

Blackwell-Nash Equilibria in Zero-Sum Stochastic Differential Games



Beatris Adriana Escobedo-Trujillo, Héctor Jasso-Fuentes,
and José Daniel López-Barrientos

Abstract Advanced-type equilibria for a general class of zero-sum stochastic differential games have been studied in part by Escobedo-Trujillo et al. (J Optim Theory Appl 153:662–687, 2012), in which a comprehensive study of the so-named *bias and overtaking equilibria* was provided. On the other hand, a complete analysis of advanced optimality criteria in the context of optimal control theory such as bias, overtaking, sensitive discount, and Blackwell optimality was developed independently by Jasso-Fuentes and Hernández-Lerma (Appl Math Optim 57:349–369, 2008; J Appl Probab 46:372–391, 2009; Stoch Anal Appl 27:363–385, 2009). In this work we try to fill out the gap between the aforementioned references. Namely, the aim is to analyze Blackwell-Nash equilibria for a general class of zero-sum stochastic differential games. Our approach is based on the use of dynamic programming, the Laurent series and the study of sensitive discount optimality.

Keywords Zero-sum stochastic differential games · Average equilibrium · Bias equilibrium · Laurent series · Blackwell-Nash equilibrium

2010 Mathematics Subject Classification 91A10, 91A15, 91A25

B. A. Escobedo-Trujillo
Engineering Faculty, Universidad Veracruzana, Coatzacoalcos, Veracruz, México
e-mail: bescobedo@uv.mx

H. Jasso-Fuentes (✉)
Departamento de Matemáticas, CINVESTAV-IPN, México, D.F., México
e-mail: hjasso@math.cinvestav.mx

J. D. López-Barrientos
Facultad de Ciencias Actuariales, Universidad Anáhuac México, Huixquilucan, Estado de México, México
e-mail: daniel.lopez@anahuac.mx

1 Introduction

Among the most common payoff functions existing in the literature in the general theory of dynamic games we can mention the (finite-horizon) Bolza-type payoff and the well-known (infinite-horizon) discounted and average payoffs. The key features of these two last criteria is that, whereas the discounted payoff only focuses on earlier revenues, the average reward ignores these and pays attention only to the asymptotic behavior of the utilities. A drawback of these points of view is that they do not consider what happens in the mid-run. For example, there can be N -tuples of strategies (N represents the number of players in the game) that might be “optimal” for all the players in the infinite-horizon average criterion, but in turn, they provide low profits (and/or high costs) to the players at *any* finite period of time. From another angle, there exist several applications in which the (infinite-horizon) discounted payoff criterion is used to model the real or present value (at the current time) of a company; the key ingredient is the use of a *discount factor*. However, in some other situations, this criterion might be used for some other purposes; one of them is to regard it as an estimate of criteria without discount.

To fix ideas, suppose we have a game such that $\bar{\pi} = (\pi_1 \cdots, \pi_N)$ represents an N -tuple associated to some choices of the players (i.e., π_i corresponds to the strategy of player i), and denote by r_i the associated payoff rate function of player i (for illustrative purposes let us assume for the moment that all players have the same reward rate; i.e., $r_i = r$ for all $i = 1, \cdots, N$). The expected undiscounted and discounted payoffs of $\bar{\pi}$ for each player are defined, respectively, as

$$V(\bar{\pi}) = E \int_0^{\infty} r(x^{\bar{\pi}}(t))dt, \quad \text{and} \quad V_{\alpha}(\bar{\pi}) = E \int_0^{\infty} e^{-\alpha t} r(x^{\bar{\pi}}(t))dt,$$

where $x^{\bar{\pi}}(t)$ represents the state of the process under the policy $\bar{\pi}$ at time t , and $\alpha > 0$ is a given constant. A very important property of V_{α} is that, under mild assumptions, it is finite-valued; whereas the former requires very strong hypotheses to possess this feature. In this sense, if one is interested in studying optimality under the criterion V , one may regard such criterion as the limit of some sequence of V_{α} in the following sense:

$$V_{\alpha_n}(\bar{\pi}) \rightarrow V(\bar{\pi}) \quad \text{as} \quad \{\alpha_n\}_n \downarrow 0. \quad (1.1)$$

However, even when one can provide optimality results (Nash equilibria) to V_{α} for some *fixed* and of course *positive* and even *small* α , it turns out that this V_{α} , regarded as an estimate of V , is acceptable at *early* periods of times, but it is very imprecise in the long run.

An alternative approach that lies in the same direction of the limit (1.1) is the use of Blackwell-Nash equilibria. This consists essentially in seeking Nash equilibria that remain *optimal* for *all* the discounted payoffs V_{α} , $0 < \alpha < \alpha^*$, for some fixed $\alpha^* > 0$ (see Definition 8.1). Due to the nature of this class of equilibria, they turn out to be good “optimizers”, when the payoff criterion under study is of type V .

The purpose of this work is to analyze Blackwell-Nash equilibria for a general class of zero-sum stochastic differential games; namely, we provide sufficient conditions for ensuring the existence and characterizations of these equilibria. This study is based on the analysis of the so-named sensitive discount equilibria introduced in Definition 8.2. It is worth noting that Blackwell-Nash equilibria have the property of being bias and overtaking equilibria too. In this sense, our present analysis is more general than [5], because we use the same set of assumptions. Finally, it is important to say that, due to the fact that our work studies only the zero-sum case, here and in the sequel, we consider only to the case $N = 2$ players.

Another interesting application concerning Blackwell games goes in the spirit of the so-named priority mean-payoff games, which are regarded as the limit of special multi-discounted games. In this type of games, Blackwell equilibria play an important role because of their stability property under small perturbations of the discount factor—see [7–9]. The study of Blackwell-Nash equilibria in zero-sum stochastic differential games also permits the extension to the theory of priority mean-payoff games in the stochastic differential games setting.

Bias and overtaking criteria have been studied in the context of zero-sum stochastic differential games; see, for example, [5, 17]. Nevertheless, to the best of our knowledge, the only works dealing with sensitive discount and Blackwell optimality, but in the context of controlled diffusions (i.e., the case of one player only) are [12, 13] and [22]. It is worth mentioning, however, that there are some works that are close to the present proposal. For instance, Arapostathis et al. [3] study a zero-sum stochastic differential game under a slightly different ergodicity assumption than ours. It states a parabolic Hamilton-Jacobi-Bellman (HJB) equation, and finds *risk-sensitive optimal selectors*, in the sense that the payoff form is “sensitive to higher moments of the running cost, and not merely its mean”. This represents an alternative approach to ours, because while they deal with the concept of risk-sensitivity (as introduced in [25]), we rather choose the notion of *sensitive discount* in a Laurent series, as presented in [12] and [21]. Other works that are related to the selective criteria we study for stochastic diffusions are [5, 11–13, 17] and the references therein.

The rest of our work comprises eight short sections. In the next section we introduce the notation that we use, our game model, the main hypotheses, and the basic type of strategies we will deal with along our developments. Section 3 presents the long-run average optimality criterion, and a very well-known result on the existence of the corresponding Nash equilibria. Section 4 is devoted to the so called *bias criterion*. This is a first refinement of the criterion introduced in Sect. 3, and we profit from it by quoting the concepts introduced in that part in further sections. In Sect. 5 we extend the results from [21, Section 3] to the zero-sum case. There, we use an exponential ergodicity condition to characterize a discounted payoff in terms of a Laurent series. Sections 6–8 are extensions of the results from [12] and represent the main contribution of this paper. In Sect. 6 we define the so-called *Poisson system* and model its solution in terms of the criterion presented before in Sect. 3. Section 7 shows a connection between the Poisson system and the dynamic programming principle. There, we lay out the concept of *canonical*

equilibria and represent it as the strategies for which certain HJB equations are met. In Sect. 8 we exhibit Blackwell-Nash and sensitive discount equilibria and relate them in some appropriated sense. We draw our conclusions in Sect. 9.

2 The Game Model and Main Assumptions

The Dynamic System Let us consider an n -dimensional diffusion process $x(\cdot)$ controlled by two players and evolving according to the stochastic differential equation

$$dx(t) = b(x(t), u_1(t), u_2(t))dt + \sigma(x(t))dW(t), \quad x(0) = x_0, \quad t \geq 0, \quad (2.1)$$

where $b : \mathbb{R}^n \times U_1 \times U_2 \rightarrow \mathbb{R}^n$ and $\sigma : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times d}$ 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 given (Borel) sets. Moreover, for $i = 1, 2$, $u_i(\cdot)$ is a U_i -valued stochastic process representing the strategy of player i at each time $t \geq 0$.

Notation For vectors x and matrices A we consider the usual Euclidean norms

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

where A' and $\text{Tr}(\cdot)$ denote the transpose and the trace of a matrix, respectively.

Assumption 2.1

- (a) *The action sets U_1 and U_2 are compact.*
 (b) *$b(x, u_1, u_2)$ is continuous on $\mathbb{R}^n \times U_1 \times U_2$, and $x \mapsto b(x, u_1, u_2)$ 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*

$$\sup_{(u_1, u_2) \in U_1 \times U_2} |b(x, u_1, u_2) - b(y, u_1, u_2)| \leq K_1|x - y| \quad \text{for all } x, y \in \mathbb{R}^n.$$

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

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

- (d) *(Uniform ellipticity.) The matrix $a(x) := \sigma(x)\sigma'(x)$ satisfies that, for some constant $K_3 > 0$,*

$$x'a(y)x \geq K_3|x|^2 \quad \text{for all } x, y \in \mathbb{R}^n.$$

For $(u_1, u_2) \in U_1 \times U_2$, and v in $C^2(\mathbb{R}^n)$, $p \geq 1$, let

$$\mathcal{L}^{u_1, u_2} v(x) := \sum_{i=1}^n b^i(x, u_1, u_2) \partial_i v(x) + \frac{1}{2} \sum_{i, j=1}^n a^{ij}(x) \partial_{ij}^2 v(x), \quad (2.2)$$

where b^i is the i -th component of b , and a^{ij} is the (i, j) -component of the matrix $a(\cdot)$ defined in Assumption 2.1(d).

2.1 Strategies

Throughout this work, we will be interested in finding saddle points (see Theorem 3.5 below). To ensure that our search leads us to this result, we use the theory of relaxed controls—see for instance, [19, 24, 26]. The use of this class of controls, along with the semi-continuity properties of the cost/reward function (see Assumption 2.8(c) below) will give us the convex structure needed to guarantee the existence of non-cooperative Nash equilibria.

For each $k = 1, 2$, let $\mathcal{P}(U_k)$ be the space of probability measures on U_k endowed with the topology of weak convergence, and denote by $\mathcal{B}(U_k)$ the Borel σ -algebra of U_k .

Definition 2.2 A randomized strategy for player k is a family $\pi^k := \{\pi_t^k, t > 0\}$ of stochastic kernels on $\mathcal{B}(U_k) \times \mathbb{R}^n$ satisfying:

- (a) for each $t \geq 0$ and $x \in \mathbb{R}^n$, $\pi_t^k(\cdot|x)$ is a probability measure on U_k such that $\pi_t^k(U_k|x) = 1$, and for each $D \in \mathcal{B}(U_k)$, $\pi_t^k(D|\cdot)$ is a Borel function on \mathbb{R}^n ; and
- (b) for each $D \in \mathcal{B}(U_k)$ and $x \in \mathbb{R}^n$, the mapping $t \mapsto \pi_t^k(D|x)$ is Borel measurable.

We now introduce the notion of *stationary strategy*.

Definition 2.3 For each $k = 1, 2$, we say that a randomized strategy is *stationary* if and only if there is a probability measure $\pi^k(\cdot|x) \in \mathcal{P}(U_k)$ such that $\pi_t^k(\cdot|x) = \pi^k(\cdot|x)$ for all $x \in \mathbb{R}^n$ and $t \geq 0$.

The set of randomized stationary strategies for player $k = 1, 2$ is denoted by Π^k . It is important to state that we suppose the existence of a topology defined on Π^k , $k = 1, 2$, such that Π^k is compact—for more details see [14, Section 2].

For each pair of probability measures $(\phi, \psi) \in \mathcal{P}(U_1) \times \mathcal{P}(U_2)$ we write the drift coefficient b in (2.1) and the operator \mathcal{L} in (2.2) in terms of these measures by

means of the following expressions:

$$b(x, \phi, \psi) := \int_{U_2} \int_{U_1} b(x, u_1, u_2) \phi(du_1) \psi(du_2), \tag{2.3}$$

$$\mathcal{L}^{\phi, \psi} h(x) := \int_{U_2} \int_{U_1} \mathcal{L}^{u_1, u_2} h(x) \phi(du_1) \psi(du_2). \tag{2.4}$$

The notation above is valid also when the strategies $\pi^1 \in \Pi^1$ or/and $\pi^2 \in \Pi^2$ in (2.3)–(2.4) are interpreted as probability measures for each fixed $x \in \mathbb{R}^n$; that is, $\pi^k(\cdot|x) \in \mathcal{P}(U^k)$. In this case, unless the context requires further clarification, we shall simply write the “variable” π^k in the left-hand side of (2.3)–(2.4), rather than $\pi^k(\cdot|x)$.

Remark 2.4 Assumption 2.1 ensures that, for each pair of strategies $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$ there exists an almost surely unique strong solution of (2.1) which is a Markov-Feller process. Furthermore, for each pair of strategies $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$, the operator $\mathcal{L}^{\pi^1, \pi^2} v$ in (2.4) becomes the infinitesimal generator of (2.1). (For more details, see the arguments of [2, Theorem 2.2.12] or [6, Theorem 2.1].)

Sometimes we write $x(\cdot)$ as $x^{\pi^1, \pi^2}(\cdot)$ to emphasize the dependence on $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$. Also, we shall denote by $\mathbb{P}^{\pi^1, \pi^2}(t, x, \cdot)$ the corresponding transition probability of the process $x^{\pi^1, \pi^2}(\cdot)$, i.e., $\mathbb{P}^{\pi^1, \pi^2}(t, x, B) := \mathbb{P}(x^{\pi^1, \pi^2}(t) \in B | x(0) = x)$ for every Borel set $B \subset \mathbb{R}^n$ and $t \geq 0$. The symbol $\mathbb{E}_x^{\pi^1, \pi^2}(\cdot)$ stands for the associated conditional expectation.

Remark 2.5 In later sections, we will restrict ourselves to the space of stationary strategies within the class of randomized strategies. The reason is that the recurrence and ergodicity properties of the state system (2.1) can be easily verified through the use of such policies, but for a more general class of strategies (for instance, that of the so-called non-anticipative strategies), the corresponding state system might be time-inhomogeneous; which might present some technical difficulties. Thus even when it is possible to work with non-anticipative policies, our hypotheses ensure the existence of Nash equilibria in the class of stationary strategies for both players (see, [2, 15, 16]).

Definition 2.6 Let $\mathcal{O} \subset \mathbb{R}^n$ be an open set. We denote by $\mathcal{B}_w(\mathcal{O})$ the Banach space of real-valued measurable functions v on \mathcal{O} with finite w -norm defined as follows:

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

2.2 Recurrence and Ergodicity

Assumption 2.7 *There exists a function $w \in C^2(\mathbb{R}^n)$, with $w \geq 1$, and constants $d \geq c > 0$ such that*

- (i) $\lim_{|x| \rightarrow \infty} w(x) = +\infty$, and
- (ii) $\mathcal{L}^{\pi^1, \pi^2} w(x) \leq -cw(x) + d$ for each $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$ and $x \in \mathbb{R}^n$.

Assumption 2.7 ensures the existence of a unique invariant probability measure μ_{π^1, π^2} for the Markov process $x^{\pi^1, \pi^2}(\cdot)$, such that

$$\mu_{\pi^1, \pi^2}(w) := \int_{\mathbb{R}^n} w(x) \mu_{\pi^1, \pi^2}(dx) < \infty \text{ for all } (\pi^1, \pi^2) \in \Pi^1 \times \Pi^2. \quad (2.5)$$

(See [2, 18] for details.) Moreover, for every $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$, $x \in \mathbb{R}^n$, and $t \geq 0$, an application of Dynkin’s formula to the function $v(t, x) := e^{ct}w(x)$, and Assumption 2.7(ii) yield

$$\mathbb{E}_x^{\pi^1, \pi^2} w(x(t)) \leq e^{-ct} w(x) + \frac{d}{c}(1 - e^{-ct}). \quad (2.6)$$

Hence, integrating both sides of (2.6) with respect to the invariant measure μ_{π^1, π^2} leads to

$$\mu_{\pi^1, \pi^2}(w) \leq \frac{d}{c}. \quad (2.7)$$

Assumption 2.8 *The process $x^{\pi^1, \pi^2}(\cdot)$ in (2.1) is uniformly w -exponentially ergodic; that is, there exist constants $C > 0$ and $\delta > 0$ such that*

$$\sup_{(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2} |\mathbb{E}_x^{\pi^1, \pi^2} [g(x(t))] - \mu_{\pi^1, \pi^2}(g)| \leq Ce^{-\delta t} \|g\|_w w(x) \quad (2.8)$$

for all $x \in \mathbb{R}^n$, $t \geq 0$, and $g \in \mathcal{B}_w(\mathbb{R}^n)$. In this case, $\mu_{\pi^1, \pi^1}(g)$ equals the integral in (2.5) with g rather than w .

Sufficient conditions for ensuring the w -exponential ergodicity of process $x^{\pi^1, \pi^2}(\cdot)$ are given in [11, Theorem 2.7].

2.3 The Payoff Rate

Let $r : \mathbb{R}^n \times U_1 \times U_2 \rightarrow \mathbb{R}$ be a measurable function, so-named the payoff (or reward/cost) rate, which satisfies the following conditions:

Assumption 2.9

(a) *The function $r(x, u_1, u_2)$ is continuous on $\mathbb{R}^n \times U_1 \times U_2$ 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*

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

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

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

(c) *$r(x, u_1, u_2)$ is upper semicontinuous (u.s.c.) and concave in $u_1 \in U_1$ for every $(x, u_2) \in \mathbb{R}^n \times U_2$, and lower semicontinuous (l.s.c.) and convex in $u_2 \in U_2$ for every $(x, u_1) \in \mathbb{R}^n \times U_1$.*

Similar to (2.3)–(2.4), for each $(\phi, \psi) \in \mathcal{P}(U_1) \times \mathcal{P}(U_2)$ we write

$$r(x, \phi, \psi) := \int_{U_2} \int_{U_1} r(x, u_1, u_2)\phi(du_1)\psi(du_2), \quad x \in \mathbb{R}^n. \tag{2.9}$$

Note that this definition remains valid when the strategies $\pi^1 \in \Pi^1$ or/and $\pi^2 \in \Pi^2$ are applied in (2.9) as they are interpreted as probability measures, for each fixed $x \in \mathbb{R}^n$; that is, $\pi^k(\cdot|x) \in \mathcal{P}(U^k)$. As was agreed earlier, we shall simply write the “variable” π^k in the left-hand side of (2.9) rather than $\pi^k(\cdot|x)$.

Remark 2.10 Under Assumptions 2.1 and 2.9, the payoff rate $r(\cdot, \phi, \psi)$ and the infinitesimal generator $\mathcal{L}^{\phi, \psi} h(\cdot)$ (with $h \in C^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$) are u.s.c. in $\phi \in \mathcal{P}(U_1)$ and l.s.c. in $\psi \in \mathcal{P}(U_2)$. For further details see [5, Lemma 3.1].

3 Average Equilibria

We devote this section to the introduction of the basic optimality criterion we will use—and refine—along this study. We present the material in the spirit of [5, 11, 12, 17], and [20].

Definition 3.1 *The long-run average payoff* (also known as the *ergodic payoff*) when the players use the pair of strategies $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$ given the initial state x is

$$J(x, \pi^1, \pi^2) := \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}_x^{\pi^1, \pi^2} \left[\int_0^T r(x(t), \pi^1, \pi^2) dt \right]. \tag{3.1}$$

Given $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$, let us define the constant

$$J(\pi^1, \pi^2) := \mu_{\pi^1, \pi^2}(r(\cdot, \pi^1, \pi^2)) = \int_{\mathbb{R}^n} r(x, \pi^1, \pi^2) \mu_{\pi^1, \pi^2}(dx). \tag{3.2}$$

with μ_{π^1, π^2} as in (2.5). Under our set of assumptions, it follows from (2.8) and (3.2) that the average payoff (3.1) coincides with the constant $J(\pi^1, \pi^2)$ for every $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$ —see [5, p. 669]. Moreover, by the definition (3.2) of $J(\pi^1, \pi^2)$, together with Assumption 2.9(b) and (2.7)

$$|J(\pi^1, \pi^2)| \leq \int_{\mathbb{R}^n} |r(x, (\pi^1, \pi^2))| \mu_{\pi^1, \pi^2}(dx) \leq M \cdot \frac{d}{c} \quad \forall (\pi^1, \pi^2) \in \Pi^1 \times \Pi^2, \tag{3.3}$$

so that the constant $J(\pi^1, \pi^2)$ is uniformly bounded on $\Pi^1 \times \Pi^2$.

Value of the Game Let

$$L := \sup_{\pi^1 \in \Pi_1} \inf_{\pi^2 \in \Pi_2} J(\pi^1, \pi^2) \quad \text{and} \quad U := \inf_{\pi^2 \in \Pi_2} \sup_{\pi^1 \in \Pi_1} J(\pi^1, \pi^2)$$

The function L is said to be the game’s *lower value* whereas U is better known as the game’s *upper value*. Clearly, we have $L \leq U$. If the upper and lower values coincide, then the game is said to have a *value*, which we will denote by \mathcal{V} ; in other words,

$$\mathcal{V} = L = U. \tag{3.4}$$

As a consequence of (3.3), L and U are finite; and hence, so is \mathcal{V} if the second equality in (3.4) holds.

Definition 3.2 We say that a pair of stationary strategies $(\pi^{*1}, \pi^{*2}) \in \Pi^1 \times \Pi^2$ is an *average Nash equilibrium* (also known as an *average saddle point*) if

$$J(\pi^1, \pi^{*2}) \leq J(\pi^{*1}, \pi^{*2}) \leq J(\pi^{*1}, \pi^2) \quad \text{for every } (\pi^1, \pi^2) \in \Pi^1 \times \Pi^2.$$

The set of average saddle points pairs is denoted by $(\Pi^1 \times \Pi^2)_{ao}$.

Remark 3.3 Note that if $(\pi^{*1}, \pi^{*2}) \in \Pi^1 \times \Pi^2$ is an average Nash equilibrium (in case it does exist), then the game has a value $J(\pi^{*1}, \pi^{*2}) =: \mathcal{V}$ —see, for instance, [10, Proposition 4.2]. However, the converse is not necessarily true.

The following definition is crucial for our developments.

Definition 3.4 We say that a constant $J \in \mathbb{R}$, a function $h \in C^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$, and a pair of strategies $(\pi^{*1}, \pi^{*2}) \in \Pi^1 \times \Pi^2$ verify the *average payoff optimality equations* if, for every $x \in \mathbb{R}^n$,

$$J = r(x, \pi^{*1}, \pi^{*2}) + \mathcal{L}^{\pi^{*1}, \pi^{*2}} h(x) \quad (3.5)$$

$$= \sup_{\phi \in \mathcal{P}(U_1)} \{r(x, \phi, \pi^{*2}) + \mathcal{L}^{\phi, \pi^{*2}} h(x)\} \quad (3.6)$$

$$= \inf_{\psi \in \mathcal{P}(U_2)} \{r(x, \pi^{*1}, \psi) + \mathcal{L}^{\pi^{*1}, \psi} h(x)\} \quad \text{for all } x \in \mathbb{R}^n. \quad (3.7)$$

In this case, the pair of strategies $(\pi^{*1}, \pi^{*2}) \in \Pi^1 \times \Pi^2$ that satisfies (3.5)–(3.7) is called a pair of *canonical strategies*. We denote by $(\Pi^1 \times \Pi^2)_{ca}$ the family of canonical strategies.

Equation (3.5) is sometimes referred to as *Poisson equation*. This is the reason for which we call Eqs. (6.1)–(6.3) below, *Poisson system*.

The following result ensures the existence of solutions of Eqs.(3.5)–(3.7). It also states the existence of average saddle points, and provides us with their characterization. For a proof see [3, 5].

Theorem 3.5 *If Assumptions 2.1, 2.7, 2.8, and 2.9 hold, then:*

- (i) *There exist solutions $(J, h, (\pi^{*1}, \pi^{*2}))$ to the average payoff equations (3.5)–(3.7). Moreover, the constant J coincides with \mathcal{V} defined in (3.4), and the function h is unique up to additive constants; in fact, h is unique under the additional condition that $h(0) = 0$.*
- (ii) *A pair of strategies is an average saddle point if, and only if, it is canonical, that is, $(\Pi^1 \times \Pi^2)_{ao} = (\Pi^1 \times \Pi^2)_{ca}$.*

Remark 3.6 One important aspect in the proof of the last result is that Remark 2.10 ensures that the mapping $\phi \mapsto r(x, \phi, \psi) + \mathcal{L}^{\phi, \psi} h(x)$ is u.s.c. on the compact set $\mathcal{P}(U_1)$, whereas $\psi \mapsto r(x, \phi, \psi) + \mathcal{L}^{\phi, \psi} h(x)$ is l.s.c. on the compact set $\mathcal{P}(U_2)$. Therefore, the existence of a canonical pair (π^{*1}, π^{*2}) as in (3.5)–(3.7) can be easily obtained from standard measurable selection theorems —see, for instance [23, Theorem 12.1].

4 Bias Equilibria

The first refinement of Definition 3.4 and Theorem 3.5 is presented in this section. Here, we will note that the set of bias equilibria is a subset of that of average equilibria. However, this section can be regarded as a list of some results that we have obtained in past works (see, for instance [5] and [17]).

Definition 4.1 Let $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$. The *bias* of (π^1, π^2) is the function $h_{\pi^1, \pi^2} \in \mathcal{B}_w(\mathbb{R}^n)$ given by

$$h_{\pi^1, \pi^2}(x) := \int_0^\infty [\mathbb{E}_x^{\pi^1, \pi^2} r(x(t), \pi^1, \pi^2) - J(\pi^1, \pi^2)] dt \quad \text{for all } x \in \mathbb{R}^n. \tag{4.1}$$

Remark 4.2

- (i) The w -exponential ergodicity of the process $x^{\pi^1, \pi^2}(\cdot)$ (see (2.8)) and the Assumption 2.9(b) ensure that the bias h_{π^1, π^2} is a finite-valued function and, in fact, it is in $\mathcal{B}_w(\mathbb{R}^n)$. Moreover, its w -norm is uniformly bounded in $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$.
- (ii) By Escobedo-Trujillo et al. [5, Proposition 5.2] we can prove that if $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$ is average optimal, then its bias h_{π^1, π^2} and any function h satisfying the average optimality equations (3.5)–(3.7) coincide up to an additive constant; that is, for all $x \in \mathbb{R}^n$,

$$h_{\pi^1, \pi^2}(x) = h(x) - \mu_{\pi^1, \pi^2}(h).$$

Definition 4.3 (Bias Equilibrium) We say that an average saddle point $(\pi^{*1}, \pi^{*2}) \in (\Pi^1 \times \Pi^2)_{ao}$ is a *bias saddle point* if

$$h_{\pi^{*1}, \pi^{*2}}(x) \leq h_{\pi^{*1}, \pi^{*2}}(x) \leq h_{\pi^{*1}, \pi^{*2}}(x)$$

for every $x \in \mathbb{R}^n$ and every pair of strategies $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$. The function $h_{\pi^{*1}, \pi^{*2}}$ is called the *optimal bias function*.

We denote by $(\Pi^1 \times \Pi^2)_{bias}$ the set of bias saddle points. By Definition 4.3, $(\Pi^1 \times \Pi^2)_{bias} \subset (\Pi^1 \times \Pi^2)_{ao}$; that is,

$$\text{Bias equilibrium} \implies \text{Average equilibrium}.$$

Let (J, h) be a solution of the average payoff optimality equations (3.5)–(3.7). We define for each $x \in \mathbb{R}^n$ the sets

$$\Gamma_0^1(x) := \{\phi \in \mathcal{P}(U_1) \mid J = \inf_{\psi \in \mathcal{P}(U_2)} \{r(x, \phi, \psi) + \mathcal{L}^{\phi, \psi} h(x)\},$$

$$\Gamma_0^2(x) := \{\psi \in \mathcal{P}(U_2) \mid J = \sup_{\phi \in \mathcal{P}(U_1)} \{r(x, \phi, \psi) + \mathcal{L}^{\phi, \psi} h(x)\}.$$

Definition 4.4 We say that the constant $J \in \mathbb{R}$, the functions $h, \tilde{h} \in C^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$, and a pair $(\pi^{*1}, \pi^{*2}) \in \Pi^1 \times \Pi^2$ verify the bias optimality equations if and only if the triplet $(J, h, (\pi^{*1}, \pi^{*2}))$ satisfies the average optimality equations

(3.5)–(3.7) together with the following equations

$$h(x) = \mathcal{L}^{\pi^{*1}, \pi^{*2}} \tilde{h}(x) \quad (4.2)$$

$$= \sup_{\phi \in \Gamma_0^1(x)} \{\mathcal{L}^{\phi, \pi^{*2}} \tilde{h}(x)\} \quad (4.3)$$

$$= \inf_{\psi \in \Gamma_0^2(x)} \{\mathcal{L}^{\pi^{*1}, \psi} \tilde{h}(x)\}. \quad (4.4)$$

The next result summarizes important results on the existence of bias equilibria. For further details, see [5, Section 5] or [17, Theorem 7.7].

Proposition 4.5 *Under Assumptions 2.1, 2.7, 2.8, and 2.9, the following holds:*

- (i) $(\Pi^1 \times \Pi^2)_{bias}$ is nonempty.
- (ii) $\Gamma_0^1(x)$ and $\Gamma_0^2(x)$ are convex compact sets.
- (iii) The triplet $(J, h_{\pi^{*1}, \pi^{*2}}, \tilde{h})$ consisting of the constant J in Definition 3.4, the optimal bias function $h_{\pi^{*1}, \pi^{*2}}$ in Definition 4.3 and some other function $\tilde{h} \in C^2(\mathbb{R}^n \cap B_w(\mathbb{R}^n))$, form the unique solution satisfying the bias optimality equations (3.5)–(3.7) and (4.2)–(4.4).
- (iv) $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$ is a bias saddle point if and only if it verifies the bias optimality equations (4.2)–(4.4).

5 The Laurent Series

This section presents an extension of the results shown in [12, Section 3] or in [21, Section 3] to the zero-sum case. Here, we use the exponential ergodicity condition from Assumption 2.8 to characterize a discounted payoff in terms of a Laurent series. This will be very useful in our later developments. This is the essence of Theorem 5.5, which is the main result of this part.

Recall the definition of w in Assumption 2.7 and let μ_{π^1, π^2} be the invariant measure whose existence is ensured by Assumption 2.7.

Definition 5.1 Let $\mathcal{B}_w(\mathbb{R}^n \times U_1 \times U_2)$ be the space of measurable functions $v : \mathbb{R}^n \times U_1 \times U_2 \rightarrow \mathbb{R}$ such that

$$\sup_{(u_1, u_2) \in U_1 \times U_2} |v(x, u_1, u_2)| \leq M_v w(x) \quad \forall x \in \mathbb{R}^n, \quad (5.1)$$

where M_v is a positive constant depending of v .

As in (2.9) for $v \in \mathcal{B}_w(\mathbb{R}^n \times U_1 \times U_2)$ and $(\phi, \psi) \in \mathcal{P}(U_1) \times \mathcal{P}(U_2)$, we write

$$v(x, \phi, \psi) := \int_{U_2} \int_{U_1} v(x, u_1, u_2) \phi(du_1) \psi(du_2) \quad \forall x \in \mathbb{R}^n.$$

Now use $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$ in lieu of $(\phi, \psi) \in \mathcal{P}(U_1) \times \mathcal{P}(U_2)$. Let us define

$$\begin{aligned} \bar{v}(\pi^1, \pi^2) &:= \int_{\mathbb{R}^n} v(x, \pi^1, \pi^2) \mu_{\pi^1, \pi^2}(dx), \quad \text{and} \\ Z_t^{\pi^1, \pi^2} v(x) &:= \mathbb{E}_x^{\pi^1, \pi^2} v(x(t), \pi^1, \pi^2) - \bar{v}(\pi^1, \pi^2). \end{aligned}$$

With these ingredients, we define the v -bias operator $G_{\pi^1, \pi^2} : \mathcal{B}_w(\mathbb{R}^n \times U_1 \times U_2) \rightarrow \mathcal{B}_w(\mathbb{R}^n)$ as follows

$$G_{\pi^1, \pi^2} v(x) := \int_0^\infty [\mathbb{E}_x^{\pi^1, \pi^2} v(x(t), \pi^1, \pi^2) - \bar{v}(\pi^1, \pi^2)] dt. \tag{5.2}$$

Remark 5.2 Note that the w -exponential ergodicity of the process $x^{\pi^1, \pi^2}(\cdot)$ established in (2.8), and (5.1) yield that

$$|Z_t^{\pi^1, \pi^2} v(x)| \leq CM_v e^{-\delta t} w(x),$$

and thus,

$$|G_{\pi^1, \pi^2} v(x)| \leq \delta^{-1} CM_v w(x) \quad \text{or equivalently} \quad \|G_{\pi^1, \pi^2} v\|_w \leq \delta^{-1} CM_v. \tag{5.3}$$

The following result shows some properties of both, the operator G_{π^1, π^2} , and the operators that result from its compositions with itself. Its proof delves into the discussion that led from (3.10) to (3.11) in [12].

Lemma 5.3 For $j \geq 0$, let G_{π^1, π^2}^{j+1} be the $j + 1$ -composition of G_{π^1, π^2} with itself. Then

$$G_{\pi^1, \pi^2}^{j+1} v \text{ is in } \mathcal{B}_w(\mathbb{R}^n), \quad \text{and} \quad \mu_{\pi^1, \pi^2} \left(G_{\pi^1, \pi^2}^{j+1} v \right) = 0.$$

Proof By (5.3), G_{π^1, π^2} is in $B_w(\mathbb{R}^n)$. Now, the fact that $\mu_{\pi^1, \pi^2}(G_{\pi^1, \pi^2} v) = 0$ is straightforward from (3.2) and (5.2). The rest of the proof easily follows by applying mathematical induction on j . \square

Definition 5.4 Given a discount factor $\alpha > 0$. The expected α -discounted v -payoff when the players use $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$, given the initial state $x \in \mathbb{R}^n$, is

$$V_\alpha(x, \pi^1, \pi^2, v) := \mathbb{E}_x^{\pi^1, \pi^2} \left[\int_0^\infty e^{-\alpha t} v(x(t), \pi^1, \pi^2) dt \right]. \tag{5.4}$$

The following result provides a useful characterization of the α -discounted v -payoff in terms of a Laurent series (see, for instance [4, Chapter 6]). The proof uses essentially the same steps of the proof of Theorem 3.1 and Proposition 3.2 in [12], so we shall omit it.

Theorem 5.5

(a) Let $\delta > 0$ be the constant in Assumption 2.8. If (π^1, π^2) an arbitrary pair of strategies in $\Pi^1 \times \Pi^2$ and v is a function in $\mathcal{B}_w(\mathbb{R}^n \times U_1 \times U_2)$, then, for $\alpha \in (0, \delta)$, the α -discounted v -payoff (5.4) can be written as

$$V_\alpha(x, \pi^1, \pi^2, v) = \frac{1}{\alpha} \bar{v}(\pi^1, \pi^2) + \sum_{j=0}^{\infty} (-\alpha)^j G_{\pi^1, \pi^2}^{j+1} v(x). \tag{5.5}$$

Moreover the above series converges in w -norm.

(b) Let $\theta \in \mathbb{R}$ be such that $0 < \theta < \delta$, where δ is the constant in Assumption 2.8. For each $v \in \mathcal{B}_w(\mathbb{R}^n \times U_1 \times U_2)$, $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$, and $i = 0, 1, \dots$ define the i -residual of the Laurent series (5.5) as

$$R_i(\pi^1, \pi^2, v, \alpha) := \sum_{j=i}^{\infty} (-\alpha)^j G_{\pi^1, \pi^2}^{j+1} v.$$

Then, for all $|\alpha| \leq \theta$ and $i = 0, 1, \dots$,

$$\sup_{(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2} \left\| R_i((\pi^1, \pi^2), v, \alpha) \right\|_w \leq \frac{CM_v}{\delta^i(\delta - \theta)} |\alpha|^k. \tag{5.6}$$

For each $v \in \mathcal{B}_w(\mathbb{R}^n \times U_1 \times U_2)$, (π^1, π^2) in $\Pi^1 \times \Pi^2$, and $i = 0, 1, \dots$, define $h_{\pi^1, \pi^2}^i v$ as

$$h_{\pi^1, \pi^2}^i v(x) := (-1)^i G_{\pi^1, \pi^2}^{i+1} v(x) \quad \text{for all } x \in \mathbb{R}^n \text{ and } i = 1, 2. \tag{5.7}$$

It is obvious that, for each $v \in \mathcal{B}_w(\mathbb{R}^n \times U_1 \times U_2)$, $h_{\pi^1, \pi^2}^i v$ belongs to $B_w(\mathbb{R}^n)$ because $G_{\pi^1, \pi^2}^{i+1} v$ does.

Notation For $v = r$, with r as in Assumption 2.9, we simply write the operator in (5.7) as $h_{\pi^1, \pi^2}^i r$; that is,

$$h_{\pi^1, \pi^2}^i r := h_{\pi^1, \pi^2}^i v.$$

Note that for $i = 0$, h_{π^1, π^2}^0 equals to the bias function defined in (4.1), i.e.,

$$h_{\pi^1, \pi^2}^0(x) = G_{\pi^1, \pi^2} r(x) = h_{\pi^1, \pi^2}(x) \quad \text{for all } x \in \mathbb{R}^n.$$

Moreover,

$$h^1_{\pi^1, \pi^2} = -G^2_{\pi^1, \pi^2} r(x) = G_{\pi^1, \pi^2}(-h^0_{\pi^1, \pi^2}),$$

is the bias of (π^1, π^2) when the payoff is $-h^0_{\pi^1, \pi^2}$. In general, using mathematical induction, we can obtain that

$$h^i_{\pi^1, \pi^2} = G_{\pi^1, \pi^2}(-h^{i-1}_{\pi^1, \pi^2}) \quad i = 1, 2, \dots$$

By Theorem 5.5(a) and the expression (5.2), the α -discounted payoff (5.4)—with r in lieu v —can be written in terms of operator $h^i_{\pi^1, \pi^2}$ as follows

$$V_\alpha(x, \pi^1, \pi^2, r) = \frac{1}{\alpha} J(\pi^1, \pi^2) + \sum_{i=0}^\infty \alpha^i h^i_{\pi^1, \pi^2}(x), \tag{5.8}$$

and, by Lemma 5.3,

$$\mu_{\pi^1, \pi^2}(h^i_{\pi^1, \pi^2}) = 0 \quad \text{for all } i = 0, 1, 2, \dots \tag{5.9}$$

6 The Poisson System

We now define the so-called *Poisson system* and characterize its solution in terms of the basic average optimality criterion, and the recursive operator G_{π^1, π^2} introduced in Sect. 5.

For the following definition, recall that Eq. (3.5) is sometimes dubbed *Poisson equation*.

Definition 6.1 Let $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$ fixed. We say that a constant $J \in \mathbb{R}$ and the functions $h^0, h^1, \dots, h^{m+1} \in C^2(\mathbb{R}^n) \cap B_w(\mathbb{R}^n)$ verify the *Poisson system* for $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$ if

$$J = r(x, \pi^1, \pi^2) + \mathcal{L}^{\pi^1, \pi^2} h^0(x), \tag{6.1}$$

$$h^0(x) = \mathcal{L}^{\pi^1, \pi^2} h^1(x), \tag{6.2}$$

...

$$h^m(x) = \mathcal{L}^{\pi^1, \pi^2} h^{m+1}(x). \tag{6.3}$$

Theorem 6.2 Let $m \geq -1$ be fixed. The constant $J \in \mathbb{R}$ and the functions $h^0, h^1, \dots, h^{m+1} \in C^2(\mathbb{R}^n) \cap B_w(\mathbb{R}^n)$ are solutions to the Poisson system (6.1)–(6.3) if and only if $J = J(\pi^1, \pi^2)$, $h^i = h^i_{\pi^1, \pi^2}$ for $0 \leq i \leq m$, and $h^{m+1} =$

$h_{\pi^1, \pi^2}^{m+1} + z$ for $z \in \mathbb{R}$, where J and h_{π^1, π^2}^i , $0 \leq i \leq m + 1$, are the functions in (3.2) and (5.7), respectively.

Proof We will use mathematical induction over Eqs. (6.1)–(6.3).

1. Case $m = -1$ follows from Lemma 3.2 and Proposition 5.1 in [5].
2. Now, suppose the result is valid for some $m \geq -1$.
3. Case $m + 1$:

The “if” part: Suppose that $J = J(\pi^1, \pi^2)$, $h^i = h_{\pi^1, \pi^2}^i$ for $0 \leq i \leq m$, and $h^{m+1} = h_{\pi^1, \pi^2}^{m+1} + z$ for $z \in \mathbb{R}$. Then, we need to prove that h_{π^1, π^2}^{m+1} verifies the $(m + 1)$ -th Poisson equation. To this end, observe that h_{π^1, π^2}^{m+2} is the bias function of (π^1, π^2) when we consider as reward rate $-h_{\pi^1, \pi^2}^{m+1}(x)$. It is easy to verify through a mathematical induction procedure that $-h_{\pi^1, \pi^2}^{m+1}$ satisfies Assumption 2.9, then we can invoke Theorem 4.1 in [5], to ensure the existence of a function $h^{m+2} \in C^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$, a constant \bar{J} and a pair of strategies (π^1, π^2) that satisfy the average optimality equation

$$\begin{aligned} \bar{J} &= -h^{m+1}(x) + \mathcal{L}^{\pi^1, \pi^2} h^{m+2}(x) \\ &= \sup_{\phi \in \mathcal{P}(U_1)} \{-h^{m+1}(x) + \mathcal{L}^{\phi, \pi^2} h^{m+2}(x)\}, \\ &= \inf_{\psi \in \mathcal{P}(U_2)} \{-h^{m+1}(x) + \mathcal{L}^{\pi^1, \psi} h^{m+2}(x)\}, \end{aligned}$$

with $\bar{J} = \mu_{\pi^1, \pi^2}(-h^{m+1}) = \mu_{\pi^1, \pi^2}(-h_{\pi^1, \pi^2}^{m+1})$. Now, Proposition 5.1 in [5] gives that the bias function with reward rate $-h^{m+1}(x) = -h_{\pi^1, \pi^2}^{m+1}(x)$ satisfies the following Poisson equation

$$\mu_{\pi^1, \pi^2}(-h_{\pi^1, \pi^2}^{m+1}) = -h_{\pi^1, \pi^2}^{m+1}(x) + \mathcal{L}^{\pi^1, \pi^2} h^{m+2}(x),$$

which implies that

$$h_{\pi^1, \pi^2}^{m+1}(x) = \mathcal{L}^{\pi^1, \pi^2} h^{m+2}(x), \tag{6.4}$$

since that (5.9) gives $\mu_{\pi^1, \pi^2}(-h_{\pi^1, \pi^2}^{m+1}) = 0$. Thus, (6.4) implies that h_{π^1, π^2}^{m+1} satisfies the $(m + 1)$ -th Poisson equation.

The “only if” part: Suppose that $J \in \mathbb{R}$ and $h^0, h^1, \dots, h^{m+1} \in C^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$ are solutions to (6.1)–(6.3). By the induction hypothesis the result holds for some $m \geq 0$, i.e.,

$$h_{\pi^1, \pi^2}^m(x) = h^m(x) = \mathcal{L}^{\pi^1, \pi^2} h^{m+1}(x). \tag{6.5}$$

Therefore, we only need to prove that $h^{m+1} = h_{\pi^1, \pi^2}^{m+1}$. Namely, the bias function $h_{\pi^1, \pi^2}^{m+1}(x)$ when the payoff rate is $-h_{\pi^1, \pi^2}^m$, verifies the following Poisson equation

$$\mu_{\pi^1, \pi^2}(-h_{\pi^1, \pi^2}^m) = -h_{\pi^1, \pi^2}^m(x) + \mathcal{L}^{\pi^1, \pi^2} h_{\pi^1, \pi^2}^{m+1}(x),$$

then, by (5.9) we obtain

$$h_{\pi^1, \pi^2}^m(x) = \mathcal{L}^{\pi^1, \pi^2} h_{\pi^1, \pi^2}^{m+1}(x). \tag{6.6}$$

Thus, subtracting equation (6.5) to (6.6) we obtain

$$0 = \mathcal{L}^{\pi^1, \pi^2} (h_{\pi^1, \pi^2}^{m+1}(x) - h^{m+1}(x)).$$

Therefore, $h_{\pi^1, \pi^2}^{m+1} - h^{m+1}$ is a harmonic function and as a consequence, Lemma 2.1 in [5], yields

$$h_{\pi^1, \pi^2}^{m+1}(x) = h^{m+1}(x) + \mu_{\pi^1, \pi^2}(h^{m+1}). \tag{6.7}$$

Since μ_{π^1, π^2} is an invariant probability measure, and $h^{m+1} \in C^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$ satisfies the $(m + 1)$ -th Poisson equation, we have

$$\begin{aligned} \mu_{\pi^1, \pi^2}(h^{m+1}) &= \int_{\mathbb{R}^n} \mathcal{L}^{\pi^1, \pi^2} h^{m+2}(y) \mu_{\pi^1, \pi^2}(dy) = 0 \quad \text{for all} \\ h^{m+2} &\in C^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n), \end{aligned} \tag{6.8}$$

where the last equality follows from a well-known result of invariant probability measures—see, for example [2]. Therefore, $h^{m+1} = h_{\pi^1, \pi^2}^{m+1}$ follows from (6.7) and (6.8).

□

7 The Average Payoff Optimality System

We devote this section to link the Poisson system (6.1)–(6.3) from Sect. 6 with the optimization problem we are trying to solve (see Definitions 8.1 and 8.2 below). We do this by means of a system of average optimality equations, and the characterization of their solutions as a sequence of *canonical equilibria* of a collection of average payoff games. This is the purpose of the main result of this part, namely, Theorem 7.4.

Definition 7.1 We say that a constant $J \in \mathbb{R}$ and functions $h^0, h^1, \dots, h^{m+1} \in C^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$ verify the -1 -th, 0 -th, \dots , m -th average payoff optimality system for $(\pi^{*1}, \pi^{*2}) \in \Pi^1 \times \Pi^2$ and $x \in \mathbb{R}^n$ if

$$J = r(x, \pi^{*1}, \pi^{*2}) + \mathcal{L}^{\pi^{*1}, \pi^{*2}} h^0(x), \quad (7.1)$$

$$= \sup_{\phi \in \mathcal{P}(U_1)} r(x, \phi, \pi^{*2}) + \mathcal{L}^{\phi, \pi^{*2}} h^0(x), \quad (7.2)$$

$$= \inf_{\psi \in \mathcal{P}(U_2)} r(x, \pi^{*1}, \psi) + \mathcal{L}^{\pi^{*1}, \psi} h^0(x) \quad (7.3)$$

$$h^0(x) = \mathcal{L}^{\pi^{*1}, \pi^{*2}} h^1(x) \quad (7.4)$$

$$= \sup_{\phi \in \Gamma_0^1(x)} \mathcal{L}^{\phi, \pi^{*2}} h^1(x) \quad (7.5)$$

$$= \inf_{\psi \in \Gamma_0^2(x)} \mathcal{L}^{\pi^{*1}, \psi} h^1(x) \quad (7.6)$$

...

$$h^m(x) = \mathcal{L}^{\pi^{*1}, \pi^{*2}} h^{m+1}(x) \quad (7.7)$$

$$= \sup_{\phi \in \Gamma_m^1(x)} \mathcal{L}^{\phi, \pi^{*2}} h^{m+1}(x) \quad (7.8)$$

$$= \inf_{\psi \in \Gamma_m^2(x)} \mathcal{L}^{\pi^{*1}, \psi} h^{m+1}(x) \quad (7.9)$$

where letting $\Gamma_{-1}^1(x) := \mathcal{P}(U_1)$ and $\Gamma_{-1}^2(x) := \mathcal{P}(U_2)$ for all $x \in \mathbb{R}^n$, then the sets $\Gamma_j^k(x)$, for $0 \leq j \leq m$ and $k = 1, 2$, consist of probability measures $\phi \in \Gamma_{j-1}^1(x)$ and $\psi \in \Gamma_{j-1}^2(x)$ attaining the maximum and minimum in the $(j-1)$ -th average payoff optimality equation, respectively; that is, for each $x \in \mathbb{R}^n$,

$$\Gamma_0^1(x) := \left\{ \phi \in \mathcal{P}(U_1) \mid J = \inf_{\psi \in \mathcal{P}(U_2)} \left[r(x, \phi, \psi) + \mathcal{L}^{\phi, \psi} h^0(x) \right] \right\},$$

$$\Gamma_0^2(x) := \left\{ \psi \in \mathcal{P}(U_2) \mid J = \sup_{\phi \in \mathcal{P}(U_1)} \left[r(x, \phi, \psi) + \mathcal{L}^{\phi, \psi} h^0(x) \right] \right\}$$

and, for $1 \leq j \leq m$,

$$\Gamma_j^1(x) := \left\{ \phi \in \Gamma_{j-1}^1(x) \mid h^{j-1}(x) = \inf_{\psi \in \Gamma_{j-1}^2(x)} \mathcal{L}^{\phi, \psi} h^j(x) \right\},$$

$$\Gamma_j^2(x) := \left\{ \psi \in \Gamma_{j-1}^2(x) \mid h^{j-1}(x) = \sup_{\phi \in \Gamma_{j-1}^1(x)} \mathcal{L}^{\phi, \psi} h^j(x) \right\}.$$

Proposition 7.2 *For each $k = 1, 2$, and $-1 \leq j \leq m$, the sets $\{\Gamma_j^k(x)\}_{j \geq 0}$ are convex compact sets.*

Proof We use mathematical induction on j :

1. Case $j = -1, 0$. Since $\mathcal{P}(U_1)$ and $\mathcal{P}(U_2)$ are compact and convex sets (see, for instance, [1, Theorem 15.11]), Lemma 5.1 in [5], gives that $\Gamma_0^1(x)$ and $\Gamma_0^2(x)$ are also convex and compact sets.
2. Suppose now that for some $0 \leq j \leq m$, $\Gamma_j^1(x)$ and $\Gamma_j^2(x)$, are convex compact sets.
3. Let us prove the result for $m = j + 1$. To this end, note that

$$\Gamma_{j+1}^1(x) := \left\{ \phi \in \Gamma_j^1(x) \mid h^j(x) = \inf_{\psi \in \Gamma_j^2(x)} \mathcal{L}^{\phi, \psi} h^{j+1}(x) \right\},$$

and

$$\Gamma_{j+1}^2(x) := \left\{ \psi \in \Gamma_j^2(x) \mid h^j(x) = \sup_{\phi \in \Gamma_j^1(x)} \mathcal{L}^{\phi, \psi} h^{j+1}(x) \right\},$$

and by induction hypothesis $\Gamma_j^1(x)$ and $\Gamma_j^2(x)$ are convex compact sets. Then, to verify if $\Gamma_{j+1}^1(x)$ and $\Gamma_{j+1}^2(x)$ are compact sets it is sufficient to prove that they are closed, but this property follows due to the compactness of $\Gamma_j^1(x)$ and $\Gamma_j^2(x)$ (induction hypothesis) and the u.s.c in ϕ (l.s.c. ψ) of $\mathcal{L}^{\phi, \psi}$ established in the Remark 2.10.

The proof that $\Gamma_{j+1}^1(x)$ and $\Gamma_{j+1}^2(x)$ are convex sets mimicks that of Lemma 4.6 in [20].

□

Since $\{\Gamma_j^k(x)\}_{j \geq 0}$, $k = 1, 2$, is a nonincreasing sequence of nonempty compact sets, the set

$$\Gamma_\infty^k(x) := \bigcap_{m \geq -1} \Gamma_m^k(x) \tag{7.10}$$

is nonempty and compact as well.

The following definition concerns the pair of strategies $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$ that attain the maximum and minimum respectively in Eqs. (7.1)–(7.9).

Definition 7.3 We define

$$\Pi_m^1 \times \Pi_m^2 := \{(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2 \mid (\pi^1(\cdot|x), \pi^2(\cdot|x)) \in \Gamma_{m+1}^1(x) \times \Gamma_{m+1}^2(x), \forall x \in \mathbb{R}^n\}.$$

A pair $(\pi^1, \pi^2) \in \Pi_m^1 \times \Pi_m^2$ will be referred to as a *canonical equilibrium* for the -1 -th, 0 -th, \dots , m -th average payoff optimality system (7.1)–(7.9).

From Definition 7.3, it is clear that $\Pi_{m+1}^1 \times \Pi_{m+1}^2 \subseteq \Pi_m^1 \times \Pi_m^2$, for all $m = -1, 0, 1, \dots$.

Theorem 7.4 *The -1 -th, 0 -th, \dots , m -th average reward HJB system (7.1)–(7.9) admits a unique solution $J \in \mathbb{R}$, $h^0, h^1, \dots, h^{m+1} \in C^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$, where J, h^0, h^1, \dots, h^m are unique, and h^{m+1} is unique up to an additive constant. Moreover, the set $\Pi_m^1 \times \Pi_m^2$ is nonempty.*

Proof We will use mathematical induction on m .

1. Case $m = 0$. It follows from Theorems 4.1, 5.1 and 5.2 in [5].
2. Suppose that the result holds for some $m = j$.
3. Now, we prove that the result holds for $m = j + 1$.

The induction hypothesis ensures the existence of $J \in \mathbb{R}$, $h^0, h^1, \dots, h^j \in C^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$ which are unique solutions of the -1 -th, 0 -th, \dots , j -th average payoff optimality system and that that Π_j^1 and Π_j^2 are nonempty.

Let us consider now a new game, so-named j -bias game, consisting in:

- The dynamic system (2.1).
- The payoff function $-h^j$. (7.11)
- The set of control actions $\Gamma_j^1(x)$ and $\Gamma_j^2(x)$.

It is easy to verify that this new game satisfies all of our hypotheses. Then, Theorem 3.5(i)–(ii) ensures the existence of solutions $(\bar{J}, h^{j+1}, (\pi^{*1}, \pi^{*2}))$ to the following average optimality equations

$$\begin{aligned} \bar{J} &= -h^j(x) + \mathcal{L}^{\pi^{*1}, \pi^{*2}} h^{j+1}(x) \\ &= \sup_{\phi \in \Gamma_j^1(x)} \{-h^j(x) + \mathcal{L}^{\phi, \pi^{*2}} h^{j+1}(x)\} \\ &= \inf_{\psi \in \Gamma_j^2(x)} \{-h^j(x) + \mathcal{L}^{\pi^{*1}, \psi} h^{j+1}(x)\}. \end{aligned}$$

The existence of a function $h^{j+2} \in C^2(\mathbb{R}^n) \cap \mathcal{B}_w(\mathbb{R}^n)$ satisfying

$$\begin{aligned} h^{j+1}(x) &= \mathcal{L}^{\pi^{*1}, \pi^{*2}} h^{j+2}(x) \\ &= \sup_{\phi \in \Gamma_j^1(x)} \mathcal{L}^{\phi, \pi^{*2}} h^{j+2}(x) \\ &= \inf_{\psi \in \Gamma_j^2(x)} \mathcal{L}^{\pi^{*1}, \psi} h^{j+2}(x). \end{aligned}$$

is ensured by Proposition 4.5, and the fact that $\bar{J} = \mu_{\pi^{*1}, \pi^{*2}}(-h^j)$. In this case h^{j+1} is unique, and h^{j+2} is unique up to additive constants. Thus, h^{j+1} satisfies the $(j + 1)$ -th average reward HJB equations.

It remains to prove that $\Pi_m^1 \times \Pi_m^2$ is nonempty. To this end, we proceed again by mathematical induction on m . Namely, for the case $m = 0$, the result follows by Theorems 5.1 and 5.2 in [5]. Now assume that $\Pi_j^1 \times \Pi_j^2$ is nonempty for some $j = 0, 1, \dots$; that is, there is at least an element $(\pi_j^1, \pi_j^2) \in \Pi_j^1 \times \Pi_j^2$ or equivalently, $(\pi_{j+1}^1(\cdot|x), \pi_{j+1}^2(\cdot|x)) \in \Gamma_{j+1}^1(x) \times \Gamma_{j+1}^2(x)$ for all $x \in \mathbb{R}^n$. We want to prove that $\Pi_{j+1}^1 \times \Pi_{j+1}^2$ is nonempty. For this, we consider again the j -bias game (7.11). Since this game satisfies all of our hypotheses, we can invoke Proposition 4.5(i) to ensure the existence of a bias equilibrium (π^1, π^2) associated to the j -bias game. Hence, Proposition 4.5(iii) yields that, in fact, such equilibrium satisfies both, the j -th and the $(j + 1)$ -th average payoff equations. This completes the proof. \square

Remark 7.5 It is worth noting the relation of the pairs $(\pi^{*1}, \pi^{*2}) \in \Pi_m^1 \times \Pi_m^2$ with the m -bias game in (7.11); namely, if we apply iteratively Proposition 4.5, we can easily verify that $(\pi^{*1}, \pi^{*2}) \in \Pi_m^1 \times \Pi_m^2$ if and only if such a pair is an average Nash equilibrium for the j -bias game (7.11) for $j = -1, \dots, m$.

We define

$$\Pi_\infty^1 \times \Pi_\infty^2 := \bigcap_{m=-1}^\infty (\Pi_m^1 \times \Pi_m^2). \tag{7.12}$$

As a consequence of (7.10) and Theorem 7.4, we deduce the following result.

Corollary 7.6 *There exists a strategy $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$ that satisfies the m -th average reward HJB equation for all $m = -1, 0, \dots$. In other words, $\Pi_\infty^1 \times \Pi_\infty^2$ is nonempty.*

8 Blackwell-Nash Equilibria

In this section we present a zero-sum type of Nash equilibrium so-named *Blackwell-Nash equilibrium*; we will also introduce a sensitive discount concept related to a family of optimality criteria so-named *m-discount equilibria*, for $m \geq -1$. We will see that a Blackwell-Nash equilibrium becomes the limit, as $m \rightarrow \infty$, of a sequence of *m-discount equilibria* and prove the existence of each element of this sequence based on the results given in previous sections. To begin with this analysis, we first define the aforementioned concepts as follows.

Definition 8.1 (Blackwell-Nash Equilibrium) A pair $(\pi^{*1}, \pi^{*2}) \in \Pi^1 \times \Pi^2$ is called Blackwell-Nash equilibrium if for each $(\pi^1, \pi^2) \in \Pi^1 \times \Pi^2$ and each state $x \in \mathbb{R}^n$, there exists a discount factor $\alpha^* = \alpha^*(x, \pi^1, \pi^2)$ such that

$$V_\alpha(x, \pi^1, \pi^{*2}) \leq V_\alpha(x, \pi^{*1}, \pi^{*2}) \leq V_\alpha(x, \pi^{*1}, \pi^2) \quad (8.1)$$

for all $0 < \alpha < \alpha^*$.

Definition 8.2 (Sensitive Discount Equilibrium)

(a) Let $m \geq -1$ be an integer. A pair $(\pi^{*1}, \pi^{*2}) \in \Pi^1 \times \Pi^2$ is called an *m-discount equilibrium* if

$$\liminf_{\alpha \rightarrow 0} \alpha^{-m} [V_\alpha(x, \pi^{*1}, \pi^{*2}) - V_\alpha(x, \pi^1, \pi^{*2})] \geq 0 \quad \text{for all } \pi^1 \in \Pi^1,$$

and

$$\limsup_{\alpha \rightarrow 0} \alpha^{-m} [V_\alpha(x, \pi^{*1}, \pi^{*2}) - V_\alpha(x, \pi^{*1}, \pi^2)] \leq 0 \quad \text{for all } \pi^2 \in \Pi^2.$$

(b) We call *sensitive discount equilibria* to the family $\{(\pi_m^{*1}, \pi_m^{*2}) \mid m \geq -1\}$ of all the *m-discount equilibria* ($m \geq -1$).

We denote by $\Pi_m^{1,d}$ and $\Pi_m^{2,d}$ the sets of strategies *m-discount optimal* for player 1 and 2, respectively.

Theorem 8.3

- (i) Let $m \geq -1$ be an integer, then $\Pi_m^1 \times \Pi_m^2 \subseteq \Pi_m^{1,d} \times \Pi_m^{2,d}$.
(ii) If $(\pi^{*1}, \pi^{*2}) \in \Pi_\infty^1 \times \Pi_\infty^2$, then it is a *Blackwell-Nash equilibrium*.

Proof

- (i) Consider the pair $(\pi^{*1}, \pi^{*2}) \in \Pi_m^1 \times \Pi_m^2$, and use the series (5.8) to deduce the following

$$\begin{aligned} & \frac{1}{\alpha^m} [V_\alpha(x, \pi^{*1}, \pi^{*2}) - V_\alpha(x, \pi^1, \pi^{*2})] \\ &= \frac{1}{\alpha} \left[\frac{1}{\alpha^m} \left(J(\pi^{*1}, \pi^{*2}) - J(\pi^1, \pi^{*2}) \right) + \frac{1}{\alpha^{m-1}} \left(h_{\pi^{*1}, \pi^{*2}}^0(x) - h_{\pi^1, \pi^{*2}}^0(x) \right) \right. \\ &+ \dots + \left. \left(h_{\pi^{*1}, \pi^{*2}}^{m-1}(x) - h_{\pi^1, \pi^{*2}}^{m-1}(x) \right) \right] + \left(h_{\pi^{*1}, \pi^{*2}}^m(x) - h_{\pi^1, \pi^{*2}}^m(x) \right) + \\ &+ \frac{1}{\alpha^m} \sum_{i=m+1}^{\infty} \alpha^i \left(h_{\pi^{*1}, \pi^{*2}}^i(x) - h_{\pi^1, \pi^{*2}}^i(x) \right), \end{aligned} \tag{8.2}$$

for all $\pi^1 \in \Pi^1$. By virtue of Remark 7.5, (π^{*1}, π^{*2}) is a Nash equilibrium for the -1 -th, 0 -th, \dots , m -th bias game (7.11). Then, the first $m + 2$ elements in equality (8.2) are greater or equal to zero. Finally, letting $\alpha \rightarrow 0$ in both sides of (8.2) and using Theorem 5.5(b), we get

$$\frac{1}{\alpha^m} [V_\alpha(x, \pi^{*1}, \pi^{*2}) - V_\alpha(x, \pi^1, \pi^{*2})] \geq 0.$$

Similar arguments yield

$$\frac{1}{\alpha^m} [V_\alpha(x, \pi^{*1}, \pi^{*2}) - V_\alpha(x, \pi^{*1}, \pi^2)] \leq 0 \text{ for all } \pi^2 \in \Pi^2.$$

Therefore, $\Pi_m^1 \times \Pi_m^2 \subset \Pi_m^{1,d} \times \Pi_m^{2,d}$, which proves (i).

- (ii) Let $\pi^1 \in \Pi^1$ and $x \in \mathbb{R}^n$ arbitrary and suppose that $(\pi^{*1}, \pi^{*2}) \in \Pi_\infty^1 \times \Pi_\infty^2$, then using again (5.8) we can write

$$\begin{aligned} V_\alpha(x, \pi^{*1}, \pi^{*2}) - V_\alpha(x, \pi^1, \pi^{*2}) &= \frac{1}{\alpha} [J(\pi^{*1}, \pi^{*2}) - J(\pi^1, \pi^{*2})] \\ &+ \sum_{i=0}^{\infty} \alpha^i [h_{\pi^{*1}, \pi^{*2}}^i(x) - h_{\pi^1, \pi^{*2}}^i(x)]. \end{aligned} \tag{8.3}$$

By virtue of (7.12), $(\pi^{*1}, \pi^{*2}) \in \Pi_m^1 \times \Pi_m^2$ for $-1 \leq m \leq \infty$. So, (π^{*1}, π^{*2}) is a Nash equilibrium for the m -bias game (7.11) for all $m = -1, 0, 1, \dots$. Therefore, the equality in (8.3) is nonnegative for every $\alpha > \alpha^*$, where α^* depends on the residual term (5.6), which yields the first inequality in (8.1). We can also mimic the same arguments but now for arbitrary $\pi^2 \in \Pi^2$ and thus to obtain the second inequality in (8.1), yielding that (π^{*1}, π^{*2}) is a Blackwell-Nash equilibrium.

□

We use Theorems 7.4, 8.3, and Corollary 7.6, to state our final claim.

Corollary 8.4 *Under Assumptions 2.1, 2.7, 2.8, and 2.9,*

- (i) *For each $m \geq -1$, the set $\Pi_m^{1,d} \times \Pi_m^{2,d}$ of m -discount optimal strategies is nonempty.*
- (ii) *There exist Blackwell optimal strategies in $\Pi^1 \times \Pi^2$.*

9 Final Remarks

In this paper we have shown the existence and provide some characterizations of the sensitive discount equilibria in a class of zero-sum stochastic differential games with a uniform ellipticity assumption. This yields a Blackwell-Nash equilibrium in the limit as $m \rightarrow \infty$. To this end, we truncated the Laurent series of the expected discounted reward/cost, and thus stated the so-called *Poisson system*, which allowed us to characterize the equilibria as the collection of strategies that meet it.

It is worth pointing out the fact that Theorem 8.3 and Corollary 8.4 show that, for a zero-sum stochastic differential game, an m -discount equilibrium is equivalent to a Blackwell-Nash equilibrium only when $m \rightarrow \infty$. This agrees with the controlled diffusion scheme (see [12, 22]).

Some possible extensions of our work are, for example, to do this same analysis but considering a more general dynamics type, such is the case of stochastic differential equations with jumps (in the context of Lévy processes) or using the same dynamic than ours but under weaker assumptions than those considered here, such is the case of *degenerate* diffusions.

Acknowledgement This research was supported in part by CONACyT grant 238045.

References

1. C. Aliprantis, K. Border, *Infinite Dimensional Analysis*, 3rd edn. (Springer, Berlin, 2006)
2. A. Arapostathis, M. Ghosh, V. Borkar, *Ergodic Control of Diffusion Processes*, vol. 143. Encyclopedia of Mathematics and Its Applications (Cambridge University Press, Cambridge, 2012)
3. A. Arapostathis, V. Borkar, K. Suresh, Relative value iteration for stochastic differential games, in *Advances in Dynamic Games: Theory, Applications, and Numerical Methods* (Springer International Publishing, Cham, 2013), pp. 3–27
4. J.B. Conway, *A Course in Functional Analysis* (Springer, New York, 1994)
5. B.A. Escobedo-Trujillo, J.D. López-Barrientos, O. Hernández-Lerma, Bias and overtaking equilibria for zero-sum stochastic differential games. *J. Optim. Theory Appl.* **153**, 662–687 (2012)
6. M. Ghosh, A. Arapostathis, S. Marcus, Optimal control of switching diffusions with applications to flexible manufacturing systems. *SIAM J. Control Optim.* **30**, 1–23 (1992)

7. H. Gimbert, W. Zielonka, Deterministic priority mean-payoff games as limits of discounted games, in *Automata, Languages and Programming*, Part II. Lecture Notes in Computer Science, vol. 4052 (Springer, Berlin, 2006), pp. 312–323
8. H. Gimbert, W. Zielonka, Applying Blackwell optimality: priority mean-payoff games as limits of multi-discounted games, in *Logic and Automata*. Texts in Logic and Games, vol. 2 (Amsterdam University Press, Amsterdam, 2008), pp. 331–355
9. H. Gimbert, W. Zielonka, Blackwell optimal strategies in priority mean-payoff games. *Int. J. Found. Comput. Sci.* **23**, 687–711 (2012)
10. H. Jasso-Fuentes, Noncooperative continuous-time Markov games. *Morfismos* **9**, 39–54 (2005)
11. H. Jasso-Fuentes, O. Hernández-Lerma, Characterizations of overtaking optimality for controlled diffusion processes. *Appl. Math. Optim.* **57**, 349–369 (2008)
12. H. Jasso-Fuentes, O. Hernández-Lerma, Blackwell optimality for controlled diffusion processes. *J. Appl. Probab.* **46**, 372–391 (2009)
13. H. Jasso-Fuentes, O. Hernández-Lerma, Ergodic control, bias and sensitive discount optimality for Markov diffusion processes. *Stoch. Anal. Appl.* **27**, 363–385 (2009)
14. H. Jasso-Fuentes, J.D. López-Barrientos, B.A. Escobedo-Trujillo, Infinite horizon nonzero-sum stochastic differential games with additive structure. *IMA J. Math. Control Inform.* **34**, 283–309 (2017)
15. A. Leizarowitz, Controlled diffusion process on infinite horizon with the overtaking criterion. *Appl. Math. Optim.* **17**, 61–78 (1988)
16. A. Leizarowitz, Optimal controls for diffusion in \mathbb{R}^d -min-max max-min formula for the minimal cost growth rate. *J. Math. Anal. Appl.* **149**, 180–209 (1990)
17. J.D. López-Barrientos, Basic and advanced optimality criteria for zero-sum stochastic differential games. Ph.D. dissertation. Department of Mathematics, CINVESTAV-IPN, 2012. Available at <http://www.math.cinvestav.mx/sites/default/files/tesis-daniel-2012.pdf>
18. S. Meyn, R. Tweedie, Stability of Markovian processes III. Foster-Lyapunov criteria for continuous-time processes. *Adv. Appl. Probab.* **25**, 518–548 (1993)
19. A.S. Nowak, Correlated relaxed equilibria in nonzero-sum linear differential games. *J. Math. Anal. Appl.* **163**, 104–112 (1992)
20. T. Prieto-Rumeau, O. Hernández-Lerma, Bias and overtaking equilibria for zero-sum continuous-time Markov games. *Math. Meth. Oper. Res.* **61**, 437–454 (2005)
21. T. Prieto-Rumeau, O. Hernández-Lerma, The Laurent series, sensitive discount and Blackwell optimality for continuous-time controlled Markov chains. *Math. Meth. Oper. Res.* **61**, 123–145 (2005)
22. M.L. Puterman, Sensitive discount optimality in controlled one-dimensional diffusions. *Ann. Probab.* **2**, 408–419 (1974)
23. M. Schäl, Conditions for optimality and for the limit of n -stage optimal policies to be optimal. *Z. Wahrs. Verw. Gerb.* **32**, 179–196 (1975)
24. J. Warga, *Optimal Control of Differential and Functional Equations* (Academic, New York, 1972)
25. P. Whittle, *Risk-Sensitive Optimal Control*. Wiley-Interscience Series in Systems and Optimization (Wiley, Chichester, 1990)
26. L.C. Young, Generalized curves and the existence of an attained absolute minimum in the calculus of variations. *Comptes Rendus de la Société des Sciences et des Lettres de Varsovie, Classe III* **30**, 212–234 (1937)