (11), make a substitution  $v(t, p) \mapsto e^{-\lambda t} v(p)$ . □

**Remark 8.** The above theorem is similar to [Sor18, Proposition 5.3], which concerns another model of games with stage duration.

### 4.4 The proof of Theorem 2

Let us first state a technical lemma.

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

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

**Lemma 1.** The family  $\{v_{T_\infty, k}(t, p)\}_{T_\infty}$  is equilipschitz-continuous and equibounded for all partitions  $T_\infty$  with  $\sup h_i$  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

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

Lemma 1 is an analogue of [Sor18, Proposition 3.11], while the deduction of Theorem 2 from it is an analogue of [Sor18, Proposition 3.12]. The proof of Lemma 1 is different from the proof of its analogue, while the deduction of Theorem 2 from it is almost identical to [Sor18, proof of Proposition 3.12].

*Proof of Lemma 1.* First, we prove equiboundedness. By Remark 2 (with  $\varepsilon = 1$ ) we have for  $\sup h_i$  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  $P^h(x, y) := Id + hq(x, y)$ . For all  $p \in \Delta(\Omega), \omega \in \Omega, x \in \Delta(I), y \in \Delta(J)$ , denote  $\widehat{p}^h(x, y)(\omega) := (p * P^h(x, y))(\omega)$ . 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  $h \in (0, 1]$ ,  $x \in \Delta(I)$ ,  $y \in \Delta(J)$

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

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

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

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

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

and since  $\|(P^h(x, y))^T\|_{\text{op}} = 1$ , we have

$$\|\widehat{p}_1^h(x, y) - \widehat{p}_2^h(x, y)\|_1 \leq \|p_1 - p_2\|_1. \quad (12)$$

Fix  $t_n \in T_\infty$ , and  $p_1, p_2 \in \Delta(\Omega)$ . By Proposition 1(5) and Remark 5 there exists<sup>2</sup> a profile of optimal Markov strategies  $(\sigma_1, \tau_1)$  (respectively there exists a profile of optimal Markov strategies  $(\sigma_2, \tau_2)$ ), if initial time is  $t_n$  and initial distribution of states is  $p_1$  (respectively  $p_2$ ). For a strategy  $(\sigma_i, \tau_i)$  ( $i = 1, 2$ ), 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  $p_i^j$  and on the stage number  $j$ ). We have

$$|v_{T_\infty, k}(t_n, p_1) - v_{T_\infty, k}(t_n, p_2)| \leq \sum_{j=n}^{\infty} h_j k(t_j) |g(x_1^j, y_1^j, p_1^j) - g(x_2^j, y_2^j, p_2^j)|, \quad (13)$$

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

$$\begin{aligned} |g(x_1^j, y_1^j, p_1^j) - g(x_2^j, y_2^j, p_2^j)| &\leq |g(x_1^j, y_2^j, p_1^j) - g(x_1^j, y_2^j, p_2^j)| = \left| \sum_{\omega \in \Omega} p_1^j(\omega) g(x_1^j, y_2^j, \omega) - \sum_{\omega \in \Omega} p_2^j(\omega) g(x_1^j, y_2^j, \omega) \right| \\ &\leq \|g\|_\infty \sum_{\omega \in \Omega} |p_1^j(\omega) - p_2^j(\omega)| = \|g\|_\infty \|p_1^j - p_2^j\|_1 \leq \|g\|_\infty \|p_1 - p_2\|_1, \end{aligned} \quad (14)$$

where the last inequality follows from (12). By Remark 2 (with  $\varepsilon = 1$ ) and by combining (13) and (14), we obtain

$$|v_{T_\infty, k}(t_n, p_1) - v_{T_\infty, k}(t_n, p_2)| \leq \|g\|_\infty \|p_1 - p_2\|_1 \sum_{j=n}^{\infty} h_j k(t_j) \leq 2 \|g\|_\infty \|p_1 - p_2\|_1, \quad (15)$$

where the second inequality holds if is small enough.

Fix  $p \in \Delta(\Omega)$ , and  $t_n, t_m \in T_\infty$  with  $t_m \geq t_n$ . Let  $(\sigma_1, \tau_1)$  (respectively  $(\sigma_2, \tau_2)$ ) be a profile of optimal Markov strategies, if starting time is  $t_n$  (respectively  $t_m$ ) and initial distribution of states is  $p$ . For a strategy  $(\sigma_i, \tau_i)$ , denote by  $(x_i^j, y_i^j)$  the profile of mixed actions played at the  $j$ -th stage (it

---

<sup>2</sup>Here, we consider the state-blind stochastic game as a stochastic game with perfect observation of the state, on the compact state space  $\Delta(\Omega)$ .

depends on the current distribution  $p_i^j$  and on the stage number  $j$ ). We have

$$\begin{aligned} |v_{T_\infty, k}(t_n, p) - v_{T_\infty, k}(t_m, p)| &\leq \sum_{j=n}^{m-1} h_j k(t_j) |g(x_1^j, y_1^j, p_1^j)| + \sum_{j=m}^{\infty} h_j k(t_j) |g(x_1^j, y_1^j, p_1^j) - g(x_2^j, y_2^j, p_2^j)| \\ &\leq \|k\|_\infty \|g\|_\infty |t_n - t_m| + \sum_{j=m}^{\infty} h_j k(t_j) |g(x_1^j, y_1^j, p_1^j) - g(x_2^j, y_2^j, p_2^j)|, \end{aligned} \quad (16)$$

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

$$\|p_2^m - p_1^m\|_1 = \left\| p - \left( \prod_{j=n}^{m-1} (1 - h_j) p + \left( 1 - \prod_{j=n}^{m-1} (1 - h_j) \right) \bar{p} \right) \right\|_1. \quad (17)$$

By Lemma 3(3) below we have

$$\begin{aligned} \left\| p - \left( \prod_{j=n}^{m-1} (1 - h_j) p + \left( 1 - \prod_{j=n}^{m-1} (1 - h_j) \right) \bar{p} \right) \right\|_1 &\leq \left\| \left( 1 - \prod_{j=n}^{m-1} (1 - h_j) \right) p \right\|_1 + \left\| \left( 1 - \prod_{j=n}^{m-1} (1 - h_j) \right) \bar{p} \right\|_1 \\ &\leq \left\| \left( 1 - \left( 1 - \sum_{j=n}^{m-1} h_j \right) \right) p \right\|_1 + \left\| \left( 1 - \left( 1 - \sum_{j=n}^{m-1} h_j \right) \right) \bar{p} \right\|_1 \\ &\leq |t_n - t_m| (\|p\|_1 + \|\bar{p}\|_1) = 2|t_n - t_m|. \end{aligned} \quad (18)$$

By combining (17) and (18), we have

$$\|p_2^m - p_1^m\|_1 \leq 2|t_n - t_m|. \quad (19)$$

By Remark 2 (with  $\varepsilon = 1$ ) and by combining (12), (16), and (19), we have

$$\begin{aligned} |v_{T_\infty, k}(t_n, p) - v_{T_\infty, k}(t_m, p)| &\leq \|k\|_\infty \|g\|_\infty |t_n - t_m| + 2|t_n - t_m| \sum_{j=n}^{\infty} h_j k(t_j) \\ &\leq \|k\|_\infty \|g\|_\infty |t_n - t_m| + 4|t_n - t_m| \\ &= (\|k\|_\infty \|g\|_\infty + 4)|t_n - t_m|. \end{aligned} \quad (20)$$

By combining (15) and (20), we have for any  $p_1, p_2 \in \Delta(\Omega)$ ,  $t^1 = t_n \in T_\infty$ , and  $t^2 = t_m \in T_\infty$

$$\begin{aligned} |v_{T_\infty, k}(t^1, p_1) - v_{T_\infty, k}(t^2, p_2)| &\leq |v_{T_\infty, k}(t^1, p_1) - v_{T_\infty, k}(t^1, p_2)| \\ &\quad + |v_{T_\infty, k}(t^1, p_2) - v_{T_\infty, k}(t^2, p_2)| \\ &\leq 2\|g\|_\infty \|p_1 - p_2\|_1 + (\|k\|_\infty \|g\|_\infty + 4)|t^1 - t^2|. \end{aligned} \quad (21)$$

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 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

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

We have

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

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

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

If  $\alpha \geq \beta$ , then  $t_{n+1} \geq t_{m+1}$  and  $t_n \geq t_m$ , and a computation similar to the above one shows that  $|v_{T_\infty, k}(t^1, p_2) - v_{T_\infty, k}(t^2, p_2)| \leq C_2 |t^2 - t^1|$ . Thus we proved (21) for any  $t^1, t^2 \in [0, T]$ , and hence we proved the lemma.  $\square$

Denote

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

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

$$v_{T_\infty, k}(t_n, p) = \mathbf{Val}_{I \times J} [h_n k(t_n) g(i, j, p) + v_{T_\infty, k}(t_{n+1}, \tilde{p}(i, j))], \quad (22)$$

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

By Lemma 1 and the Arzelà–Ascoli theorem we have  $U \neq \emptyset$ . We are going to prove that any  $\mathcal{U} \in U$  is a viscosity solution of the partial differential equation

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

Let  $\psi(t, p)$  be a  $\mathcal{C}^1$  function such that  $\mathcal{U} - \psi$  has a strict local maximum at  $(\bar{t}, \bar{p}) \in [0, T] \times \Delta(\Omega)$ . Consider a sequence of partitions  $\{T_\infty(m)\}_{m \in \mathbb{N}^*}$  such that in the partition  $T_\infty(m)$  we have  $\sup_{i \in \mathbb{N}^*} h_i \rightarrow 0$  as  $m \rightarrow \infty$ , and such that a sequence  $W_m = v_{T_\infty(m), \lambda}$  converging uniformly to  $\mathcal{U}$  as  $m \rightarrow \infty$ , and let  $(t^*(m), p(m))$  be a maximizing (locally near  $(\bar{t}, \bar{p})$ ) sequence for  $(W_m - \psi)(t, p)$ , where  $t^*(m) \in T_\infty(m)$ . In particular,  $(t^*(m), p(m))$  converges to  $(\bar{t}, \bar{p})$  as  $m \rightarrow \infty$ . Given an optimal in (22) mixed strategy  $x^*(m) \in \Delta(I)$ , one has with  $t^*(m) = t_n \in T_\infty(m)$

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

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

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

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

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

where  $o(h_n)/h_n \rightarrow 0$  when  $h_n \rightarrow 0$ . 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  $m \rightarrow \infty$ , one obtains, for some accumulation point  $x^* \in \Delta(I)$  (we use again the continuity of  $k$  and  $\psi$  being  $\mathcal{C}^1$ )

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

Analogously one can prove that if  $\psi(t, p)$  is a  $\mathcal{C}^1$  function such that  $\mathcal{U} - \psi$  has a strict local minimum at  $(\bar{t}, \bar{p}) \in [0, T) \times \Delta(\Omega)$ , and  $y^* \in \Delta(J)$  is optimal in (22), then

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

Thus  $\mathcal{U}$  is a viscosity solution of  $0 = \frac{d}{dt} V(t, p) + \text{Val}_{I \times J} [k(t)g(i, j, p) + \langle p * q(i, j), \nabla V(t, p) \rangle]$ . The uniqueness follows from [Sor18, Proposition 3.9].  $\square$

**Remark 9.** In the proof of Theorem 1, 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  $v_{T_\infty, k}^{\text{cont}}$  of the discretization of a continuous-time Markov game, which is known to converge when  $\sup h_i \rightarrow 0$ . The analogous was 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 need to estimate the difference  $f(t_{n+1}, \tilde{p}_1(i, j)) - f(t_{n+1}, \tilde{p}_2(i, j))$ , where  $\tilde{p}_1(i, j) = p * (Id + h_n q(i, j))$  and  $\tilde{p}_2(i, j) = p * \exp \{h_n q(i, j)\}$ . It is not clear how to find an appropriate estimate (which is small in comparison with  $\sup h_i$ ).

## 5 Discounted stochastic games with stage duration

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

In this section, we consider in more details the discounted case. In previous sections, we said that in a discounted game with  $n$ -th stage duration  $h_n$ , the total payoff is  $\lambda \sum_{i=1}^{\infty} \exp\{-\lambda t_i\} h_i g_i$ . However, previous articles [Ney13] and [SV16] about games with stage duration considered the total payoff  $\lambda \sum_{i=1}^{\infty} \left( \prod_{j=1}^{i-1} (1 - \lambda h_j) \right) h_i g_i$ . We did not use this total payoff because it is not a particular case of Definition 2. This section shows that if  $\sup h_i \rightarrow 0$ , then these two total payoffs are essentially equivalent.

For each  $h \in (0, 1]$ , consider a discount factor  $\alpha_h \in [0, 1)$ . We want to impose some natural condition on  $\alpha_h$ , 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 6.** (cf. [Ney13, p. 240]). A family of  $h$ -dependent discount factors  $\alpha_h$  is called *admissible* if  $\lim_{h \rightarrow 0^+} \frac{\alpha_h}{h}$  exists. The limit is called the *asymptotic discount rate*.

**Example 3** (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  $\lambda \in (0, +\infty)$ ).

Now, we define discounted games with stage duration.

**Definition 7** ( $\lambda$ -discounted stochastic game with stage duration). Let  $\alpha_h$  be an admissible family of discount factors with asymptotic discount rate  $\lambda$ .



The  $\lambda$ -discounted stochastic game with  $n$ -th stage duration  $h_n$  is the stochastic game with payoff  $\sum_{i=1}^{\infty} \left( \prod_{j=1}^{i-1} (1 - \alpha_{h_j}) \right) h_i g_i$ . Denote by  $V_{T_{\infty}, \lambda}^{\alpha}$  the value of such a game, and denote by  $v_{T_{\infty}, \lambda}^{\alpha} = V_{T_{\infty}, \lambda}^{\alpha} / \sum_{i=1}^{\infty} \left( \prod_{j=1}^{i-1} (1 - \alpha_{h_j}) \right) h_i$  its normalization.

**Remark 10.** In the above definition:

1. If  $\alpha_h = 1 - e^{-\lambda h}$ , then  $\sum_{i=1}^{\infty} \left( \prod_{j=1}^{i-1} (1 - \alpha_{h_j}) \right) h_i g_i = \sum_{i=1}^{\infty} \left( \prod_{j=1}^{i-1} \exp\{-\lambda h_j\} \right) h_i g_i = \sum_{i=1}^{\infty} \exp\{-\lambda t_i\} h_i g_i$ ;
2. If  $\alpha_h = \lambda h$ , then  $\sum_{i=1}^{\infty} \left( \prod_{j=1}^{i-1} (1 - \alpha_{h_j}) \right) h_i g_i = \sum_{i=1}^{\infty} \left( \prod_{j=1}^{i-1} (1 - \lambda h_j) \right) h_i g_i$ , 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 3(2) for a proof).

**Lemma 2.** Fix  $\lambda \in (0, +\infty)$ . Consider two admissible families  $\alpha_h$  and  $\beta_h$  of discount factors with asymptotic discount rate  $\lambda$ . Then for all families (parametrized by partitions  $T_{\infty}$  of  $\mathbb{R}_+$ ) of streams  $x(T_{\infty}) = (g_1, g_2, \dots)$  with  $|g_i| \leq C h_i$  ( $i \in \mathbb{N}^*$ ), 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 2, we first state and prove the following simple lemma.

**Lemma 3.** 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.  $\sum_{i=1}^{\infty} h_i = +\infty \Rightarrow \prod_{i=1}^{\infty} (1 - \lambda h_i) = 0$ ;
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

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

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 immediate. 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

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

We are ready to prove Lemma 2.

*Proof of Lemma 2.* Fix a family of streams  $x(H_{\infty}) = (g_1, g_2, \dots)$  with  $|g_i| \leq C h_i$  ( $i \in \mathbb{N}^*$ ), and assume that  $\alpha_h$  and  $\beta_h$  are two families of discount factors with asymptotic discount rate  $\lambda$ . In that case we

have  $1 - \alpha_h = 1 - \lambda h + m(h)$  and  $1 - \beta_h = 1 - \lambda h + n(h)$ , where  $m(h)/h \rightarrow 0$  and  $n(h)/h \rightarrow 0$  as  $h \rightarrow 0$ . Consider the sum

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

where  $k_i(h_j)$  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  $H_{\infty}$  with  $0 < h_i \leq h'$  we have by the second assertion of Lemma 3

$$\begin{aligned} \sum_{i=1}^{\infty} \left( \prod_{j=1}^{i-1} (1 - \lambda h_j + k_i(h_j)) \right) h_i &\leq \sum_{i=1}^{\infty} \left( \prod_{j=1}^{i-1} \left[ 1 - \left( \lambda - \sup_{j \in \mathbb{N}} \left\{ \frac{|m(h_j)|}{h_j}, \frac{|n(h_j)|}{h_j} \right\} \right) h_j \right] \right) h_i \\ &= \frac{1}{\lambda - \sup_{j \in \mathbb{N}} \left\{ \frac{|m(h_j)|}{h_j}, \frac{|n(h_j)|}{h_j} \right\}}; \\ \sum_{i=1}^{\infty} \left( \prod_{j=1}^{i-1} (1 - \lambda h_j + k_i(h_j)) \right) h_i &\geq \sum_{i=1}^{\infty} \left( \prod_{j=1}^{i-1} \left[ 1 - \left( \lambda + \sup_{j \in \mathbb{N}} \left\{ \frac{|m(h_j)|}{h_j}, \frac{|n(h_j)|}{h_j} \right\} \right) h_j \right] \right) h_i \\ &= \frac{1}{\lambda + \sup_{j \in \mathbb{N}} \left\{ \frac{|m(h_j)|}{h_j}, \frac{|n(h_j)|}{h_j} \right\}}. \end{aligned}$$

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

$$\begin{aligned} &\left| \sum_{i=1}^{\infty} \left( \prod_{j=1}^{i-1} (1 - \alpha_j) \right) g_i - \sum_{i=1}^{\infty} \left( \prod_{j=1}^{i-1} (1 - \beta_j) \right) g_i \right| \\ &\leq C \sum_{i=1}^{\infty} \left| \prod_{j=1}^{i-1} (1 - \lambda h_j + m(h_j)) h_i - \prod_{j=1}^{i-1} (1 - \lambda h_j + n(h_j)) h_i \right| \\ &= C \sum_{i=1}^{\infty} \left( \max \left\{ \prod_{j=1}^{i-1} (1 - \lambda h_j + m(h_j)) h_i, \prod_{j=1}^{i-1} (1 - \lambda h_j + n(h_j)) h_i \right\} \right) \\ &\quad - C \sum_{i=1}^{\infty} \left( \min \left\{ \prod_{j=1}^{i-1} (1 - \lambda h_j + m(h_j)) h_i, \prod_{j=1}^{i-1} (1 - \lambda h_j + n(h_j)) h_i \right\} \right) \\ &= C \sum_{i=1}^{\infty} \left( \prod_{j=1}^{i-1} (1 - \lambda h_j + s_i(h_j)) \right) h_i - C \sum_{i=1}^{\infty} \left( \prod_{j=1}^{i-1} (1 - \lambda h_j + r_i(h_j)) \right) h_i, \end{aligned}$$

where  $s_i(h_j)$  and  $r_i(h_j)$  are either  $m(h_j)$  or  $n(h_j)$ . By the above discussion we have

$$\sum_{i=1}^{\infty} \prod_{j=1}^{i-1} (1 - \lambda h_j + s_i(h_j)) h_i - \sum_{i=1}^{\infty} \prod_{j=1}^{i-1} (1 - \lambda h_j + r_i(h_j)) h_i \xrightarrow{\sup h_j \rightarrow 0} \frac{1}{\lambda} - \frac{1}{\lambda} = 0. \quad \square$$

**Corollary 2.** If  $\alpha_h$  and  $\beta_h$  are two families of discount factors with asymptotic discount rate  $\lambda$ , then we have:

$$\lim_{\substack{\sup_{i \in \mathbb{N}^*} h_i \rightarrow 0 \\ h_1 + h_2 + \dots = +\infty}} V_{T_{\infty}, \lambda}^{\alpha} = \lim_{\substack{\sup_{i \in \mathbb{N}^*} h_i \rightarrow 0 \\ h_1 + h_2 + \dots = +\infty}} V_{T_{\infty}, \lambda}^{\beta}; \quad (23)$$

$$\lim_{\substack{\sup_{i \in \mathbb{N}^*} h_i \rightarrow 0 \\ h_1 + h_2 + \dots = +\infty}} \lambda V_{T_\infty, \lambda}^\alpha = \lim_{\substack{\sup_{i \in \mathbb{N}^*} h_i \rightarrow 0 \\ h_1 + h_2 + \dots = +\infty}} v_{T_\infty, \lambda}^\alpha = \lim_{\substack{\sup_{i \in \mathbb{N}^*} h_i \rightarrow 0 \\ h_1 + h_2 + \dots = +\infty}} v_{T_\infty, \lambda}^\beta = \lim_{\substack{\sup_{i \in \mathbb{N}^*} h_i \rightarrow 0 \\ h_1 + h_2 + \dots = +\infty}} \lambda V_{T_\infty, \lambda}^\beta. \quad (24)$$

*Proof of Corollary 2.* (23) follows from Lemma 2, and (24) follows from (23) and Remark 10.  $\square$

A particular case of the above lemma was established in [Ney13] for games in which each stage has the same duration.

## 6 Final comments

**Remark 11** (The case of infinite stochastic games with stage duration). In this article, we assumed that action spaces are finite. However, it is not necessary, and we may assume that the action spaces are compact metric spaces, as long as the state space is still finite. Indeed, it is straightforward to define games with stage duration for this more general case. Under standard assumptions (see, for example, assumptions in [Sor02]) on the payoff and transition probability functions, there is a value for the game with stage duration, for any fixed partition  $T_\infty$ .

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

Theorem 1 also holds in this setting, because it is based on a result from [Sor18], which still holds if action sets are compact metric spaces and the state space is finite.

**Remark 12** (Stochastic games with public signals and its limit values). 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]. A distinctive feature of games with stage duration and public signals is the fact that there is no connection between the limit value (as the discount factor  $\lambda$  tends to 0) of a game with stage duration 1 and limit values of corresponding games with vanishing stage duration. In the case of perfect observation of the state, there is a connection, see [SV16, §7.3].

Let us finish with two propositions that show the connection between the model of games with stage duration presented here and the model of games with stage duration from [Sor18]. Recall that  $v_{T_\infty, k}^{\text{cont}}$  is the value of the discretization of a continuous-time Markov game, see §3.4.2.

**Proposition 2.** Fix the discretization of a continuous-time finite Markov game  $(\Omega, I, J, g, q_1)$ , where  $q_1$  is an infinitesimal generator, and fix a stochastic game  $(\Omega, I, J, g, q_2)$ , where  $q_2$  is the kernel. If  $q_1 = q_2 = q$ , then the uniform limits  $\lim_{\substack{\sup_{i \in \mathbb{N}^*} h_i \rightarrow 0 \\ h_1 + h_2 + \dots = T}} v_{T_\infty, k}^{\text{cont}}$  and  $\lim_{\substack{\sup_{i \in \mathbb{N}^*} h_i \rightarrow 0 \\ h_1 + h_2 + \dots = T}} v_{T_\infty, k}$  both exist and are equal.

*Proof.* By Theorem 1 and [Sor18, Proposition 4.3] any of these limit functions is the unique viscosity solution of  $0 = \frac{d}{dt}v(t, \omega) + \text{Val}_{I \times J}[k(t)g(i, j, \omega) + \langle q(i, j)(\omega, \cdot), v(t, \cdot) \rangle]$ .  $\square$

**Proposition 3.** Fix the discretization of a continuous-time state-blind finite Markov game  $(\Omega, I, J, g, q_1)$ , where  $q_1$  is its infinitesimal generator, and fix a state-blind stochastic game  $(\Omega, I, J, g, q_2)$ , where  $q_2$  is its kernel. If  $q_1 = q_2 = q$ , then the uniform limits  $\lim_{\substack{\sup_{i \in \mathbb{N}^*} h_i \rightarrow 0 \\ h_1 + h_2 + \dots = +\infty}} v_{T_\infty, k}^{\text{cont}}$  and  $\lim_{\substack{\sup_{i \in \mathbb{N}^*} h_i \rightarrow 0 \\ h_1 + h_2 + \dots = +\infty}} v_{T_\infty, k}$  both exist and

are equal.

*Proof.* By Theorem 1 and [Sor18, Proposition 5.3] any of these limit functions is the unique viscosity solution of  $0 = \frac{d}{dt}v(t, p) + \text{Val}_{I \times J}[k(t)g(i, j, p) + \langle p * q(i, j), \nabla v(t, p) \rangle]$ .  $\square$

## 7 Acknowledgements

The author is grateful to his research advisor Guillaume Vigeral for constant attention to this work. The author is grateful to Sylvain Sorin for useful discussions. The author is grateful to anonymous reviewers for many helpful comments.

## 8 References

- [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. doi:10.1239/jap/1053003547.
- [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. doi:10.3150/bj/1137421638.
- [LRS19] Rida Laraki, Jérôme Renault, and Sylvain Sorin. *Mathematical Foundations of Game Theory*. Universitext. Springer International Publishing, 2019. doi:10.1007/978-3-030-26646-2.
- [MSZ15] Jean-François Mertens, Sylvain Sorin, and Shmuel Zamir. *Repeated Games*. Econometric Society Monographs. Cambridge University Press, 2015. doi:10.1017/CB09781139343275.
- [Ney13] Abraham Neyman. Stochastic games with short-stage duration. *Dynamic Games and Applications*, 3(2):236–278, 2013. doi:10.1007/s13235-013-0083-x.
- [Nov24] Ivan Novikov. Limit value in zero-sum stochastic games with vanishing stage duration and public signals. Preprint, 2024. doi:10.48550/arXiv.2401.10572.
- [Sha53] Lloyd Stowell Shapley. Stochastic games. *Proceedings of the National Academy of Sciences*, 39(10):1095–1100, 1953. doi:10.1073/pnas.39.10.1095.
- [Sor02] Sylvain Sorin. *A First Course on Zero-Sum Repeated Games*, volume 37 of *Mathématiques et Applications*. Springer-Verlag Berlin Heidelberg, first edition, 2002. doi:10.1007/978-3-030-26646-2.
- [Sor18] Sylvain Sorin. Limit value of dynamic zero-sum games with vanishing stage duration. *Mathematics of Operations Research*, 43(1):51–63, 2018. doi:10.1287/moor.2017.0851.
- [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. doi:10.1007/s00182-015-0512-8.
- [Zac64] Lars Erik Zachrisson. *Advances in Game Theory. (AM-52), Volume 52*, pages 211–254. Princeton University Press, 1964. doi:10.1515/9781400882014-014.
- [Zil16] Bruno Ziliotto. Zero-sum repeated games: Counterexamples to the existence of the asymptotic value and the conjecture  $\max\min = \lim v_n$ . *The Annals of Probability*, 44(2):1107 – 1133, 2016. doi:10.1214/14-AOP997.