Correlated Equilibria for Infinite Horizon Nonzero-Sum Stochastic Differential Games

Beatris A. Escobedo-Trujillo and Héctor Jasso-Fuentes

Abstract This chapter is about two-person nonzero-sum stochastic differential games with discounted and long-run average (a.k.a. ergodic) payoffs. Our aim is to give conditions for the existence of *feedback* correlated randomized equilibria for each aforementioned payoff that are natural generalizations of the well-known Nash equilibria. To do so, we rewrite our original problem in terms of an auxiliary zero-sum game, so that the way to find correlated equilibria is based on some properties of this later game. Key ingredients to achieve the desired results are the continuity properties of the payoffs.

1 Introduction

Nash equilibrium is a very useful concept in game theory, however it is well known that under standard conditions the existence of Nash equilibria in nonzero-sum games with uncountable state-action spaces is not necessarily guarantied within the set of randomized strategies.

During the past decades, there has been works that have dealt to game models with particular features in order to ensure the existence of Nash equilibria; for instance, games with an additive structure (see, e.g. [8, 16, 17]). Other works have explored an alternative method, which consists of "relaxing" the idea of Nash equilibrium. The idea is to extend the set of strategies into a bigger one, giving rise to the concept of *correlated strategies* as well as to the concept of *correlated equilib-*

Beatris A. Escobedo-Trujillo
Engineering Faculty. Universidad Veracruzana, Coatzacoalcos, Ver., México. Tel. (921) 218-77-83, 211-57-07. e-mail: bescobedo@uv.mx

Héctor Jasso-Fuentes
Departamento de Matemáticas, CINVESTAV-IPN. Apartado Postal 14-740, México. D.F., 07000, México. e-mail: hjasso@math.cinvestav.mx

G. Yin and Q. Zhang (eds.), *Modeling, Stochastic Control, Optimization, and Applications*, The IMA Volumes in Mathematics and its Applications 164,
https://doi.org/10.1007/978-3-030-25498-8_9

ria (see e.g. [4, 5, 20, 21]). This approach is the one we have focused on in this manuscript, whose details will be explained in later sections.

Recall that the Nash equilibrium concept means that if one player tries to alter his strategy unilaterally, he cannot improve his performance by such a change. If players choose their strategies according to the Nash equilibrium concept, they are said to play non-cooperatively, i.e., each player is only interested in maximizing his own utility. The correlated equilibrium concept means that all players, before taking a decision over the strategies, receive a global (or joint) recommendation, that is drawn randomly according to a joint distribution μ, then no player has an incentive to divert from the recommendation, provided that all other players follow theirs.

The main distinguishing feature of the concept of correlated equilibrium, unlike the definition of Nash equilibrium is that those recommendations do not need to be independent; i.e., the joint distributions do not need to be a product of marginal distributions. It turns out that a correlated equilibrium μ is a Nash equilibrium if and only if μ is a product measure.

It is well recognized that correlated equilibria were introduced by Aumman in 1974 for nonzero-sum games in normal form, extending the Nash equilibrium concept, [4, 5]. There exists a vast number of manuscripts that are focused on correlated strategies concept providing conditions for the existence of correlated equilibria [4, 5, 10, 20, 21, 22, 23, 24], this is, in some part, because it is easier to prove the existence and characterize correlated equilibria compared with Nash equilibria.

The work is inspired by the paper [20] which deals with correlated relaxed equilibria in nonzero-sum linear differential games with finite-horizon payoffs. Our aim here is to prove the existence of *feedback* correlated equilibria for a more general dynamic when the payoffs are of the (infinite horizon) discounted and average type. A key point to obtain our desired equilibria is to guarantee the continuity to both payoff functions (discounted and average payoffs) within the set of correlated strategies.

Although we restrict ourselves to the case of two players, its relatively easy to extend our results to the more general context of N players.

The main novelty of the manuscript lies in the fact that we are working with infinite-horizon (discounted and ergodic) payoff criteria under a considerable general diffusion process. Furthermore, the set of correlated equilibria are shown to be *feedback*, meaning that they are dependent on the current state of the game. To the best of our knowledge, this treatment has not been already studied in the current literature.

This chapter is organized as follows: In section 2, we introduce both the game and the payoffs we are trying to optimize. We also define the Nash equilibrium concept. Section 3 is devoted to the introduction of correlated strategies. We extend the domain of our payoffs over these strategies and will prove the continuity of those criteria. By last, in section 4 we introduce the concept of correlated equilibria and prove the existence of them. To do so, we rewrite the original game as a zero-sum game and we explore some of its properties. Correlated equilibria will be obtained through the use of some min-max theorems as well as for the continuity of our payoffs.

Notation and terminology.

- For some m, $n \in \mathbb{N}$, let $\mathscr{O} \subset \mathbb{R}^m$ and $V \subset \mathbb{R}^n$ be given open and Borel sets, respectively. We define:
  - $\mathbb{W}^{l,p}(\mathscr{O})$ the Sobolev space consisting of all real-valued measurable functions h on $\mathscr{O}$ such that $D^\lambda h$ exists for all $|\lambda| \leq l$ in the weak sense and it belongs to $\mathbb{L}^p(\mathscr{O})$, where

$$D^\lambda h := \frac{\partial^{|\lambda|} h}{\partial x_1^{\lambda_1}, \ldots, \partial x_m^{\lambda_m}} \quad \text{with } \lambda = (\lambda_1, \cdots, \lambda_m), \quad \text{and} \quad |\lambda| := \sum_{i=1}^{m} \lambda_i.$$

  - $\mathbb{C}^k(\mathscr{O})$ the space of all real-valued continuous functions on $\mathscr{O}$ with continuous l-th partial derivative in $x_i \in \mathbb{R}$, for $i = 1, ..., m$, $l = 0, 1, ..., k$. In particular, when $k = 0$, $\mathbb{C}^0(\mathscr{O})$ stands for the space of real–valued continuous functions on $\mathscr{O}$.
  - $\mathbb{C}^{k,\beta}(\mathscr{O})$ the subspace of $\mathbb{C}^k(\mathscr{O})$ consisting of all those functions h such that $D^\lambda h$ satisfies a Hölder condition with exponent $\beta \in (0,1]$, for all $|\lambda| \leq k$.
  - $\mathbb{C}_b(\mathscr{O} \times V)$ the space consisting of all continuous bounded functions on $\mathscr{O} \times V$.
- For vectors x and matrices A we use the usual Euclidean norms

$$|x|^2 := \sum_i x_i^2 \quad \text{and} \quad |A|^2 := Tr(AA^{'}) = \sum_{i,j} A_{ij}^2,$$

  where $A^{'}$ and $Tr(\cdot)$ denote the transpose and the trace of a square matrix, respectively.
- For any two strategies, say π^1 and π^2, the notation $\pi^1 \times \pi^2$ means the product measure associated to this pair.

2 The game model

Consider an $m-$dimensional diffusion process $x(\cdot)$ controlled by two players and evolving according to the stochastic differential equation

$$dx(t) = b(x(t), u_1(t), u_2(t))dt + \sigma(x(t))dW(t), \qquad x(0) = x_0, \tag{1}$$

where $b : \mathbb{R}^m \times U_1 \times U_2 \to \mathbb{R}^m$, $\sigma : \mathbb{R}^m \to \mathbb{R}^{m \times d}$ are given functions, and $W(\cdot)$ is a d-dimensional standard Brownian motion. The sets $U_1 \subset \mathbb{R}^{m_1}$ and $U_2 \subset \mathbb{R}^{m_2}$ are Borel sets called the action set for player 1 and player 2, respectively. Moreover, for $k = 1, 2$, $u_k(\cdot)$ is a non-anticipative U_k-valued stochastic process representing the control actions of player k at each time $t \geq 0$.

For $(u_1, u_2) \in U_1 \times U_2$, and h in $\mathbb{W}^{2,p}(\mathbb{R}^m)$, let

$$L^{u_1,u_2}h(x) := \sum_{i=1}^{m} b_i(x,u_1,u_2)\frac{\partial h}{\partial x_i}(x) + \frac{1}{2}\sum_{i,j}^{m} a^{ij}(x)\frac{\partial^2 h}{\partial x_i \partial x_j}(x), \tag{2}$$

where b_i is the i–th component of b, and a^{ij} is the (i,j)–component of the matrix $a(\cdot) := \sigma(\cdot)\sigma'(\cdot)$.

Let us now proceed to define the sets of strategies allowed for each player.

Randomized Markov strategies. Let $\mathscr{B}(U_1)$ be the Borel $\sigma-$algebra of U_1 , and let $\mathscr{P}(U_1)$ be the space of probability measures on U_1. In the same way, we define $\mathscr{B}(U_2)$ and $\mathscr{P}(U_2)$ associated to player 2.

Definition 1. A *randomized Markov strategy* for player k ($k = 1,2$) is defined as a family $\pi^k := \{\pi_t^k : t > 0\}$ of stochastic kernels on $\mathscr{B}(U_k)\times\mathbb{R}^m$; that is:

(a) for each $t \geq 0$ and $x \in \mathbb{R}^m$, $\pi_t^k(\cdot|x)$ is in $\mathscr{P}(U_k)$, satisfying $\pi_t^k(U_k|x) = 1$;
(b) for each $D \in \mathscr{B}(U_k)$ and $t \geq 0$, $\pi_t^k(D|\cdot)$ is a Borel function on $\mathbb{R}^m$; and
(c) for each $B \in \mathscr{B}(U_k)$ and $x \in \mathbb{R}^m$, the function $t \mapsto \pi_t^k(B|x)$ is a Borel measurable function.

Definition 2. A randomized strategy $\pi^k = \{\pi_t^k : t \geq 0\}$ ($k = 1,2$) is said to be *stationary* if there is a stochastic kernel π^k on $\mathscr{B}(U_k)\times\mathbb{R}^m$ such that $\pi_t^k(\cdot|x) = \pi^k(\cdot|x)$ for all $x \in \mathbb{R}^m$, $t \geq 0$.

The set of randomized stationary strategies for player k is denoted by Π_k, $k = 1,2$.

Next we define the payoff functions that each player wants to "optimize.''

Payoff rates. For each player $k = 1,2$, let $r_k : \mathbb{R}^m \times U_1 \times U_2 \to \mathbb{R}$ be a measurable function, which we will call the payoff rate of player k; in this sense, at each $t \geq 0$, $r_k(x(t),u_1,u_2)$ is the payoff of player k at time t, when the actions $u_1 \in U_1$ and $u_2 \in U_2$ are decided by players 1 and 2, respectively.

Throughout this manuscript we will use the notation $\pi^1 \times \pi^2$, representing the product measure of π^1 and π^2.

Let the function ψ be either b or r_k, $k = 1,2$. When players use a stationary randomized strategy $(\pi^1 \times \pi^2) \in \Pi_1 \times \Pi_2$, we write:

$$\psi(x,\pi^1\times\pi^2) := \int_{U_1}\int_{U_2} \psi(x,u_1,u_2)\pi^1(du_1|x)\pi^2(du_2|x), \quad x \in \mathbb{R}^m.$$

With the above notation, the infinitesimal generator (2) is written as

$$L^{\pi^1\times\pi^2}h(x) := \sum_{i=1}^{m} b_i(x,\pi^1\times\pi^2)\frac{\partial h}{\partial x_i}(x) + \frac{1}{2}\sum_{i,j}^{m} a^{ij}(x)\frac{\partial^2 h}{\partial x_i \partial x_j}(x), \quad x \in \mathbb{R}^m.$$

Assume for the moment the existence of the probability measure $\mathbb{P}_x^{\pi^1\times\pi^2}$ for each $x \in \mathbb{R}^m$ and $\pi^1 \times \pi^2 \in \Pi_1 \times \Pi_2$, associated to the process $x(\cdot)$. We will also denote by $\mathbb{E}_x^{\pi^1\times\pi^2}(\cdot)$ its respective expectation. Next define the payoff criteria each player would be interested to optimize.

Definition 3 (Discounted payoff criterion). Let $\alpha > 0$, and consider the payoff rates r_1 and r_2. For each player $k = 1,2$, the expected $\alpha-$discounted payoff for player k when players use the strategy $(\pi^1 \times \pi^2) \in \Pi_1 \times \Pi_2$ given the initial state $x \in \mathbb{R}^m$, is

$$V_k(x,\pi^1 \times \pi^2) := \mathbb{E}_x^{\pi^1 \times \pi^2}\left[\int_0^\infty e^{-\alpha t} r_k(x(t),\pi^1 \times \pi^2)dt\right]. \tag{3}$$

Definition 4 (Average payoff criterion). For each player $k = 1,2$, the expected average payoff for player k when players use the strategy $(\pi^1 \times \pi^2) \in \Pi_1 \times \Pi_2$ given the initial state $x \in \mathbb{R}^m$, is

$$J_k(x,\pi^1 \times \pi^2) := \liminf_{T\to\infty} \frac{1}{T}\mathbb{E}_x^{\pi^1 \times \pi^2}\left[\int_0^T r_k(x(t),\pi^1 \times \pi^2)dt\right]. \tag{4}$$

In a noncooperative N-person nonzero-sum stochastic differential game, each player tries to maximize (or minimize) his/her individual performance criterion (in particular, criteria of type (3) and (4)). A Nash equilibrium, in this case, is a strategy such that once it is chosen by the players, no player will profit unilaterally by simply changing his/her own strategy. More specifically, in the maximization context, we have the next definition:

Definition 5 (Nash equilibrium). Let F_k, $k = 1,2$, be either the discounted payoff in (3) or the average payoff (4). A randomized pair of strategies $(\pi_*^1 \times \pi_*^2) \in \Pi_1 \times \Pi_2$ is a Nash equilibrium if and only if

$$F_1(x,\pi_*^1 \times \pi_*^2) \geq F_1(x,\pi^1 \times \pi_*^2), \quad \forall\, \pi^1 \in \Pi_1,$$

$$F_2(x,\pi_*^1 \times \pi_*^2) \geq F_2(x,\pi_*^1 \times \pi^2), \quad \forall\, \pi^2 \in \Pi_2.$$

It is also well recognized that the existence of Nash equilibria in nonzero-sum games with uncountable state-action spaces is not necessarily guarantied in the set of randomized policies $\Pi_1 \times \Pi_2$ under standard conditions (for example, the conditions established in this chapter); however, such existence is achieved under special cases. For instance, we can assume the drift b in the dynamic (1) and the payoff rates r_k both satisfy an additive structure property (see, for instance, [8, 16, 17]). One alternative to avoid restrictive assumptions to the model, is to extend the concept of Nash equilibrium into a bigger set of $\Pi_1 \times \Pi_2$, which entails, in particular, to the concept of *correlated strategies* and its corresponding *correlated equilibrium*.

In the next section we will provide, among other things, conditions so that the dynamic (1) can attain a unique solution in some sense and that the payoffs in (3) and (4) are finite valued on the set of correlated strategies.

3 Correlated strategies.

In this work we extend the set of strategies available to the players. These strategies allow players to correlate their decisions during a pre-play communication process (see Section 4).

Definition 6 (Correlated strategy). A correlated (stationary) randomized strategy μ is a stochastic kernel on $\mathscr{B}(U_1 \times U_2) \times \mathbb{R}^m$ such that:

(a) for each $x \in \mathbb{R}^m$, $\mu(\cdot|x)$ is a joined probability measure on $U_1 \times U_2$ and such that $\mu(U_1 \times U_2|x) = 1$.
(b) For each $D \in \mathscr{B}(U_1 \times U_2)$, $\mu(D|\cdot)$ is Borel measurable on $\mathbb{R}^m$.

We will denote by Γ the set of all correlated randomized strategies. On the other hand, the marginal distributions of μ are defined as:

$$\mu_1(B_1|x) := \mu(B_1 \times U_2|x) \qquad \text{and} \qquad \mu_2(B_2|x) := \mu(U_1 \times B_2|x),$$

for each Borel set $B_k \in \mathscr{B}(U_k)$, $k = 1,2$, and $x \in \mathbb{R}^m$.

Let ψ be either b, r_1 or r_2. When players use a correlated strategy $\mu \in \Gamma$, we write

$$\psi(x,\mu) := \int_{U_1 \times U_2} \psi(x,u_1,u_2)\mu(d(u_1,u_2)|x),$$

Furthermore, the generator (2) turns out to be

$$L^{\mu}h(x) := \sum_{i=1}^{m} b_i(x,\mu)\frac{\partial h}{\partial x_i}(x) + \frac{1}{2}\sum_{i,j}^{m} a^{ij}(x)\frac{\partial^2 h}{\partial x_i \partial x_j}(x), \quad x \in \mathbb{R}^m. \tag{5}$$

We denote by Γ_k the set of k-marginal measures associated to Γ, for $k = 1,2$.

Remark 1. Throughout this work we will assume that the players choose only randomized *stationary* strategies. The reason is that, even when it is possible to work in a more general class of strategies (for instance that of the so-named non-anticipative policies), our present hypotheses (stated later on) ensure the existence of optimal policies in the class of stationary strategies for all players. Further, it is worth to mention that recurrence and ergodicity properties of the state system (1) can be easily verified through the use of stationary strategies, but for general non-anticipative strategies, the corresponding state system might be time-inhomogeneous; a fact that can be hard to handle.

Assumption 1 *Recall the elements of the dynamic* (1). *We assume:*

(a) The action sets U_1 and U_2 are compact.
(b) The function $b : \mathbb{R}^m \times U_1 \times U_2 \to \mathbb{R}^m$ satisfies the following conditions:

(i) it is continuous on $\mathbb{R}^m \times U_1 \times U_2$.

(ii) it satisfies a Lipschitz condition uniformly in $(u_1,u_2)\in U_1\times U_2$; that is, there exists a positive constant K_1 such that, for all $x,y\in\mathbb{R}^m$,

$$\sup_{(u_1,u_2)\in U_1\times U_2}|b(x,u_1,u_2)-b(y,u_1,u_2)|\leq K_1|x-y|.$$

(c) There exists a positive constant K_2 such that for all $x,y\in\mathbb{R}^m$,

$$|\sigma(x)-\sigma(y)|\leq K_2|x-y|.$$

(d) (Uniform ellipticity). The matrix $a(x)=\sigma(x)\sigma'(x)$ satisfies that, for some constant $c_0>0$

$$xa(y)x'\geq c_0|x|^2 \text{ for all } x,y\in\mathbb{R}^m.$$

Remark 2. (a) Assumption 1 ensures that there exists an almost surely unique strong solution of (1), for each strategy $\mu\in\Gamma$, which is a Markov–Feller process and whose infinitesimal generator coincides with L^{μ} in (5). (For more details, see the arguments of [2, Theorem 2.2.7]).

(b) The aforementioned existence and uniqueness remain valid for special types of joint kernels of either form $\mu=\pi_1\times\pi_2$, or $\mu=\pi_1\times\mu_2$ or $\mu=\mu_1\times\pi_2$, for every $\pi_1\in\Pi_1$, $\pi_2\in\Pi_2$ $\mu_1\in\Gamma_1$, $\mu_2\in\Gamma_2$. This implies that the dynamic (1) is well defined even when a usual pair of strategies $(\pi_1\times\pi_2)\in\Pi_1\times\Pi_2$ as that introduced in Definition 2 is applied.

The following assumption is a Lyapunov–like condition that guaranties, in particular, that the discounted payoff criterion (3) is finite, among other facts such as the positive recurrence property of the diffusion (1) and the existence of an invariant measure, each of them for a suitable type of controls (or strategies) $u_1(\cdot)$ and $u_2(\cdot)$.

Assumption 2 *There exists a function $w\in\mathbb{C}^2(\mathbb{R}^m)$, with $w\geq 1$, and constants $d\geq c>0$ such that*

(i) $\lim_{|x|\to\infty}w(x)=+\infty$, *and*
(ii) $L^{u_1,u_2}w(x)\leq -cw(x)+d$ *for each $(u_1,u_2)\in U_1\times U_2$ and $x\in\mathbb{R}^m$.*

Remark 3. An easy application of Ito's formula to $e^{ct}w(x(t))$ along with Assumption $2(ii)$, give us that

$$\sup_{\mu\in\Gamma}\mathbb{E}_x^{\mu}(w(x(t))\leq e^{-ct}w(x)+\frac{d}{c}(1-e^{-ct}).$$

Definition 7. Let $w\geq 1$ be the function in Assumption 2 and $\mathscr{O}\subset\mathbb{R}^m$ be an open set. We define the Banach space $\mathbb{B}_w(\mathscr{O})$ consisting of real–valued measurable functions h on $\mathscr{O}$ with finite w–norm defined as follows:

$$||h||_w:=\sup_{x\in\mathscr{O}}\frac{|h(x)|}{w(x)}.$$

We also include another set of hypotheses related to the payoffs r_k that uses the above definition.

Assumption 3 *(a) The function $r_k(x,u_1,u_2)$ is continuous on $\mathbb{R}^m \times U_1 \times U_2$ and locally Lipschitz in x uniformly with respect to $(u_1,u_2) \in U_1 \times U_2$; that is, for each $R > 0$, there exists a constant $K(R) > 0$ such that*

$$\sup_{(u_1,u_2)\in U_1\times U_2} |r_k(x,u_1,u_2) - r_k(y,u_1,u_2)| \leq K(R)|x-y| \text{ for all } |x|,|y| \leq R.$$

(b) $r_k(\cdot,u_1,u_2)$ is in $\mathbb{B}_w(\mathbb{R}^m)$ uniformly in (u_1,u_2); that is, there exists $M > 0$ such that for all $x \in \mathbb{R}^m$

$$\sup_{(u_1,u_2)\in U_1\times U_2} |r_k(x,u_1,u_2)| \leq Mw(x).$$

We will extend the discounted payoff criteria (3) and (4) on the set Γ.

Extended discounted criterion: Given r_k as in Assumption 3, $k = 1,2$, and for any initial state $x \in \mathbb{R}^m$, the extended $\alpha-$discounted payoff for player k when the strategy $\mu \in \Gamma$ is applied is defined as

$$V_k(x,\mu) := \mathbb{E}_x^{\mu}\left[\int_0^{\infty} e^{-\alpha t} r_k(x(t),\mu)dt\right]. \tag{6}$$

The following proposition is a direct consequence of Assumptions 2, 3(b), and Remark 3, so we shall omit the proof; similar arguments can be founded in [9, Proposition 9.1].

Proposition 1. *Under the Assumptions 1, 2, 3, the payoff (6) belongs to the space $\mathbb{B}_w(\mathbb{R}^m)$ for each correlated strategy μ; in fact, for each x in $\mathbb{R}^m$ we have*

$$|V_k(x,\mu)| \leq M(\alpha)w(x) \tag{7}$$

with $M(\alpha) := M\frac{(\alpha+d)}{\alpha c}$. Here, c and d are the constants in Assumption 2(b), and M is the constant in Assumption 3(b).

Extended average criterion: Let r_k be as in Assumption 3, $k = 1,2$, and $x \in \mathbb{R}^m$. We define the extended average payoff for player k when the strategy $\mu \in \Gamma$ is used and initial state x, as follows

$$J_k(x,\mu) := \liminf_{T\to\infty} \frac{1}{T}\mathbb{E}_x^{\mu}\left[\int_0^T r_k(x(t),\mu)dt\right]. \tag{8}$$

Note that the above limit always exists. Actually, we will impose an ergodicity condition so that this limit becomes a constant in some sense.

From the arguments in [2, 12], for each $\mu \in \Gamma$, the Markov process $x(\cdot)$ is positive recurrent and admits a unique invariant probability measure η_μ, for which

$$\eta_\mu(w) := \int_{\mathbb{R}^m} w(x)\eta_\mu(dx) < \infty, \tag{9}$$

where w is the function defined in Assumption 2. The next assumption corresponds to the well-known assymptotic behaviour of $x(t)$ when t goes to the infinite. Sufficient conditions for this assumption can be seen in Theorem 2.7 in [14].

Assumption 4 *For every $\mu \in \Gamma$, the process $x(\cdot)$ is uniformly w-exponentially ergodic; that is, there exist positive constants k_1 and k_2 such that*

$$\sup_{\mu \in \Gamma} \left| \mathbb{E}_x^{\mu} \left[\nu(x(t)) \right] - \eta_{\mu}(\nu) \right| \leq k_1 \|\nu\|_w e^{-k_2 t} w(x), \tag{10}$$

for all $x \in \mathbb{R}^m$, $t \geq 0$, and $\nu \in \mathbb{B}_w(\mathbb{R}^m)$.

In (10), the notation $\eta_{\mu}(\nu)$ has the same meaning as (9) with ν instead of w.

Remark 4. Under Assumptions 1, 2, 3, and 4, the extended average payoff criterion (8) satisfies the following: For each $k = 1,2$:

(a) $J_k(x,\mu) = \int_{\mathbb{R}^m} r_k(y,\mu)\eta_{\mu}(dy)$, for all $x \in \mathbb{R}^m$, $\mu \in \Gamma$; actually the limit in (8) does exist in a strong sense (i.e., $\liminf = \limsup$) and does not depend on the initial condition x.

(b) $\sup_{\mu \in \Gamma} |J_k(x,\mu)| \leq M \cdot d/c$, , for all $x \in \mathbb{R}^m$, with M and d,c the constants appearing in our previous Assumptions 3 and 2, respectively.

For a proof of these two assertions, we can quote Section 3 in [9] or Section 2 in [14].

3.1 Continuity properties

In this part we will ensure that the functions $\mu \mapsto V_k(x,\mu)$ and $\mu \mapsto J_k(x,\mu)$ are continuous. To this end, we endow the set Γ with the topology of joint strategies (see e.g. [3, Lemma 3.4] or [6]).

Definition 8 (Topology of join strategies). We say that a sequence $\{\mu_n : n = 1,2,...\} \subset \Gamma$ converges to $\mu \in \Gamma$ (and we will denote such convergence as $\mu_n \stackrel{W}{\to} \mu$) if and only if for all $h \in \mathbb{C}_b(\mathbb{R}^m \times U_1 \times U_2)$ and $g \in \mathbb{L}^1(\mathbb{R}^m)$

$$\int_{\mathbb{R}^m} g(x) \int_{U_1 \times U_2} h(x,u_1,u_2)\mu_n(d(u_1,u_2)|x)dx \underset{n\to\infty}{\longrightarrow} \int_{\mathbb{R}^m} g(x) \int_{U_1 \times U_2} h(x,u_1,u_2)\mu(d(u_1,u_2)|x)dx.$$

Remark 5. The space Γ is a convex compact metric space endowed with the previous topology; see [25, Theorem IV.3.11] or [6, Section 3]. Furthermore, as was mentioned in [16, Remark 2.11(b)], the set $\Pi_1 \times \Pi_2$ is compact too. The convexity of this last product set easily follows from the convexity of Π_k, $k = 1,2$.

The following proposition gives a characterization for the α-discounted reward (6). For a proof we quote [15, Proposition 3.1.5].

Proposition 2. *Assume that the Assumptions 1, 2, 3 hold true. Then, for every $\mu \in \Gamma$, the associated total expected α-discounted function $V_k(\cdot,\mu)$ ($k = 1,2$) is in $\mathbb{W}^{2,p}(\mathbb{R}^m) \cap \mathbb{B}_w(\mathbb{R}^m)$ and it satisfies the equation*

$$\alpha V_k(x,\mu) = r_k(x,\mu) + L^{\mu} V_k(x,\mu). \tag{11}$$

Conversely, if some function $\varphi_k \in \mathbb{W}^{2,p}(\mathbb{R}^n) \cap \mathbb{B}_w(\mathbb{R}^m)$ verifies (11), then

$$\varphi_k(x) = V_k(x,\mu) \quad \text{for all } x \in \mathbb{R}^m.$$

Moreover, if the equality in (11) is replaced by "$\leq$" or "$\geq$", then (11) *holds with the respective inequality.*

The following result addresses a continuity property of the total expected α-discounted payoffs.

Proposition 3 (Continuity of V_k). *For $k = 1,2$, the mapping $\mu \longmapsto V_k(x,\mu)$ is continuous on Γ, for each $x \in \mathbb{R}^m$.*

Proof. Let $\{\mu_n\} \in \Gamma$ such that $\mu_n \xrightarrow{W} \mu$. Observe that Proposition 2 ensures that, for each $n \geq 1$, $V_k(x,\mu_n)$ satisfies the equation

$$\alpha V_k(x,\mu_n) = r_k(x,\mu_n) + L^{\mu_n} V_k(x,\mu_n) \quad x \in \mathbb{R}^m. \tag{12}$$

This last equation in terms of the operator $\mathscr{L}_{\alpha}^{\mu_n}$ given in (30) becomes

$$0 = \mathscr{L}_{\alpha}^{\mu_n} V_k(x,\mu_n) \quad x \in \mathbb{R}^m. \tag{13}$$

Next we will check that the hypotheses (a)-(e) of Theorem 2 provided in the appendix of this chapter are satisfied.

(a) This hypothesis trivially follows from (13) (or by (12)).

(b) To prove this hypotheses, let $R > 0$, and take the ball $B_R := \{x \in \mathbb{R}^m \mid |x| < R\}$. By [13, Theorem 9.11], there exists a constant C_0 independent of R such that, for a fixed $p > m$ (m being the dimension of (1)), we have

$$\begin{aligned}\|V_k(\cdot,\mu_n)\|_{\mathbb{W}^{2,p}(B_R)} &\leq C_0 \left(\|V_k(\cdot,\mu_n)\|_{\mathbb{L}^p(B_{2R})} + \|r_k(\cdot,\mu_n)\|_{\mathbb{L}^p(B_{2R})}\right)\\ &\leq C_0 \left(M(\alpha)\|w\|_{\mathbb{L}^p(B_{2R})} + M\|w\|_{\mathbb{L}^p(B_{2R})}\right)\\ &\leq C_0 \left(M(\alpha) + M\right) |\bar{B}_{2R}|^{1/p} \max_{x \in \bar{B}_{2R}} w(x) < \infty,\end{aligned}$$

where $|\bar{B}_{2R}|$ represents the volume of the closed ball with radious $2R$, and M and $M(\alpha)$ are the constants in Assumption 3(b) and in (7), respectively.

(c)-(e) The parts (c) and (d) of Theorem 2 trivially hold by taking $\xi_n \equiv 0$ and $\alpha_n \equiv \alpha$, whereas that part (e) is part of our hypotheses.

Then, for $k = 1,2$, we get the existence of a function $h_{\mu}^k \in \mathbb{W}^{2,p}(B_R)$ together with a subsequence $\{n_j\}$ such that $V_k(\cdot,\mu_{n_j}) \to h_{\mu}^k(\cdot)$ *uniformly* in B_R and pointwise on $\mathbb{R}^m$ as $j \to \infty$ and $\mu_{n_j} \xrightarrow{W} \mu$. Furthermore, h_{μ}^k satisfies

$$\alpha h_\mu^k(x) = r_k(x,\mu) + L^\mu h_\mu^k(x), \quad x \in B_R.$$

Since the radious $R > 0$ was arbitrary, we can extend our analysis to all of $x \in \mathbb{R}^m$. Thus, Proposition 2 asserts that $h_\mu^k(x)$ actually coincides with $V_k(x,\mu)$. This proves the continuity of V_k. □

We are going to focus on the continuity of the extended average payoff (8). To begin with, we shall use a characterization of this criterion, whose proof is identical to that in [14, Lemma 4.1] (see also [9, Proposition 5.1]).

Proposition 4 (Poisson equation). *For each $k = 1,2$, and each fixed strategy $\mu \in \Gamma$, we denote by $g^k(\mu) := \int_{\mathbb{R}^m} r_k(y,\mu)\eta_\mu(dy)$. Then, under the Assumptions 1, 2, 3, and 4, there exists a function $\varphi_\mu^k \in \mathbb{W}^{2,p}(\mathbb{R}^n) \cap \mathbb{B}_w(\mathbb{R}^m)$, such that the pair $(g^k(\mu), \varphi_\mu^k)$ satisfies the so-named Poisson equation*

$$g^k(\mu) = r_k(x,\mu) + L^\mu \varphi_\mu^k(x), \quad k = 1,2, \quad x \in \mathbb{R}^m, \tag{14}$$

as long with the transversality condition

$$\int_{\mathbb{R}^m} \varphi_\mu^k(x)\eta_\mu(dx) = 0. \tag{15}$$

Moreover, $g^k(\mu)$ is equal to the extended average payoff $J_k(x,\mu)$, for all $x \in \mathbb{R}^m$.

Now let us show the continuity of $g^k(\mu)$:

Proposition 5 (Continuity of g^k). *For $k = 1,2$, the mapping $\mu \longmapsto g^k(\mu)$ is continuous on Γ.*

Proof. The proof is similar to that given in Proposition 3. Indeed, take again a ball B_R for some $R > 0$ and use $\mu_n \in \Gamma$ such that $\mu_n \xrightarrow{W} \mu$. By Proposition 4, for each n the pair $(g^k(\mu_n), \varphi_{\mu_n}^k)$ satisfies the equation (14) with $\varphi_{\mu_n}^k \in \mathbb{W}^{2,p}(\mathbb{R}^m) \cap \mathbb{B}_w(\mathbb{R}^m)$. This equation in terms of operator $\mathscr{L}_\alpha^{\mu_n}$ in (30) becomes

$$g^k(\mu_n) = \mathscr{L}_0^{\mu_n} \varphi_{\mu_n}^k(x). \tag{16}$$

We will check that hypotheses (a)-(e) of Theorem 2 are satisfied. For this end, note by Assumption 3(b) and Proposition 4, that the functions $r_k(\cdot,\mu_n)$ and $\varphi_{\mu_n}^k$ are both in $\mathbb{B}_w(\mathbb{R}^m)$. Thus, using again the result in [13, Theorem 9.11], we can ensure the existence of some $\bar{C}_0$ (independent of R) such that

$$\begin{aligned}
||\varphi_{\mu_n}^k||_{\mathbb{W}^{2,p}(B_R)} &\leq \bar{C}_0(||\varphi_{\mu_n}^k||_{\mathbb{L}^p(B_{2R})} + ||r_k(\cdot,\mu_n)||_{\mathbb{L}^p(B_{2R})} \\
&\leq \bar{C}_0(M_1||w||_{\mathbb{L}^p(B_{2R})} + M||w||_{\mathbb{L}^p(B_{2R})}) \\
&\leq \bar{C}_0(M_1 + M)|\overline{B}_{2R}|^{1/p} \max_{x\in\overline{B}_{2R}} w(x) < \infty,
\end{aligned} \tag{17}$$

where $|\overline{B}_{2R}|$ is defined as in the proof of Proposition 3 and M_1 is some given constant. The hypotheses (a) and (b) follows from (16) and (17), respectively. As for part (c), we take $\xi_n = g^k(\mu_n)$ and noting that $|g^k(\mu_n)| \leq Md/c$ (see Remark 4(b)), we get the existence of a constant g^k such that $g^k(\mu_n) \to g^k$ (under a suitable subsequence), hence part (c) of Theorem 2 trivially holds. Also, part (d) is satisfied by taking $\alpha_n \equiv 0$. Part (e) is part of our hypotheses. In this way, Theorem 2 ensures the existence of a function $\varphi_\mu^k \in \mathbb{W}^{2,p}(B_R)$ together with a subsequence $\{n_j\}$ such that $\varphi_{\mu_{n_j}}^k(\cdot) \to \varphi_\mu^k(\cdot)$ *uniformly* in B_R and pointwise on $\mathbb{R}^m$ as $j \to \infty$ and $\mu_{n_j} \xrightarrow{W} \mu$. Moreover, φ_μ^k satisfies

$$g^k = r_k(x,\mu) + L^\mu \varphi_\mu^k(x) = 0, \quad x \in B_R. \tag{18}$$

Since the radious $R > 0$ was arbitrary, we can extend our analysis to all of $x \in \mathbb{R}^m$.

Finally, let $\overline{\varphi}_\mu^k(\cdot)$ be the bias function of μ, see [9, Definition 5.1]. By [9, Proposition 5.1], the pair $(g^k(\mu), \overline{\varphi}_\mu^k(\cdot))$ is the unique solution of the Poisson equation (14), i.e., $\varphi_\mu^k(\cdot) = \overline{\varphi}_\mu^k(\cdot) + c$ for some constant $c \in \mathbb{R}$, and $g^k = g^k(\mu)$. This implies that

$$g^k(\mu) = r_k(\cdot,\mu) + \mathscr{L}^\mu \overline{\varphi}_\mu^k(x) = 0 \quad x \in \mathbb{R}^m.$$

Furthermore, [9, Proposition 5.1] also ensures that the bias $\overline{\varphi}_\mu^k(\cdot)$ satisfies the transversality condition (15). Hence, a simple use of Proposition 4 provides us the continuity of the mapping $\mu \longmapsto g^k(\mu)$ on Γ. □

4 Correlated equilibria

As mentioned in [20], a correlated strategy limits the freedom of the players in selecting their strategies, because a process of pre-play communication is needed to carry out a correlated strategy. However, any player is free to choose any strategy, regardless of the results of the communication process.

Suppose that a correlated strategy $\mu \in \Gamma$ is fixed by the players during a pre-play communication process. Then players make their final decisions independently of each other. As a consequence, we obtain the following cases.

1. Both players accept $\mu \in \Gamma$, then the system (1) evolves by applying the control strategy μ.
2. Both players do not accept $\mu \in \Gamma$, then the system (1) evolves according to some $(\pi^1, \pi^2) \in \Pi_1 \times \Pi_2 \subset \Gamma$.
3. Player 1 does not accept $\mu \in \Gamma$ and decides to use a stationary randomized strategy $\pi^1 \in \Pi^1$ instead, while player 2 approves the use of μ. Then, π^1 and μ are taken into account and the system (1) evolves with the control strategy $(\pi^1, \mu_2) \in \Pi_1 \times \Gamma_2 \subset \Gamma$, with μ_2 as the marginal distribution of μ on U_2.
4. Player 2 does not accept $\mu \in \Gamma$ and decides to use a stationary randomized strategy $\pi^2 \in \Pi^2$ instead, while player 1 approves the use of μ. Then, π^2 and μ are taken into account and the system (1) evolves according to the pair $(\mu_1, \pi^2) \in \Gamma_1 \times \Pi_2$, with μ_1 as the marginal distribution of μ on U_1.

The next definition extends the concept of a Nash equilibrium for the larger set of join strategies Γ.

Definition 9 (Correlated equilibria). Let F_k be either payoffs V_k or J_k, defined in (6) and (8), respectively. A correlated randomized strategy $\mu \in \Gamma$ is a correlated equilibrium for the extended payoff F_k if and only if

$$F_1(x,\mu) \geq F_1(x,\pi^1 \times \mu_2) \quad \forall\, \pi^1 \in \Pi_1,$$

$$F_2(x,\mu) \geq F_2(x,\mu_1 \times \pi^2) \quad \forall\, \pi^2 \in \Pi_2.$$

Existence of correlated equilibria always exists under our present hypotheses as it is established next:

Theorem 1. *(a) Under Assumptions 1, 2, and 3, there exists a correlated equilibrium associated to the payoff V_k in* (6).

(b) If Assumption 4 is also considered, then the existence of a correlated equilibrium for the payoff J_k in (8) *is also achieved.*

To prove this theorem, we are going to describe some auxiliary results.

The auxiliary zero-sum game: Consider the set

$$\Theta = \left\{(\pi^1 \times \pi^2, \lambda_1, \lambda_2) \;:\; \lambda_1, \lambda_2 > 0,\; \lambda_1 + \lambda_2 = 1,\; \pi^1 \times \pi^2 \in \Pi_1 \times \Pi_2\right\}.$$

We assume that we have two virtual players, say players A and B, so that the set of correlated randomized strategies Γ is the set of strategies for player A, whereas that Θ is the set of strategies for player B. The common payoff for both players is given by

$$\begin{aligned}
&G_F(x,\mu,\pi^1 \times \pi^2, \lambda_1, \lambda_2) := \\
&\lambda_1 \left[\int_{U_1 \times U_2} F_1(x,u_1,u_2)\mu(d(u_1,u_2)|x) - \int_{U_1 \times U_2} F_1(x,u_1,u_2)\pi^1(du_1|x)\mu_2(du_2|x)\right] \\
&+ \lambda_2 \left[\int_{U_1 \times U_2} F_2(x,u_1,u_2)\mu(d(u_1,u_2)|x) - \int_{U_1 \times U_2} F_2(x,u_1,u_2)\mu_1(du_1|x)\pi^2(du_2|x)\right],
\end{aligned} \tag{19}$$

where F_k denotes either V_k or J_k, for $k = 1,2$ and the subscript F of G simply refers the dependence of G with the $F_k' s$.

In virtue of the notation in (6) or (8), the payoff given in (19) can be rewritten as

$$\begin{aligned}
G_F(x,\mu,\pi^1 \times \pi^2, \lambda_1, \lambda_2) = {}& \lambda_1 \left[F_1(x,\mu) - F_1\left(x,\pi^1 \times \mu_2\right)\right] \\
&+ \lambda_2 \left[F_2(x,\mu) - F_2\left(x,\mu_1 \times \pi^2\right)\right].
\end{aligned} \tag{20}$$

Value of the game: In zero-sum games, the functions

$$\mathtt{U}(x) := \inf_{(\pi^1 \times \pi^2, \lambda_1, \lambda_2) \in \Theta} \sup_{\mu \in \Gamma} G_F(x,\mu,\pi^1 \times \pi^2, \lambda_1, \lambda_2) \quad \text{and}$$

$$\mathtt{L}(x) := \sup_{\mu \in \Gamma} \inf_{(\pi^1 \times \pi^2, \lambda_1, \lambda_2) \in \Theta} G_F(x,\mu,\pi^1 \times \pi^2, \lambda_1, \lambda_2),$$

play an important role. The function $\mathtt{L}$ is called the game's *lower value*, and $\mathtt{U}$ is the game's *upper value*. Clearly, we have $\mathtt{L} \leq \mathtt{U}$. If the upper and lower values coincide, then the game is said to have a *value*, and the *value of the game*, denoted as $\mathtt{V}$, is the common value of $\mathtt{L}$ and $\mathtt{U}$, i.e.,

$$\mathtt{V} := \mathtt{L} = \mathtt{U}.$$

Definition 10. Let X be a nonempty Hausdorff space and let $g : X \mapsto \mathbb{R}$ be a real-valued function. We say that g is affine-like function if and only if, for every $x_1, x_2 \in X$ and $\beta \in [0,1]$, there exists $x_\beta \in X$ such that $g(x_\beta) = \beta g(x_1) + (1-\beta) g(x_2)$.

The following proposition shows some properties of the payoff functions G_V and G_J.

Proposition 6. *(a) Suppose that Assumptions 1, 2 and 3 hold true. Then, the mapping $\mu \mapsto G_V(\cdot,\mu,\cdot,\cdot,\cdot)$ is continuous and affine-like on Γ. Furthermore, the mapping $(\pi^1 \times \pi^2, \lambda_1, \lambda_2) \mapsto G_V(\cdot,\cdot,\pi^1 \times \pi^2, \lambda_1, \lambda_2)$ is affine-like on Θ.*
(b) If in addition Assumption 4 is satisfied, then the same assertion in (a) is true for the payoff G_J.

Proof. (a) First, let us prove the continuity: Consider the sequence $\{\mu_n\} \subset \Gamma$ such that $\mu_n \stackrel{W}{\to} \mu$. Observe that

$$\begin{aligned} 0 &\leq |G_V(x,\mu_n,\pi^1 \times \pi^2,\lambda_1,\lambda_2) - G_V(x,\mu,\pi^1 \times \pi^2,\lambda_1,\lambda_2)| \\ &\leq \lambda_1 |V_1(x,\mu_n) - V_1(x,\mu)| + \lambda_2 |V_2(x,\mu_n) - V_2(x,\mu)| + \lambda_1 |V_1(x,\pi^1 \times \mu_{2n}) \\ &\quad - V_1(x,\pi^1 \times \mu_2)| + \lambda_2 |V_2(x,\mu_{1n} \times \pi^2) - V_2(x,\mu_1 \times \pi^2)|. \end{aligned} \tag{21}$$

Then, in virtue of Proposition 3, the terms in the right-hand side of (21) converge to zero as $\mu_n \stackrel{W}{\to} \mu$. So, G_V is continuous in Γ.

On the other hand, it is well-known that the discount payoff V_k can be seen as a linear mapping between r_k and the so-named occupation measure $\nu[x;\mu]$; i.e., $V_k(x,\mu) = \int r_k \, d\nu[x;\mu]$, for every $x \in \mathbb{R}^m$ and $\mu \in \Gamma$, where $\nu[x;\mu]$ is defined as

$$\int r_k \, \nu[x;\mu] = \alpha \mathbb{E}_x^\mu \Big[\int_0^\infty e^{-\alpha t} \int_{U_1 \times U_2} r_k(x(t),u_1,u_2) \mu\big(d(u_2,u_2)|x(t)\big) dt \Big]. \tag{22}$$

The details of this last fact can be extracted from page 1191 in [11] or from page 102 in [7]. Then by rewritting the payoff function V_k in the way of (22), it can be proved (see, for instance [11], page 1195) that, for any two strategies $\mu, \overline{\mu} \in \Gamma$ and $\beta \in [0,1]$, there exists another $\mu^\beta \in \Gamma$ so that

$$\nu[x;\mu^\beta] = \beta \nu[x;\mu] + (1-\beta)\nu[x;\overline{\mu}]. \tag{23}$$

This last property together with (22) yield that $V_k(x,\mu^\beta) = \beta V_k(x,\mu) + (1-\beta) V_k(x,\overline{\mu})$ and the choice of μ^β is independent of $k = 1,2$. With the previous ingredients, let us use the strategy $\mu^\beta \in \Gamma$ obtained by the affine-like property for both criteria V_1 and V_2, for some arbitrary choose of two strategies $\mu, \overline{\mu} \in \Gamma$ and $\beta \in [0,1]$. Then, the following is satisfied

$$\begin{aligned}G_V(x,\mu^\beta,\pi^1\times\pi^2,\lambda_1,\lambda_2) &= \lambda_1\left[V_1(x,\mu^\beta)-V_1\left(x,\pi^1\times\mu_2^\beta\right)\right]+\\ &\quad \lambda_2\left[V_2(x,\mu^\beta)-V_2\left(x,\mu_1^\beta\times\pi^2\right)\right]\\ &= \lambda_1\left[\{\beta V_1(x,\mu)+(1-\beta)V_1(x,\overline{\mu})\}-V_1\left(x,\pi^1\times\mu_2^\beta\right)\right]+\\ &\quad +\lambda_2\left[\{\beta V_2(x,\mu)+(1-\beta)V_2(x,\overline{\mu})\}-V_2\left(x,\mu_1^\beta\times\pi^2\right)\right].\end{aligned} \tag{24}$$

In addition, by following the same arguments of page 1195 in [11], it can be also verified that

$$\begin{aligned}V_1\left(x,\pi^1\times\mu_2^\beta\right) &= \beta V_1\left(x,\pi^1\times\mu_2\right)+(1-\beta)V_1\left(x,\pi^1\times\overline{\mu}_2\right) \quad \text{and}\\ V_2\left(x,\mu_1^\beta\times\pi^2\right) &= \beta V_2\left(x,\mu_1\times\pi^2\right)+(1-\beta)V_2\left(x,\overline{\mu}_1\times\pi^2\right).\end{aligned} \tag{25}$$

Combining (25) with (24) we deduce

$$\begin{aligned}G_V(x,\mu^\beta,\pi^1\times\pi^2,\lambda_1,\lambda_2) &= \beta G_V(x,\mu,\pi^1\times\pi^2,\lambda_1,\lambda_2)\\ &\quad +(1-\beta)G_V(x,\overline{\mu},\pi^1\times\pi^2,\lambda_1,\lambda_2),\end{aligned}$$

for all $(\pi^1\times\pi^2,\lambda_1,\lambda_2)\in\Theta$. This proves the affine-like property of G_V on Γ.

On the other hand, for any $\beta\in[0,1]$ and any $(\pi^1\times\pi^2,\lambda_1,\lambda_2)$, $(\overline{\pi}^1\times\overline{\pi}^2,\overline{\lambda}_1,\overline{\lambda}_2)$ $\in\Theta$ consider the following strategies $\pi_\beta^k\in\Pi_k$ and constants $\lambda_k^\beta\in\mathbb{R}$ $(k=1,2)$:

$$\pi_\beta^k := \frac{\beta\lambda_k\pi^k+(1-\beta)\overline{\lambda}_k\overline{\pi}_k}{\beta\lambda_k+(1-\beta)\overline{\lambda}_k}, \quad \text{and} \quad \lambda_k^\beta := \beta\lambda_k+(1-\beta)\overline{\lambda}_k.$$

Plugging these elements into G_V, it is easy to check that for each $\mu\in\Gamma$ and $x\in\mathbb{R}^m$,

$$\begin{aligned}G_V(x,\mu,\pi_\beta^1\times\pi_\beta^2,\lambda_1^\beta,\lambda_2^\beta) &= \beta G_V(x,\mu,\pi^1\times\pi^2,\lambda_1,\lambda_2)+\\ &\quad (1-\beta)G_V(x,\mu,\overline{\pi}^1\times\overline{\pi}^2,\overline{\lambda}_1,\overline{\lambda}_2).\end{aligned}$$

This proves that G_V is affine-like on Θ.

(b) As for the continuity of G_J on Γ, the proof is the same as in part (a), the only difference lies in replacing V_k by J_k and just use Proposition 5 in lieu of Proposition 3. Furthermore, there are works asserting that the average payoff J_k $(k=1,2)$ can be rewritten in terms of an occupation measure $\rho[\mu]$; i.e., for all $x\in\mathbb{R}^m$ and each $\mu\in\Gamma$, $J_k(x,\mu)=\int r_k d\rho[\mu]$, where $\rho[\mu](dy,du):=\eta_\mu(dy)\mu(du|y)$, with η_μ being the invariant measure defined in (9) (for further details see for instance, [2], page 87 or [7], page 91). Using the argumets as in page 92 of [7], we can obtain exactly the same property as (23) for ρ rather that ν. Then, it is straightforward that the mapping $\mu\mapsto J_k(\cdot,\mu)$ is affine-like on Γ. To prove the affine-like property of G_J, we proceed in the same way as (24). We can use also the same procedures of page 92 of [7] to get a similar relation of (25) associated to J_k. These previous properties would prove that G_J is affine-like on Γ after doing basic estimates. The proof that G_J is affine-like on Θ is similar to the one presented for G_V so we shall omit it. □

Proposition 7. *The upper value* $\mathtt{U}$ *is nonnegative.*

Proof. Clearly we know that

$$\sup_{\mu\in\Gamma} G_F(\cdot,\mu,\pi^1\times\pi^2,\lambda_1,\lambda_2) \geq G_F(\cdot,\mu,\pi^1\times\pi^2,\lambda_1,\lambda_2)$$
$$\forall\, \mu\in\Gamma,\ (\pi^1\times\pi^2,\lambda_1,\lambda_2)\in\Theta.$$

Then taking in particular $\hat{\mu} := \pi^1\times\pi^2$, we obtain that $G_F(\cdot,\hat{\mu},\pi^1\times\pi^2,\lambda_1,\lambda_2) = 0$. Therefore,

$$\sup_{\mu\in\Gamma} G_F(\cdot,\mu,\pi^1\times\pi^2,\lambda_1,\lambda_2) \geq G_F(\cdot,\hat{\mu},\pi^1\times\pi^2,\lambda_1,\lambda_2) = 0,$$
$$\forall\, (\pi^1\times\pi^2,\lambda_1,\lambda_2)\in\Theta.$$

This implies that

$$\mathtt{U}(\cdot) := \inf_{(\pi^1\times\pi^2,\lambda_1,\lambda_2)\in\Theta}\ \sup_{\mu\in\Gamma} G_F(\cdot,\mu,\pi^1\times\pi^2,\lambda_1,\lambda_2) \geq 0. \tag{26}$$

□

Proof of Theorem 1. (a) First note that Proposition 6 gives the hypotheses to get the Isaac's condition (see, for instance pages 108-109 in [20])

$$\begin{aligned}
&\inf_{(\pi^1\times\pi^2,\lambda_1,\lambda_2)\in\Theta}\ \sup_{\mu\in\Gamma} G_V(x,\mu,\pi^1\times\pi^2,\lambda_1,\lambda_2) = \\
&= \sup_{\mu\in\Gamma}\ \inf_{(\pi^1\times\pi^2,\lambda_1,\lambda_2)\in\Theta} G_V(x,\mu,\pi^1\times\pi^2,\lambda_1,\lambda_2), \quad x\in\mathbb{R}^m.
\end{aligned} \tag{27}$$

Relations (27) and (26) gives us that

$$\sup_{\mu\in\Gamma}\ \inf_{(\pi^1\times\pi^2,\lambda_1,\lambda_2)\in\Theta} G_V(x,\mu,\pi^1\times\pi^2,\lambda_1,\lambda_2) \geq 0 \quad \forall\, x\in\mathbb{R}^m.$$

As $\mu\mapsto G_V(\cdot,\mu,\cdot,\cdot,\cdot)$ is continuous, then it easy to verify that

$$\mu \mapsto \inf_{(\pi^1\times\pi^2,\lambda_1,\lambda_2)\in\Theta} G_V(\cdot,\mu,\pi^1\times\pi^2,\lambda_1,\lambda_2)$$

is upper semi-continuous. This last property together with the compactness of Γ imply the existence of $\mu^*\in\Gamma$ (that depends only of $x\in\mathbb{R}^m$) such that

$$G_V(x,\mu^*,\pi^1\times\pi^2,\lambda_1,\lambda_2) \geq 0, \quad \forall\, (\pi^1\times\pi^2,\lambda_1,\lambda_2)\in\Theta,\ x\in\mathbb{R}^m. \tag{28}$$

In virtue of (28), if we let $\lambda_2\to 0$ in (19) or (20) (yielding that $\lambda_1\to 1$), we get

$$V_1(x,\mu^*) - V_1(x,\pi^1\times\mu_2^*) \geq 0, \text{ for all } \pi^1\in\Pi^1.$$

Similarly, by letting $\lambda_1\to 0$ (yielding that $\lambda_2\to 1$), we can also deduce

$$V_2(x,\mu^*) - V_2(x,\mu_1^* \times \pi^2) \geq 0 \text{ for all } \pi^2 \in \Pi^2.$$

Thus, from the Definition 9, $\mu^* \in \Gamma$ becomes a correlated equilibrium.

The proof of part (b) is identical than (a), the only difference lies in the fact that we need Assumption 4 as an extra hypothesis to guarantee the continuity for G_J. □

Appendix

The main objective of this appendix is to prove that the convergence $\mu_n \xrightarrow{W} \mu$, $\alpha_n \to \alpha$, and $h_n \to h$ (this later convergence in a suitable sense), yield that, for each $k = 1,2$,

$$\lim_{n\to\infty} \left\{ r_k(\cdot,\mu_n) + L^{\mu_n} h_n - \alpha_n h_n \right\} = r_k(\cdot,\mu) + L^{\mu} h - \alpha h. \tag{29}$$

Let $\mathscr{O}$ be an open, bounded and connected subset of $\mathbb{R}^m$. We denote the closure of this set by $\bar{\mathscr{O}}$.

For every $x \in \mathbb{R}^m$, $\mu \in \Gamma$, $\alpha > 0$, h in $\mathbb{W}^{2,p}(\mathscr{O})$, we define

$$\hat{\Psi}(x,\mu,\alpha;h) := r_k(x,\mu) + \sum_{i=1}^{n} b_i(x,\mu) \frac{\partial h}{\partial x_i}(x) - \alpha h(x),$$

$$\mathscr{L}_{\alpha}^{\mu} h(x) := \hat{\Psi}(x,\mu,\alpha;h) + \frac{1}{2} \sum_{i,j=1}^{m} a^{ij}(x) \frac{\partial^2 h}{\partial x_i \partial x_j}(x), \tag{30}$$

where b_i is the i-th component of the function b defined in (1) and a as in Assumption 1(d).

The following theorem establishes the limit result referred in (29).

Theorem 2. *Let $\mathscr{O}$ be a bounded $\mathscr{C}^2$ domain. Suppose that there exist sequences $\{h_n\} \in \mathbb{W}^{2,p}(\mathscr{O})$, $\{\xi_n\} \in \mathbb{L}^p(\mathscr{O})$, with $p > m$ (m is the dimension of (1)), $\{\mu_n\} \in \Gamma$, and $\{\alpha_n\} \geq 0$, satisfying the following:*

(a) $\mathscr{L}_{\alpha_n}^{\mu_n} h_n = \xi_n$ in $\mathscr{O}$ for $n = 1,2,\ldots$
(b) There exists a constant $\tilde{M}_1$ such that $\|h_n\|_{\mathbb{W}^{2,p}(\mathscr{O})} \leq \tilde{M}_1$ for $n = 1,2,\ldots$
(c) ξ_n converges in $\mathbb{L}^p(\mathscr{O})$ to some function ξ.
(d) α_n converges to some constant $\alpha \geq 0$.
(e) $\mu_n \xrightarrow{W} \mu \in \Gamma$.

Then, there exist a function $h \in \mathbb{W}^{2,p}(\mathscr{O})$ and a subsequence $\{n_r\} \subset \{1,2,\ldots\}$ such that $h_{n_r} \to h$ in the norm of $\mathbb{C}^{1,\eta}(\bar{\mathscr{O}})$ for $\eta < 1 - \frac{m}{p}$ as $r \to \infty$. Moreover,

$$\mathscr{L}_{\alpha}^{\mu} h = \xi \;\; \text{in } \mathscr{O}. \tag{31}$$

Proof. We first show that there exist a function h in $\mathbb{W}^{2,p}(\mathscr{O})$ and a subsequence $\{n_r\} \subset \{1,2,...\}$ such that, as $r \to \infty$, $h_{n_r} \to h$ weakly in $\mathbb{W}^{2,p}(\mathscr{O})$ and strongly in

$\mathbb{C}^{1,\eta}(\bar{\mathscr{O}})$. Namely, since $\mathbb{W}^{2,p}(\mathscr{O})$ is reflexive (see [1, Theorem 3.5]), then, using Theorem 1.17 in the same reference [1], the ball

$$H := \left\{ h \in \mathbb{W}^{2,p}(\mathscr{O}) : \|h\|_{\mathbb{W}^{2,p}(\mathscr{O})} \leq \tilde{M} \right\} \tag{32}$$

is weakly sequentially compact. On the other hand, since $p > m$, by [1, Theorem 6.2, Part III], the imbedding $\mathbb{W}^{2,p}(\mathscr{O}) \hookrightarrow \mathbb{C}^{1,\eta}(\bar{\mathscr{O}})$, for $0 \leq \eta < 1 - \frac{m}{p}$ is compact; hence, it is also continuous, and thus the set H in (32) is relatively compact in $\mathbb{C}^{1,\eta}(\bar{\mathscr{O}})$. This fact ensures the existence of a function $h \in \mathbb{W}^{2,p}(\mathscr{O})$ and a subsequence $\{h_{n_r}\} \equiv \{h_n\} \subset H$ such that

$$h_n \to h \text{ weakly in } \mathbb{W}^{2,p}(\mathscr{O}) \text{ and strongly in } \mathbb{C}^{1,\eta}(\bar{\mathscr{O}}). \tag{33}$$

The second step is to show that, as $n \to \infty$,

$$\int_{\mathscr{O}} g(x)\hat{\Psi}(x,\mu_n,\alpha_n,;h_n)dx \to \int_{\mathscr{O}} g(x)\hat{\Psi}(x,\mu,\alpha;h)dx \quad \text{for all } g \in \mathbb{L}^1(\mathscr{O}). \tag{34}$$

To this end, given $x \in \mathscr{O}$, $k = 1,2$, functions $h \in \mathbb{W}^{2,p}(\mathscr{O})$ and $h_n \in H$, $\mu, \mu_n \in \Gamma$, and constants $\alpha_n,\ \alpha \geq 0$, the following holds for all $g \in \mathbb{L}^1(\mathscr{O})$.

$$\begin{aligned}
&\left| \int_{\mathscr{O}} g(x)\hat{\Psi}(x,\mu_n,\alpha_n;h_n)dx - \int_{\mathscr{O}} g(x)\hat{\Psi}(x,\mu,\alpha;h)dx \right| \\
&\leq \left| \int_{\mathscr{O}} g(x)\left[r_k(x,\mu_n) - r_k(x,\mu)\right]dx \right| \\
&+ \sum_{i=1}^{m} \left| \int_{\mathscr{O}} g(x)\left[b_i(x,\mu_n)\frac{\partial h_n}{\partial x_i}(x) - b_i(x,\mu)\frac{\partial h}{\partial x_i}(x)\right]dx \right| \\
&+ \left| \int_{\mathscr{O}} g(x)\left[\alpha_n h_n(x) - \alpha h(x)\right]dx \right| \\
&\leq \left| \int_{\mathscr{O}} g(x) r_k(x,\mu_n)dx - \int_{\mathscr{O}} g(x) r_k(x,\mu)dx \right| \\
&+ \sum_{i=1}^{m} \left| \int_{\mathscr{O}} g(x)\frac{\partial h_n}{\partial x_i}(x)\left[b_i(x,\mu_n) - b_i(x,\mu)\right]dx \right| \\
&+ \sum_{i=1}^{m} \left| \int_{\mathscr{O}} g(x) b_i(x,\mu_n)\left[\frac{\partial h_n}{\partial x_i}(x) - \frac{\partial h}{\partial x_i}(x)\right]dx \right| + |\alpha_n - \alpha| \left| \int_{\mathscr{O}} g(x) h_n(x)dx \right| \\
&+ \alpha \left| \int_{\mathscr{O}} g(x)\left[h_n(x) - h(x)\right]dx \right|.
\end{aligned}$$

Since the embedding $\mathbb{W}^{2,p}(\mathscr{O}) \hookrightarrow \mathbb{C}^{1,\eta}(\bar{\mathscr{O}})$ is continuous, hypothesis (b) together with the definition of the norm $\|\cdot\|_{\mathbb{C}^{1,\eta}(\bar{\mathscr{O}})}$, imply that there is a constant $\bar{M} > 0$ such that

$$\max\left\{ |h_n|, \max_{1\leq i\leq m} \left|\frac{\partial h_n}{\partial x_i}\right| \right\} \leq \|h_n\|_{\mathbb{C}^{1,\eta}(\bar{\mathscr{O}})} \leq \bar{M}\,\|h_n\|_{\mathbb{W}^{2,p}(\mathscr{O})} \leq \bar{M}\tilde{M}_1.$$

On the other hand, it is easy to verify that Assumptions 1 and 3, yield that $|b(\cdot,\mu)|+|r_k(\cdot,\mu)|\leq K(\bar{\mathscr{O}})$. Hence,

$$\begin{aligned}
&\left|\int_{\mathscr{O}} g(x)\hat{\Psi}(x,\mu_n,\alpha_n;h_n)dx-\int_{\mathscr{O}} g(x)\hat{\Psi}(x,\mu,\alpha;h)dx\right|\leq\\
&\left|\int_{\mathscr{O}} g(x)r_k(x,\mu_n)dx-\int_{\mathscr{O}} g(x)r_k(x,\mu)dx\right|\\
&+\bar{M}\tilde{M}_1 m\max_{1\leq i\leq m}\left|\int_{\mathscr{O}} g(x)\left[b_i(x,\mu_n)-b_i(x,\mu)\right]dx\right|\\
&+\|g\|_{\mathbb{L}^1(\mathscr{O})}\|h_n-h\|_{\mathbb{C}^{1,\eta}(\bar{\mathscr{O}})}(mK(\bar{\mathscr{O}})+\alpha)+|\alpha_n-\alpha|\bar{M}\tilde{M}_1\|g\|_{\mathbb{L}^1(\mathscr{O})}. \qquad (35)
\end{aligned}$$

Observe that $r_k(\cdot,\mu)$ $k=1,2$, and $b_i(\cdot,\mu)$ $i=1,\cdots,m$ are bounded on $\bar{\mathscr{O}}$. Then, hypotheses (d) to (e), together with (33), lead to the right hand side of (35) goes to zero as $n\to\infty$, thus proving (34).

The existence of the constant $K(\bar{\mathscr{O}})$ used for the analysis in (35) can be also used to get also that $|\sigma(x)|\leq K(\bar{\mathscr{O}})$, then we can affirm that for each g in $\mathbb{L}^{\frac{p}{p-1}}(\mathscr{O})$,

$$\begin{aligned}
&\tfrac{1}{2}\left|\int_{\mathscr{O}} g(x)\left[\sum_{i,j=1}^m a^{ij}(x)\tfrac{\partial^2 h_n}{\partial x_i\partial x_j}(x)-\sum_{i,j=1}^m a^{ij}(x)\tfrac{\partial^2 h}{\partial x_i\partial x_j}(x)\right]dx\right|\\
&\quad\leq\tfrac{m^2}{2}\left[K(\bar{\mathscr{O}})\right]^2\sum_{i,j=1}^m\left|\int_{\mathscr{O}} g(x)\left[\tfrac{\partial^2 h_n}{\partial x_i\partial x_j}(x)-\tfrac{\partial^2 h}{\partial x_i\partial x_j}(x)\right]dx\right|. \qquad (36)
\end{aligned}$$

Thus the weak convergence of $\{h_n\}$ to h in $\mathbb{W}^{2,p}(\mathscr{O})$ yields that the right–hand side of (36) converges to zero as $n\to\infty$. Notice also that the convergence of (34) is also valid for all $g\in\mathbb{L}^{\frac{p}{p-1}}(\mathscr{O})$. The reason is because $\mathbb{L}^{\frac{p}{p-1}}(\mathscr{O})\subset\mathbb{L}^1(\mathscr{O})$ (recall the Lebesgue measure on $\mathscr{O}$ is bounded). This last fact together with (36) and hypothesis (c), yield that for every g in $\mathbb{L}^{\frac{p}{p-1}}(\mathscr{O})$,

$$\int_{\mathscr{O}} g(x)\left[\mathscr{L}_\alpha^\mu h(x)-\xi(x)\right]dx=\lim_{n\to\infty}\int_{\mathscr{O}} g(x)\left[\mathscr{L}_{\alpha_n}^{\mu_n}(x)-\xi_n(x)\right]dx=0.$$

The above limit, along with Theorem 2.10 in [18], implies (31), i.e.

$$\mathscr{L}_\alpha^\mu h=\xi \text{ in } \mathscr{O}.$$

This completes the proof. □

References

1. Adams, R.A.: *Sobolev Spaces*. Academic Press. New York, (1975).
2. Arapostathis, A.; Ghosh, M.K.; Borkar, V.S.: *Ergodic control of diffusion processes*. Encyclopedia of Mathematics and its Applications, 143. Cambridge University Press, (2012).
3. Arapostathis, A.; Borkar, V. S. : Uniform recurrence properties of controlled diffusions and applications to optimal control. SIAM J. Control Optim. 48, 4181-4223 (2010).

4. Aumann, R. J.: *Subjectivity and correlation in randomized games*. Econometrica 30, 445462, (1974).
5. Aumann, R. J.: *Correlated equilibrium as an expression of Bayesian rationality*. Econometrica 55, 118, (1987).
6. Borkar, V.S.: *A topology for Markov controls.* Appl. Math. Optim. 20, 55-62, (1989).
7. Borkar, V. S.; Ghosh, M. K.: *Controlled diffusion processes with constraints*. J. Math.Anal. Appl., 152, 88-108. (1990)
8. Borkar, V. S.; Ghosh, M. K. (1992): *Stochastic differential games: occupation measure based approach*. J. Optim. Theory Appl., 73, 359385. Correction: 88, 251252, (1996).
9. Escobedo-Trujillo, B.A.; López-Barrientos, J.D.; Hernández-Lerma, O.: *Bias and overtaking equilibria for zero-sum stochastic differential games*. J. Optim. Theory Appl. 153, 662-687, (2012).
10. Forges, F.: *Five legitimate definitions of correlated equilibrium in games with incomplete information*.Theory and Decision 35, 277-310, (1993).
11. Ghosh, M.K.; Arapostathis, A.; Marcus, S.I.: *Optimal control of switching diffusions with applications to flexible manufacturing systems*. SIAM J. Control Optim. 30, 1-23, (1992).
12. Ghosh, M.K.; Arapostathis, A.; Marcus, S.I.: *Ergodic control of switching diffusions*. SIAM J. Control Optim. 35, 1962-1988, (1997).
13. Gilbarg, D.; Trudinger, N.S.: *Elliptic partial differential equations of second order*. Reprinted version. Heildelberg, Springer, (1998).
14. Jasso-Fuentes, H.; Hernández-Lerma, O.: *Characterizations of overtaking optimality for controlled diffusion processes*. Appl. Math. Optim. 57, 349-369, (2008).
15. Jasso-Fuentes, H.; Yin, G.G.: *Advanced criteria for controlled Markov-modulated diffusions in an infinite horizon: overtaking, bias and Blackwell optimality*. Science Press, Beijing China, (2013).
16. Jasso-Fuentes, H.; López-Barrientos D.; Escobedo-Trujillo B.A.: *Infinite horizon nonzero-sum stochastic differential games with additive structure*. IMA J. Math. Control Inform. 34, 283-309,(2017).
17. Küenle, H. U.: *Equilibrium strategies in stochastic games with additive cost and transition structure*. Int. Game Theory Rev., 1, 131147, (1999) .
18. Lieb, E.H.; Loss, M.: *Analysis*. Second Edition. AMS. Providence, Rhode Island, (2001)..
19. Neyman, A.:*Correlated Equilibrium and potential games*. International Journal of Game Theory. 26, 223-227, (1997).
20. Nowak, A. S.: *Correlated relaxed equilibria in nonzero-sum linear differential games*. J. Math. Anal. Appl. 163, 104–112, (1992).
21. Nowak, A. S.: *Correlated equilibria in nonzero-sum differential games*. J. Math. Anal. Appl. 174, 539–549, (1993) (1994).
22. Solan, E.: *Characterizacion of correlated equilibria in stochastic games*. Int. J. Game Theory 30, 259-277, (2001).
23. Soltz, G.; Lugosi G.: *Learning correlated equilibria in games with compact sets of strategies.* Games and Economic Behavior 59, 187-208, (2007).
24. Stein, N.D.; Parrilo P. A.; Ozdaglar A.: *Correlated equilibria in continuous games: Characterization and computation.* Games and Economic Behavior 71, 436-455, (2011).
25. Warga, J.:*Optimal control of differential and functional equations*. Academic Press, New York, (1972).