ZERO-SUM GAMES FOR PURE JUMP PROCESSES WITH RISK-SENSITIVE DISCOUNTED COST CRITERIA

CHANDAN PAL

Department of Mathematics
Indian Institute of Technology Guwahati
Guwahati, Assam - 781039, India

SOMNATH PRADESHAN

Department of Mathematics
Indian Institute of Science Education and Research Pune
Pune, Maharashtra - 411008, India

(Communicated by Onésimo Hernández-Lerma)

SUMMARY. In this paper we study zero-sum stochastic games for pure jump processes on a general state space with risk sensitive discounted criteria. We establish a saddle point equilibrium in Markov strategies for bounded cost function. We achieve our results by studying relevant Hamilton-Jacobi-Isaacs equations.

1. Introduction. This article is a sequel to [14] where risk-sensitive control problem for continuous time Markov decision processes (MDPs) is studied on a Borel state space. In this article, we extend the result of [14] to risk sensitive zero sum stochastic games for continuous time controlled pure jump processes on a Borel state space. Stochastic games for continuous time Markov jump processes have been widely studied under different optimality criteria. The existing literatures on stochastic games for continuous time Markov jump processes can be roughly classified into two main groups. The first deals with risk-neutral stochastic games for continuous time Markov jump processes (see, for example, [9], [7], [8], [18], [20] etc); on the other with risk-sensitive stochastic games for continuous time Markov jump processes (see, for instance, [4], [17]). In risk-neutral stochastic games players ignore the risk since they usually consider the expectation of the integral of costs. In risk-sensitive stochastic games, the cost criterion is the expectation of the exponential of the integral of costs. Due to this, the analysis of the risk-sensitive case is technically more involved. We refer to [13] for an excellent note on risk-sensitive Nash-equilibria. To our knowledge, the risk-sensitive criterion was first introduced by Bellman [1]; see [19] and the references therein. For stochastic optimal control problems this criterion has been studied widely, see [5], [10], [11], [15], [16], [14], [21] and the references therein. The corresponding literature in the context of risk sensitive stochastic games is rather limited. Some exceptions are [4], [17].

2020 Mathematics Subject Classification. Primary: 91A15, 91A25; Secondary: 60J75.

Key words and phrases. Pure jump process, Markov strategy, value of the game, saddle point equilibrium, HJI equation.

* Corresponding author: Chandan Pal.
To prove the existence of a saddle-point equilibrium for a zero-sum risk sensitive game problem, the optimality equation (also known as the Shapley equation or the Hamilton-Jacobi-Isaacs (HJI) equation) plays a key role; see, [4]. Using HJI equation, one can completely characterize saddle point equilibrium in the space of Markov strategies in Section 3, and problems. The zero-sum game is analyzed in Section 3. We prove existence and give us the existence and characterization of a saddle point equilibrium.

The rest of this paper is organized as follows. In Section 2, we describe the problems. The zero-sum game is analyzed in Section 3. We prove existence and characterization of a saddle point equilibrium in Markov strategies in Section 3, and illustrated with one example in Section 4.

2. Problem description. The two parson zero-sum continuous-time Markov game model we consider here is given by

\[[S, U_1 \times U_2, q(\cdot|x, u_1, u_2), r(x, u_1, u_2)]\]

where each component is described below:

The state space S: The state space S is the set of states of the process under observation which is assumed to be a Borel space, i.e., a Borel subset of a Polish space (a complete and separable metric space with a metric d) with Borel σ-algebra \(\mathcal{B}(S)\).

The action space U: The players dynamically take their actions from the action space \(U_i, i = 1, 2\). We assume that \(U_i\) are given compact metric spaces and \(U = U_1 \times U_2\).

The transition rate \(q(\cdot|x, u_1, u_2)\), satisfies the following properties:

(A0)(i) For each fixed \((x, u_1, u_2) \in S \times U\), \(q(\cdot|x, u_1, u_2)\) is a signed measure on \(\mathcal{B}(S)\) and for each fixed \(A \in \mathcal{B}(S)\), \(q(A)\) is a Borel-measurable map on \(S \times U\).

(ii) \(0 \leq q(A|x, u_1, u_2) < \infty\) for all \((x, u_1, u_2) \in S \times U\), if \(x \notin A \in \mathcal{B}(S)\).

(iii) \(q(S|x, u_1, u_2) = 0\) for all \((x, u_1, u_2) \in S \times U\). (Hence \(0 \leq -q(\{x\}|x, u_1, u_2) = q(S \setminus \{x\}|x, u_1, u_2) < \infty\) for all \((x, u_1, u_2) \in S \times U)\)

(A1)(i) We assume that our model is stable, i.e.,

\[
\sup_{x \in S, (u_1, u_2) \in U} [-q(\{x\}|x, u_1, u_2)] := M < \infty.
\]

(ii) For each fixed \(x \in S\), the functions \(\int_S f(y)q(dy|x, u_1, u_2)\) are continuous in \(u_1, u_2\) for all bounded measurable functions \(f\) on \(S\).

The cost rate: Let \(\bar{r}: S \times U_1 \times U_2 \rightarrow [0, \infty)\) be a measurable function which denotes the running cost function. Throughout this paper, we assume that the running cost functions \(\bar{r}: S \times U \rightarrow [0, \infty)\), is continuous in the second and third arguments for each of the first argument \(x \in S\) and \(\|\bar{r}\|_{\infty} := \sup_{(x, u_1, u_2) \in S \times U} |\bar{r}(x, u_1, u_2)| < \infty\).

We now want to introduce the concept of a randomized Markov strategy. Let \(V_i = \mathcal{P}(U_i)\ i = 1, 2\), the spaces of probability measures on \(U_i\) with Prohorov topology.
A family $v_i = (v_i(t)) \ (t \geq 0)$ is said to be a randomized Markov strategy for the player $i$, for $i = 1, 2$, if the following properties hold.

1. For each $t \geq 0$, $v_i(D|t, \cdot)$ is a Borel measurable function on $S$ where $D \in \mathcal{B}(U_i)$, and for each $x \in S$, $v_i(|t, x)$ is a probability measure on $U_i$. That is $v_i(t)$ is a stochastic kernel on $U_i$ given $S$.

2. $v_i(D|t, x)$ is a Borel measurable function in $t \geq 0$ for each $D \in \mathcal{B}(U_i)$ and $x \in S$.

Let $\mathcal{M}_i$ denotes the set of all randomized Markov strategies for player $i$. A Markov strategy $v_i$ is said to be a stationary Markov strategy if for some time invariant stochastic kernel $v_i(\cdot|x)$, on $U_i$ given $S$, satisfying $v_i(\cdot|t, x) = v_i(\cdot|x)$, $\forall x \in S$, and $t \geq 0$. The set of all stationary Markov strategies is denoted by $\mathcal{M}_i$ for $i = 1, 2$. In relaxed control framework the running payoff and transition rate are given by $r: S \times V_1 \times V_2 \rightarrow \mathbb{R}_+$, $q(A|\cdot, \cdot): S \times V_1 \times V_2 \rightarrow \mathbb{R}_+$, defined as follows

$$r(x, v_1(t), v_2(t)) = \int_{U_2} \int_{U_1} r(x, u_1, u_2) v_1(t)(du_1)v_2(t)(du_2)$$

$$q(A|x, v_1(t), v_2(t)) = \int_{U_2} \int_{U_1} q(A|x, u_1, u_2) v_1(t)(du_1)v_2(t)(du_2)$$

for each $x \in S$ $t \geq 0$ and $A \in \mathcal{B}(S)$.

Assumptions (A0)-(A1), guarantee the existence of a Markov process $X(\cdot)$ with transition rate $q(\cdot|x, u_1, u_2)$, corresponding to $(v_1, v_2) \in \mathcal{M}_1 \times \mathcal{M}_2$, for a given initial distribution $\mu \in \mathcal{P}(S)$, see for example [6, Theorem 3.2].

In this paper we consider $\alpha$-discounted zero-sum risk sensitive stochastic game problems which we describe now. The $\alpha$-discounted risk-sensitive payoff for a pair of Markov strategies $(v_1, v_2)$ is given by

$$J^{v_1,v_2}_\alpha(\theta, x) = \frac{1}{\theta} \ln E^{v_1,v_2}_x \left[ e^{\theta \int_0^\infty e^{-\alpha t} r(X(t), v_1(t), v_2(t), X(t)) } dt \right]$$

for some $\theta \in (0, \Theta)$, for a fixed $\Theta > 0$, $\alpha > 0$ is the discount factor, $X(\cdot)$ is the Markov process corresponding to $(v_1, v_2) \in \mathcal{M}_1 \times \mathcal{M}_2$ with $X(0) = x$ and $E^{v_1,v_2}_x$ denote the expectation with respect to the law of the process $X(\cdot)$.

Let $\theta \in (0, \Theta)$ be the “risk-sensitive parameter” chosen by the minimizer. When the state of the system is $x$ and players 1, 2, choose strategies $v_1 \in \mathcal{M}_1$, $v_2 \in \mathcal{M}_2$ respectively, the minimizer (player 1) tries to minimize his infinite-horizon discounted risk-sensitive cost $J^{v_1,v_2}_\alpha(\theta, x)$ over his strategies whereas the maximizer (player 2) tries to maximize the same over his strategies. Such a model is relevant in worst-case scenarios, e.g., in financial applications when a risk-averse investor is trying to minimize his long-term portfolio loss against the market which, by default, is the maximizer in this case.

A strategy $v_1^* \in \mathcal{M}_1$ is called optimal for player 1 for $(\theta, x) \in (0, \Theta) \times S$, if

$$J^{v_1^*,v_2}_\alpha(\theta, x) \leq \sup_{v_2 \in \mathcal{M}_2} \inf_{v_1 \in \mathcal{M}_1} J^{v_1,v_2}_\alpha(\theta, x) (:= J^{\alpha}(\theta, x)) \ \text{for all} \ v_2 \in \mathcal{M}_2.$$ 

Similarly a strategy $v_2^* \in \mathcal{M}_2$ is called optimal for player 2 for $(\theta, x) \in (0, \Theta) \times S$, if

$$J^{v_1,v_2^*}_\alpha(\theta, x) \geq \inf_{v_1 \in \mathcal{M}_1} \sup_{v_2 \in \mathcal{M}_2} J^{v_1,v_2}_\alpha(\theta, x) (:= J^{\alpha}(\theta, x)) \ \text{for all} \ v_1 \in \mathcal{M}_1.$$ 

If $J^{\alpha}(\theta, x) = J^{\alpha}(\theta, x) = J^{\alpha}(\theta, x)$ for all $(\theta, x) \in (0, \Theta) \times S$ then we say that the game has value. A pair of strategies $(v_1^*, v_2^*)$ at which the value of the game is achieved is called saddle-point equilibrium, then $v_2^*$ is optimal for player 2 and $v_1^*$ is
optimal for player 1. Our aim is to prove the existence of saddle-point equilibrium strategies in the class of Markov strategies.

We now list the commonly used notations below.

- $B(S)$ denotes the set of all function $f: S \rightarrow \mathbb{R}$ which are bounded, measurable.
- $\hat{C}_b([a,b] \times S)$ denotes the set of all functions $f: [a,b] \times S \rightarrow \mathbb{R}$ which are measurable and $f(\cdot, x) \in C_b[a,b]$, for each $x \in S$ and $\|f\|_\infty := \sup_{(t,x) \in [a,b] \times S} |f(t,x)| < \infty$.
- $\hat{C}_1((a,b) \times S)$ denotes the set of all functions $f: (a,b) \times S \rightarrow \mathbb{R}$ which are measurable and $f(\cdot, x) \in C_1(a,b)$, for each $x \in S$.
- $C^\infty_c(a,b)$ denotes the set of all infinitely differentiable functions on $(a,b)$ with compact support.

3. Analysis of zero-sum game. In this section we study zero-sum $\alpha$-discounted risk sensitive stochastic game for pure jump processes with general state space and we prove the existence of saddle point equilibrium. We carry out our analysis of the discounted payoff evolution criterion via the criterion:

$$\beta^\alpha_{v_1,v_2} (\theta, x) := E^\alpha_x \left[ e^{\theta \int_0^\infty e^{-\alpha t} r(X(t^-), v_1(t), v_2(t, X(t^-))) dt} \right].$$

(3.1)

Since logarithm is an increasing function, therefore the optimal strategies for the criterion (2.1) are optimal for the criterion (3.1).

Corresponding to the payoff evolution criterion (3.1), the value function is defined as

$$\overline{\psi}_\alpha (\theta, x) = \inf_{v_1 \in M_1} \sup_{v_2 \in M_2} \beta^\alpha_{v_1,v_2} (\theta, x),$$

and

$$\underline{\psi}_\alpha (\theta, x) = \sup_{v_2 \in M_2} \inf_{v_1 \in M_1} \beta^\alpha_{v_1,v_2} (\theta, x).$$

Using dynamic programming heuristics, the Hamilton-Jacobi-Isaacs (HJI) equations for the discounted payoff criterion is given by

$$\alpha \theta \frac{\partial \psi_\alpha}{\partial \theta} (\theta, x) = \inf_{v_1 \in \mathcal{V}_1} \sup_{v_2 \in \mathcal{V}_2} \left[ \mathcal{L}_{v_1,v_2} \psi_\alpha (\theta, x) + \theta r(x, v_1, v_2) \psi_\alpha (\theta, x) \right],$$

$$\psi_\alpha (0, x) = 1,$$

(3.2)

where $\mathcal{L}_{v_1,v_2} f(\theta, x) := \int_S f(\theta, y) q(dy|x, v_1, v_2)$ (second equality in (3.2) follows from Fan mini-max Theorem [3]).

Now we prove that the equation (3.2) has a unique solution in $\hat{C}_b((0, \Theta) \times S) \cap \hat{C}_1((0, \Theta) \times S)$. In order to prove this, first we show that certain ordinary integro-differential equation (ODE) admits a unique solution.

**Lemma 3.1.** Suppose that assumptions (A0)-(A1) hold. Then for any fixed $\epsilon \in (0, \Theta]$ the following ordinary integro-differential equation (ODE)

$$\alpha \theta \frac{\partial \psi_\alpha^\epsilon}{\partial \theta} (\theta, x) = \inf_{v_1 \in \mathcal{V}_1} \sup_{v_2 \in \mathcal{V}_2} \left[ \mathcal{L}_{v_1,v_2} \psi_\alpha^\epsilon (\theta, x) + \theta r(x, v_1, v_2) \psi_\alpha^\epsilon (\theta, x) \right],$$

$$\psi_\alpha^\epsilon (\epsilon, x) = e^{\frac{\epsilon}{\alpha} \| r \|_\infty} (:= h_\epsilon),$$

(3.3)
where $\| \cdot \|_{\infty}$ denotes the supnorm, admits a unique solution $\psi_{\alpha}^\epsilon \in \tilde{C}_b((\epsilon, \Theta) \times S) \cap \tilde{C}^1((\epsilon, \Theta) \times S)$.  

Proof. Let $\delta > 0$. Define the nonlinear operator $T : \tilde{C}_b((\epsilon, \epsilon + \delta) \times S) \to \tilde{C}_b((\epsilon, \epsilon + \delta) \times S)$ by

$$
Tf(\eta, x) := e^{\delta \|r\|_{\infty}} + \frac{1}{\alpha} \int_\epsilon^\eta \inf_{v_1 \in V_1} \sup_{v_2 \in V_2} \left[ \frac{1}{\alpha} \mathcal{L}_{v_1, v_2} f(\theta, x) + r(x, v_1, v_2) f(\theta, x) \right] d\theta.
$$

By (A1) and the fact $r$ is bounded, we have

$$
\| Tf_1 - Tf_2 \|_{\infty} \leq \frac{1}{\alpha} \left[ \| r \|_{\infty} \delta + 2M \ln \left( 1 + \frac{\delta}{\epsilon} \right) \right] \| f_1 - f_2 \|_{\infty}.
$$

Choose $\delta$ such that $\frac{1}{\alpha} \left[ \| r \|_{\infty} \delta + 2M \ln \left( 1 + \frac{\delta}{\epsilon} \right) \right] < 1$. Then $T$ is a contraction operator. Therefore, the contraction mapping theorem yields the existence of a unique solution $\psi_{\alpha}^\epsilon$ in $\tilde{C}_b((\epsilon, \epsilon + \delta) \times S)$ for the OIDE (3.3). Now using (A1) and the continuity of $r$ it follows that $\psi_{\alpha}^\epsilon \in \tilde{C}_b((\epsilon, \epsilon + \delta) \times S) \cap \tilde{C}^1((\epsilon, \epsilon + \delta) \times S)$. Proceeding in this way we get a $\tilde{C}^1((\epsilon, \Theta) \times S) \cap \tilde{C}_b((\epsilon, \Theta) \times S)$ solution for the OIDE (3.3).

Now, by (A1)(ii) and continuity of $\tilde{r}$, an application of standard measurable selection theorem [2], we have there exist measurable functions $\tilde{v}_i : (0, \Theta) \times S \to V_i, i = 1, 2$, such that

$$
\inf_{v_1 \in V_1} \sup_{v_2 \in V_2} \left[ \mathcal{L}_{v_1, v_2} \psi_{\alpha}^\epsilon (\theta, x) + \alpha r(x, v_1, v_2) \psi_{\alpha}^\epsilon (\theta, x) \right] =
\sup_{v_2 \in V_2} \left( \inf_{v_1 \in V_1} \left[ \mathcal{L}_{v_1, v_2} \psi_{\alpha}^\epsilon (\theta, x) + \alpha r(x, v_1, v_2) \psi_{\alpha}^\epsilon (\theta, x) \right] \right)

\text{and}

\inf_{v_1 \in V_1} \sup_{v_2 \in V_2} \left[ \mathcal{L}_{v_1, v_2} \psi_{\alpha}^\epsilon (\theta, x) + \alpha r(x, v_1, v_2) \psi_{\alpha}^\epsilon (\theta, x) \right] =
\sup_{v_2 \in V_2} \left( \inf_{v_1 \in V_1} \left[ \mathcal{L}_{v_1, v_2} \psi_{\alpha}^\epsilon (\theta, x) + \alpha r(x, v_1, v_2) \psi_{\alpha}^\epsilon (\theta, x) \right] \right). \tag{3.4}
$$

Let

$$
\tilde{v}_i^\ast : \mathbb{R}_+ \times S \to V_i, i = 1, 2,
$$

be defined as

$$
\tilde{v}_i^\ast (t, x) = \tilde{v}_i (\theta e^{-\alpha t}, x), \ i = 1, 2.
$$

Define $\theta(t) = \theta e^{-\alpha t}$ and

$$
T_\epsilon = \inf \{ t \geq 0 : \theta e^{-\alpha t} = \epsilon \}.
$$

Then by applying Dynkin’s formula to

$$e^{\int_0^{T_\epsilon} \theta(s) r(X(s), \tilde{v}_1^\ast (s, X(s)), v_2(s, X(s))) ds} \psi_{\alpha}^\epsilon (\theta(t), X(t))$$

for $(\tilde{v}_1^\ast, v_2) \in \mathcal{M}_1 \times \mathcal{M}_2$ and initial condition $x$, we get

$$
E_x \left[ e^{\int_0^{T_\epsilon} \theta(s) r(X(s), \tilde{v}_1^\ast (s, X(s)), v_2(s, X(s))) ds} \psi_{\alpha}^\epsilon (\theta(T_\epsilon), X(T_\epsilon)) \right] - \psi_{\alpha}^\epsilon (\theta(x))
$$

$$
= E_x \left[ \int_0^{T_\epsilon} e^{\int_0^{s} \theta(r) r(X(r), \tilde{v}_1^\ast (r, X(r)), v_2(s, X(s))) ds} \mathcal{L}_{\tilde{v}_1^\ast, v_2} \psi_{\alpha}^\epsilon (\theta(t), X(t)) dt \right]
$$

$$
- \alpha \theta(t) \frac{\partial \psi_{\alpha}^\epsilon}{\partial \theta}(\theta(t), X(t)) + \mathcal{L}_{\tilde{v}_1^\ast, v_2} \psi_{\alpha}^\epsilon (\theta(t), X(t))
$$

$$
+ \theta(t) r(X(t), \tilde{v}_1^\ast (t, X(t)), v_2(t, X(t))) \psi_{\alpha}^\epsilon (\theta(t), X(t)) dt \right].
$$
Since \( \psi_\alpha \) satisfies (3.3) and (3.4), the right-hand side term is less or equals to zero. Therefore we obtain

\[
\psi_\alpha(\theta, x) \geq E^\alpha_{\psi, v_2} \left[ e^{\int_{T_0}^{T_\theta} \theta(s)r(X(s-), \bar{\psi}(s, X(s-), v_2(s, X(s-)))) ds} \psi_\alpha'(\theta(T_\theta), X(T_\theta)) \right]
\]

\[
= E^\alpha_{\psi, v_2} \left[ h_\epsilon e^{\int_{T_0}^{T_\theta} \theta(s)r(X(s-), \bar{\psi}(s, X(s-), v_2(s, X(s-)))) ds} \right].
\]

Since \( v_2(\cdot) \) is arbitrary, it follows that

\[
\psi_\alpha(\theta, x) \geq \sup_{v_2 \in \mathcal{M}_2} E^\alpha_{\psi, v_2} \left[ h_\epsilon e^{\int_{T_0}^{T_\theta} e^{-\alpha r} r(X(s-), \bar{\psi}(s, X(s-), v_2(s, X(s-)))) ds} \right]. 
\]  

(3.6)

Combining (3.6) and (3.7) we have

\[
\psi_\alpha(\theta, x) = \sup_{v_1 \in \mathcal{M}_1} \sup_{v_2 \in \mathcal{M}_2} E^\alpha_{\psi, v_1, v_2} \left[ h_\epsilon e^{\int_{T_0}^{T_\theta} e^{-\alpha r} r(X(s-), v_1(s, X(s-)), v_2(s, X(s-))) ds} \right]
\]

\[
= \sup_{v_2 \in \mathcal{M}_2} \inf_{v_1 \in \mathcal{M}_1} E^\alpha_{\psi, v_1, v_2} \left[ h_\epsilon e^{\int_{T_0}^{T_\theta} e^{-\alpha r} r(X(s-), v_1(s, X(s-)), v_2(s, X(s-))) ds} \right]. 
\]  

(3.7)

(3.8)

From the above representation it is easy to see that the solution \( \psi_\alpha \) of (3.3) is unique. This completes the proof.

Thus, from (3.8), we deduce that

\[
1 \leq \psi_\alpha(\theta, x) = \sup_{v_1 \in \mathcal{M}_1} \sup_{v_2 \in \mathcal{M}_2} E^\alpha_{\psi, v_1, v_2} \left[ h_\epsilon e^{\int_{T_0}^{T_\theta} e^{-\alpha r} r(X(s-), v_1(s, X(s-)), v_2(s, X(s-))) ds} \right]
\]

\[
\leq e^{\frac{\alpha}{\theta} \| \theta \|_\infty} \quad \forall \theta \in (\epsilon, \Theta).
\]

Now we show that \( \frac{\partial \psi_\alpha}{\partial \theta} \) is bounded uniformly in \( \epsilon > 0 \).

**Lemma 3.2.** Suppose that assumptions (A0) – (A1) hold. Then for fixed \( \alpha > 0 \), the derivative \( \frac{\partial \psi_\alpha}{\partial \theta} \) is bounded uniformly in \( \epsilon \in (0, \Theta) \).

**Proof.** For a fixed \( \alpha > 0 \) and \( \delta > 0 \) small enough, consider

\[
|\psi_\alpha(\theta + \delta, x) - \psi_\alpha(\theta, x)|
\]

\[
\leq \sup_{v_1 \in \mathcal{M}_1} \sup_{v_2 \in \mathcal{M}_2} \left| E^\alpha_{\psi, v_1, v_2} \left[ h_\epsilon e^{\int_{T_0}^{T_{\theta+\delta}} e^{-\alpha r} r(X(t-), v_1(t, X(t-)), v_2(t, X(t-))) dt} \right] - E^\alpha_{\psi, v_1, v_2} \left[ h_\epsilon e^{\int_{T_0}^{T_{\theta}} e^{-\alpha r} r(X(t-), v_1(t, X(t-)), v_2(t, X(t-))) dt} \right] \right| \leq I_1 + I_2,
\]

where \( (\theta + \delta)e^{-\alpha T_{\theta+\delta}} = \epsilon \), and

\[
I_1 = \sup_{v_1 \in \mathcal{M}_1} \sup_{v_2 \in \mathcal{M}_2} \left| E^\alpha_{\psi, v_1, v_2} \left[ h_\epsilon e^{\int_{T_0}^{T_{\theta+\delta}} e^{-\alpha r} r(X(t-), v_1(t, X(t-)), v_2(t, X(t-))) dt} \right] - E^\alpha_{\psi, v_1, v_2} \left[ h_\epsilon e^{\int_{T_0}^{T_{\theta}} e^{-\alpha r} r(X(t-), v_1(t, X(t-)), v_2(t, X(t-))) dt} \right] \right|
\]

\[
I_2 = \sup_{v_1 \in \mathcal{M}_1} \sup_{v_2 \in \mathcal{M}_2} \left| E^\alpha_{\psi, v_1, v_2} \left[ h_\epsilon e^{\int_{T_0}^{T_{\theta+\delta}} e^{-\alpha r} r(X(t-), v_1(t, X(t-)), v_2(t, X(t-))) dt} \right] \right|
\]

\[
- E^\alpha_{\psi, v_1, v_2} \left[ h_\epsilon e^{\int_{T_0}^{T_{\theta}} e^{-\alpha r} r(X(t-), v_1(t, X(t-)), v_2(t, X(t-))) dt} \right],
\]
Now for $\epsilon$, we have

$$I_1 \leq |h_\epsilon| \sup_{v_1 \in \mathcal{M}_1} \sup_{v_2 \in \mathcal{M}_2} \left| E_x^{v_1,v_2} \left[ e^{\theta \int_0^{T^*_\epsilon} e^{-\alpha t} r(X(t),v_1(t),X(t)),v_2(t)) dt} \right] \right| \leq |h_\epsilon| |r| \|r\| \epsilon \frac{e^{\frac{\theta}{\alpha} \|r\|}}{\epsilon \|r\|},$$

for $\delta > 0$ small enough. For $I_2$, we have

$$I_2 \leq |h_\epsilon| \sup_{v_1 \in \mathcal{M}_1} \sup_{v_2 \in \mathcal{M}_2} \left| E_x^{v_1,v_2} \left[ e^{\theta \int_0^{T^*_\epsilon} e^{-\alpha t} r(X(t),v_1(t),X(t)),v_2(t)) dt} \right] \right| \leq |h_\epsilon| |r| \|r\| (e^{\frac{\theta}{\alpha} \|r\|} - 1).$$

But $\theta (e^{-\alpha T^*_\epsilon} - e^{-\alpha T^*_\epsilon}) = \delta e^{-\alpha T^*_\epsilon} = \frac{\epsilon \delta}{\epsilon + \delta}$. Hence we have

$$I_2 \leq |h_\epsilon| |r| \|r\| \epsilon \frac{e^{\frac{\theta}{\alpha} \|r\|}}{\epsilon \|r\|},$$

for some constant $C > 0$ and for $\delta > 0$ small enough. Hence we conclude that for $\delta > 0$ small enough

$$|\psi_\alpha^\epsilon(\theta + \delta, x) - \psi_\alpha^\epsilon(\theta, x)| \leq C_1 |h_\epsilon| |r| \|r\| \epsilon \frac{e^{\frac{\theta}{\alpha} \|r\|}}{\epsilon \|r\|},$$

for some constant $C_1 > 0$.

Similarly for $\delta < 0$, small enough, we can get an estimate of the same type

$$|\psi_\alpha^\epsilon(\theta + \delta, x) - \psi_\alpha^\epsilon(\theta, x)| \leq C_1 |h_\epsilon| |r| \|r\| \epsilon \frac{e^{\frac{\theta}{\alpha} \|r\|}}{\epsilon \|r\|}.$$
Case 1. If $\epsilon + \delta < \theta$ then (using the representation as in (3.8))
\[
\psi_{e+\delta}^\theta(\theta, x) - \bar{\psi}_e^\theta(\theta, x)
\geq E_{\bar{x}}\int_{\epsilon+\delta}^{\bar{\alpha}} \int_{\epsilon+\delta}^{2\epsilon} \left[ h_{\epsilon} e^{\theta} \int_{0}^{T_{e+\delta}^{\theta}} e^{-\alpha t} r(X(t-), \bar{v}_{e+\delta}^\theta, t, (X(t-)))dt \right]
- E_{\bar{x}}\int_{\epsilon+\delta}^{\bar{\alpha}} \int_{\epsilon+\delta}^{2\epsilon} \left[ h_{\epsilon} e^{\theta} \int_{0}^{T_{e+\delta}^{\theta}} e^{-\alpha t} r(X(t-), \bar{v}_{e+\delta}^\theta, t, (X(t-)))dt \right]
\geq E_{\bar{x}}\int_{\epsilon+\delta}^{\bar{\alpha}} \int_{\epsilon+\delta}^{2\epsilon} \left[ h_{\epsilon} e^{\theta} \int_{0}^{T_{e+\delta}^{\theta}} e^{-\alpha t} r(X(t-), \bar{v}_{e+\delta}^\theta, t, (X(t-)))dt \right]
- E_{\bar{x}}\int_{\epsilon+\delta}^{\bar{\alpha}} \int_{\epsilon+\delta}^{2\epsilon} \left[ h_{\epsilon} e^{\theta} \int_{0}^{T_{e+\delta}^{\theta}} e^{-\alpha t} r(X(t-), \bar{v}_{e+\delta}^\theta, t, (X(t-)))dt \right]
\geq 0,
\]
where $v_{e+\delta}^\theta$ and $\bar{v}_{e+\delta}^\theta$ are optimal strategies as in (3.4) and (3.5) corresponding to the functions $\psi_{e+\delta}^\theta, \bar{\psi}_e^\theta$ respectively.

Case 2. If $\epsilon < \theta \leq \epsilon + \delta$ then
\[
\bar{\psi}_{e+\delta}^\theta(\theta, x) - \bar{\psi}_e^\theta(\theta, x)
= h_{\epsilon+\delta} - E_{\bar{x}}\int_{\epsilon+\delta}^{\bar{\alpha}} \int_{\epsilon+\delta}^{2\epsilon} \left[ h_{\epsilon} e^{\theta} \int_{0}^{T_{e+\delta}^{\theta}} e^{-\alpha t} r(X(t-), \bar{v}_{e+\delta}^\theta, t, (X(t-)))dt \right]
- E_{\bar{x}}\int_{\epsilon+\delta}^{\bar{\alpha}} \int_{\epsilon+\delta}^{2\epsilon} \left[ h_{\epsilon} e^{\theta} \int_{0}^{T_{e+\delta}^{\theta}} e^{-\alpha t} r(X(t-), \bar{v}_{e+\delta}^\theta, t, (X(t-)))dt \right]
\geq h_{\epsilon} \left( h_{\epsilon+\delta} - E_{\bar{x}}\int_{\epsilon+\delta}^{\bar{\alpha}} \int_{\epsilon+\delta}^{2\epsilon} \left[ h_{\epsilon} e^{\theta} \int_{0}^{T_{e+\delta}^{\theta}} e^{-\alpha t} r(X(t-), \bar{v}_{e+\delta}^\theta, t, (X(t-)))dt \right] \right) \geq 0.
\]

Case 3. If $\theta \leq \epsilon$ then
\[
\bar{\psi}_{e+\delta}^\theta(\theta, x) - \bar{\psi}_e^\theta(\theta, x) = h_{\epsilon+\delta} - h_{\epsilon} = h_{\epsilon}(h_{\delta} - 1) \geq 0.
\]

From the above calculation it is clear that $\psi_{e+\delta}^\theta(\theta, x)$ is decreasing as $\epsilon \to 0$ for any $(\theta, x)$. Since $\psi_{e+\delta}^\theta, \frac{\partial \psi_{e+\delta}^\theta}{\partial \theta}$ are uniformly bounded (in $\epsilon > 0$) and $\bar{\psi}_e^\theta$ is Lipschitz continuous ( w.r.t $\theta$) with the Lipschitz constant uniformly bounded in $\epsilon > 0$ and $x \in S$. Moreover $\psi_{e+\delta}^\theta(\theta, x)$ is decreasing as $\epsilon \to 0$ for any $(\theta, x)$, therefore there exists a function $\psi_e(\theta, x) \in C_b((0, \Theta) \times S)$ such that along a subsequence $\epsilon_m \to 0$, we have $\psi_{e+\delta}^\theta(\theta, x) \to \psi_e(\theta, x)$ uniformly over compact subset of $(0, \Theta)$ and $\psi_{e+\delta}^\theta(\theta, x) \to \psi_e(\theta, x)$ for each $(\theta, x)$. Now for $\varphi \in C_c^\infty(0, \Theta)$, we have
\[
-\int_0^\Theta \frac{d\varphi}{d\theta}(\theta) \psi_{e+\delta}^\theta(\theta, x) d\theta = \int_0^\Theta \frac{d\varphi}{d\theta}(\theta) \psi_e(\theta, x) d\theta
= \int_0^\Theta \inf_{v_1 \in V_1, v_2 \in V_2} \left[ \mathcal{L}_{v_1, v_2} \bar{\psi}_{e+\delta}^\theta(\theta, x) + \theta E_{v_1, v_2} \bar{\psi}_{e+\delta}^\theta(\theta, x) \right] \varphi(\theta) d\theta
\]
\[-\int_0^{\epsilon_m} \inf_{v_1 \in V_1} \sup_{v_2 \in V_2} \left[ \mathcal{L}_{v_1, v_2} \tilde{\psi}_\alpha^{v_m}(\theta, x) + \theta r(x, v_1, v_2) \tilde{\psi}_\alpha^{v_m}(\theta, x) \right] \varphi(\theta) d\theta \]
\[= \int_0^{\Theta} \inf_{v_1 \in V_1} \sup_{v_2 \in V_2} \left[ \mathcal{L}_{v_1, v_2} \tilde{\psi}_\alpha^{v_m}(\theta, x) + \theta r(x, v_1, v_2) \tilde{\psi}_\alpha^{v_m}(\theta, x) \right] \varphi(\theta) d\theta \]
\[= \int_0^{\epsilon_m} \inf_{v_1 \in V_1} \sup_{v_2 \in V_2} \left[ \theta r(x, v_1, v_2) \tilde{\psi}_\alpha^{v_m}(\theta, x) \right] \varphi(\theta) d\theta. \]

For each \(x \in S\) and \(v_i \in V_i, i = 1, 2\), define

\[Q(D|x, v_1, v_2) := \frac{q(D|x, v_1, v_2)}{M + 1} + I_D(x), \text{ for all } D \in \mathcal{B}(S).\]

We claim that \(Q(\cdot|x, v_1, v_2)\) is a stochastic kernel on \(S\). It suffices to show that \(Q(\cdot|x, v_1, v_2)\) is a probability measure on \(\mathcal{B}(S)\) for fixed \(x \in S\) and \(v_i \in V_i, i = 1, 2\).

Note that \(Q(S|x, v_1, v_2) = 1\) and by assumption (A0)(i), \(Q(\cdot|x, v_1, v_2)\) is countably additive. Also note that

\[Q(D|x, v_1, v_2) = \begin{cases} 
1 + \frac{q(D|x, v_1, v_2)}{M + 1} & \text{if } x \in D \\
\frac{q(D|x, v_1, v_2)}{M + 1} & \text{if } x \in D^c. \end{cases} \]  

(3.9)

For \(x \in D\), by (A0), we obtain

\[1 > -\frac{q(\{x\}|x, v_1, v_2)}{M + 1} \geq -\frac{q(S \setminus \{x\}|x, v_1, v_2)}{M + 1} - \frac{q(D \setminus \{x\}|x, v_1, v_2)}{M + 1} = -\frac{q(D|x, v_1, v_2)}{M + 1} = \frac{q(S \setminus D|x, v_1, v_2)}{M + 1} \geq 0.\]

Therefore, \(-1 < -\frac{q(D|x, v_1, v_2)}{M + 1} \leq 0\). If \(x \notin D\), then

\[1 > -\frac{q(\{x\}|x, v_1, v_2)}{M + 1} = \frac{q(S \setminus \{x\}|x, v_1, v_2)}{M + 1} \geq \frac{q(D|x, v_1, v_2)}{M + 1} \geq 0.\]

Hence, from the above we can conclude that \(Q(\cdot|x, v_1, v_2)\) is a probability measure on \(S\). Now rewriting the above equation in terms of \(Q\), it follows that

\[-\int_0^{\Theta} \left\{ \frac{\alpha}{M + 1} \tilde{\psi}_\alpha^{v_m}(\theta, x) \frac{d(\theta \varphi)}{d\theta}(\theta, x) - \tilde{\psi}_\alpha^{v_m}(\theta, x) \right\} d\theta \]
\[= \int_0^{\Theta} \inf_{v_1 \in V_1} \sup_{v_2 \in V_2} \left[ \int_S \tilde{\psi}_\alpha^{v_m}(\theta, y)Q(\text{d}y|x, v_1, v_2) + \frac{\theta}{M + 1} r(x, v_1, v_2) \tilde{\psi}_\alpha^{v_m}(\theta, x) \right] \varphi(\theta) d\theta \]
\[-\frac{1}{M + 1} \int_0^{\epsilon_m} \inf_{v_1 \in V_1} \sup_{v_2 \in V_2} \left[ \theta r(x, v_1, v_2) \tilde{\psi}_\alpha^{v_m}(\theta, x) \right] \varphi(\theta) d\theta. \]  

(3.10)

From (3.4) for the function \(\tilde{\psi}_\alpha^{v_m}\) we have for any \(v_2 \in \mathcal{M}_2\)

\[-\int_0^{\Theta} \left\{ \frac{\alpha}{M + 1} \tilde{\psi}_\alpha^{v_m}(\theta, x) \frac{d(\theta \varphi)}{d\theta}(\theta, x) - \tilde{\psi}_\alpha^{v_m}(\theta, x) \right\} d\theta \]
\[\geq \int_S \tilde{\psi}_\alpha^{v_m}(\theta, y)Q(\text{d}y|x, \tilde{\psi}_\alpha^{v_m}(\theta, x), v_2(\theta, x)) \]
\[+ \frac{\theta}{M + 1} r(x, \tilde{\psi}_\alpha^{v_m}(\theta, x), v_2(\theta, x)) \tilde{\psi}_\alpha^{v_m}(\theta, x) \right\} \varphi(\theta) d\theta \]
\[-\frac{1}{M + 1} \int_0^{\epsilon_m} \inf_{v_1 \in V_1} \sup_{v_2 \in V_2} \left[ \theta r(x, v_1, v_2) \tilde{\psi}_\alpha^{v_m}(\theta, x) \right] \varphi(\theta) d\theta. \]  

(3.11)

Since \(V_1\) is compact, hence along a subsequence, denoted by the same notation with an abuse of notation, we have \(\tilde{v}_{1, v_m} \rightarrow v_1^*\). Letting \(m \uparrow \infty\) along the above
subsequence and using generalized Fatou’s Lemma from [12], Lemma 8.3.7, we obtain
\[
\lim_{m \to \infty} \int_S \tilde{\psi}_m^m(\theta, y)Q(dy|x, \tilde{v}_m^m(\theta, x), v_2(\theta, x)) + \frac{\theta r(x, \tilde{v}_m^m(\theta, x), v_2(\theta, x))}{M + 1} \tilde{\psi}_m^m(\theta, x)
\]
\[
= \int_S \psi_\alpha(\theta, y)Q(dy|x, \psi_\alpha^\dagger(\theta, x), v_2(\theta, x)) + \frac{\theta r(x, \psi_\alpha^\dagger(\theta, x), v_2(\theta, x))}{M + 1} \psi_\alpha(\theta, x).
\]
Now using the dominated convergence theorem, we have
\[
- \int_0^\Theta \left\{ \frac{\alpha}{M + 1} \psi_\alpha(\theta, x) \frac{d(\theta \varphi)}{d\theta}(\theta, x) - \psi_\alpha(\theta, x) \right\} d\theta 
\geq \int_0^\Theta \left[ \int_S \psi_\alpha(\theta, y)Q(dy|x, \psi_\alpha^\dagger(\theta, x), v_2(\theta, x)) + \frac{\theta r(x, \psi_\alpha^\dagger(\theta, x), v_2(\theta, x))}{M + 1} \psi_\alpha(\theta, x) \right] \varphi(\theta) d\theta.
\]
Therefore using definition of $Q$, we obtain
\[
- \int_0^\Theta \alpha \psi_\alpha(\theta, x) \frac{d(\theta \varphi)}{d\theta}(\theta, x) d\theta = \int_0^\Theta \frac{\partial \psi_\alpha}{\partial \theta}(\theta, x) \theta \varphi d\theta 
\geq \int_0^\Theta \left[ \int_S \psi_\alpha(\theta, y)q(dy|x, \psi_\alpha^\dagger(\theta, x), v_2(\theta, x)) + \theta r(x, \psi_\alpha^\dagger(\theta, x), v_2(\theta, x)) \psi_\alpha(\theta, x) \right] \varphi(\theta) d\theta.
\]
This implies that
\[
\int_0^\Theta \frac{\partial \psi_\alpha}{\partial \theta}(\theta, x) \theta \varphi d\theta 
\geq \sup_{v_2 \in V_2} \int_0^\Theta \left[ \int_S \psi_\alpha(\theta, y)q(dy|x, \psi_\alpha^\dagger(\theta, x), v_2) + \theta r(x, \psi_\alpha^\dagger(\theta, x), v_2) \psi_\alpha(\theta, x) \right] \varphi(\theta) d\theta.
\]
Arguing as above and using (3.7) instead of (3.6) we get
\[
\int_0^\Theta \frac{\partial \psi_\alpha}{\partial \theta}(\theta, x) \theta \varphi d\theta 
\leq \inf_{v_1 \in V_1} \int_0^\Theta \left[ \int_S \psi_\alpha(\theta, y)q(dy|x, v_1, v_2) + \theta r(x, v_1, v_2) \psi_\alpha(\theta, x) \right] \varphi(\theta) d\theta.
\]
Combining (3.12) and (3.13) we have
\[
\int_0^\Theta \frac{\partial \psi_\alpha}{\partial \theta}(\theta, x) \theta \varphi d\theta 
= \inf_{v_1 \in V_1} \sup_{v_2 \in V_2} \int_0^\Theta \left[ \int_S \psi_\alpha(\theta, y)q(dy|x, v_1, v_2) + \theta r(x, v_1, v_2) \psi_\alpha(\theta, x) \right] \varphi(\theta) d\theta 
= \sup_{v_2 \in V_2} \inf_{v_1 \in V_1} \left[ L_{v_1, v_2} \psi_\alpha(\theta, x) + \theta r(x, v_1, v_2) \psi_\alpha(\theta, x) \right] 
\]
Therefore we have
\[
\alpha \theta \frac{\partial \psi_\alpha}{\partial \theta}(\theta, x) = \inf_{v_1 \in V_1} \sup_{v_2 \in V_2} \left[ L_{v_1, v_2} \psi_\alpha(\theta, x) + \theta r(x, v_1, v_2) \psi_\alpha(\theta, x) \right] 
\]
\[
= \sup_{v_2 \in V_2} \inf_{v_1 \in V_1} \left[ L_{v_1, v_2} \psi_\alpha(\theta, x) + \theta r(x, v_1, v_2) \psi_\alpha(\theta, x) \right]
\]
Moreover, a saddle point equilibrium exists in $\mathcal{M}$ in the sense of distribution. By (A1) and the continuity of $r$ it follows that $\psi_{\alpha} \in \hat{C}_b((0,\Theta) \times S) \cap \hat{C}^1((0,\Theta) \times S)$. Thus (3.2) has a solution in $\hat{C}_b((0,\Theta) \times S) \cap \hat{C}^1((0,\Theta) \times S)$.

Next we prove that any mini-max selector of (3.2) is a saddle point equilibrium.

**Theorem 3.2.** Suppose that assumptions (A0) − (A1) hold. Then the solution $\psi_{\alpha}$ of (3.2) is unique, in particular, we have $\psi_{\alpha}$ admits the following representation

$$
\psi_{\alpha}(\theta, x) = \inf_{v_1 \in \mathcal{M}_1} \sup_{v_2 \in \mathcal{M}_2} E_{x}^{v_1,v_2} \left[ e^{\theta \int_0^\infty e^{-\alpha s} r(X(s),v_1(s,X(s)),v_2(s,X(s)))ds} \right]
$$

and

$$
\psi_{\alpha}(\theta, x) = \sup_{v_2 \in \mathcal{M}_2} \inf_{v_1 \in \mathcal{M}_1} E_{x}^{v_1,v_2} \left[ e^{\theta \int_0^\infty e^{-\alpha s} r(X(s),v_1(s,X(s)),v_2(s,X(s)))ds} \right].
$$

Moreover, a saddle point equilibrium exists in $\mathcal{M}_1 \times \mathcal{M}_2$.

**Proof.** By standard measurable selection theorem, as above, there exists a measurable function $\bar{v}_i : (0,\Theta) \times S \to V_i, i = 1, 2$, such that

$$
\inf_{v_1 \in \mathcal{V}_1} \sup_{v_2 \in \mathcal{V}_2} \left[ L_{v_1,v_2} \psi_{\alpha}(\theta, x) + \theta r(x, v_1, v_2) \psi_{\alpha}(\theta, x) \right] =
$$

$$
\sup_{v_2 \in \mathcal{V}_2} \left[ L_{\bar{v}_1,v_2} \psi_{\alpha}(\theta, x) + \theta r(x, \bar{v}_1, v_2) \psi_{\alpha}(\theta, x) \right] \quad (3.14)
$$

and

$$
\sup_{v_2 \in \mathcal{V}_2} \inf_{v_1 \in \mathcal{V}_1} \left[ L_{v_1,v_2} \psi_{\alpha}(\theta, x) + \theta r(x, v_1, v_2) \psi_{\alpha}(\theta, x) \right] =
$$

$$
\inf_{v_1 \in \mathcal{V}_1} \left[ L_{\bar{v}_1,v_2} \psi_{\alpha}(\theta, x) + \theta r(x, \bar{v}_1, v_2) \psi_{\alpha}(\theta, x) \right]. \quad (3.15)
$$

Let

$$
v_i^* : \mathbb{R}_+ \times S \to V_i, \ i = 1, 2,
$$

be defined as

$$
v_i^* = \bar{v}_i(\theta e^{-\alpha t}, x), \ i = 1, 2.
$$

For $(v_1^*, v_2) \in \mathcal{M}_1 \times \mathcal{M}_2$, applying Itô formula to the function

$$
e^{\int_0^T \theta(s)r(X(s),v_1^*(s,X(s)),v_2(s,X(s)))ds} \psi_{\alpha}(\theta(t), X(t))
$$

and using (3.14), we get

$$
E_{x}^{v_1^*,v_2} \left[ e^{\int_0^T \theta(s)r(X(s),v_1^*(s,X(s)),v_2(s,X(s)))ds} \psi_{\alpha}(\theta(T), X(T)) \right] - \psi_{\alpha}(\theta, x) \leq 0.
$$

Since $v_2 \in \mathcal{M}_2$ is arbitrary we have

$$
\psi_{\alpha}(\theta, x) \geq \sup_{v_2 \in \mathcal{M}_2} E_{x}^{v_1^*,v_2} \left[ e^{\int_0^T \theta(s)r(X(s),v_1^*(s,X(s)),v_2(s,X(s)))ds} \psi_{\alpha}(\theta(T), X(T)) \right]
$$

Since $1 \leq \psi_{\alpha} \leq e^{\frac{\alpha}{\theta} ||r||_\infty}$ for all $\epsilon > 0$, hence $\psi_{\alpha}(\theta(T), X(T)) \to 1$ uniformly as $T \to \infty$. Therefore by using dominated convergence theorem, letting $T \to \infty$ in the above we obtain

$$
\psi_{\alpha}(\theta, x) \geq \sup_{v_2 \in \mathcal{M}_2} E_{x}^{v_1^*,v_2} \left[ e^{\int_0^\infty \theta(s)r(X(s),v_1^*(s,X(s)),v_2(s,X(s)))ds} \right]. \quad (3.16)
$$

By the similar arguments as above and using (3.15) we can prove that

$$
\psi_{\alpha}(\theta, x) \leq \inf_{v_1 \in \mathcal{M}_1} E_{x}^{v_1^*,v_2} \left[ e^{\int_0^\infty \theta(s)r(X(s),v_1(s,X(s)),v_2^*(s,X(s)))ds} \right]. \quad (3.17)
$$
Thus (3.16) and (3.17) implies that
\[
\psi_\alpha(\theta, x) = \inf_{v_1 \in \mathcal{M}_1} \sup_{v_2 \in \mathcal{M}_2} E_x^{v_1, v_2} \left[ e^{\int_0^\infty \theta(s) r(X(s-), v_1(s, X(s-)), v_2(s, X(s-))) ds} \right]
\]
\[
= \sup_{v_2 \in \mathcal{M}_2} \inf_{v_1 \in \mathcal{M}_1} E_x^{v_1, v_2} \left[ e^{\int_0^\infty \theta(s) r(X(s-), v_1(s, X(s-)), v_2(s, X(s-))) ds} \right]
\]
\[
= E_x^{v_1^*, v_2^*} \left[ e^{\int_0^\infty \theta(s) r(X(s-), v_1^*(s, X(s-)), v_2^*(s, X(s-))) ds} \right],
\]
for any \( v_1 \in \mathcal{M}_1 \) and \( v_2 \in \mathcal{M}_2 \). This implies that \( (v_1^*, v_2^*) \) defined above is a saddle point equilibrium. This completes the proof. \( \square \)

4. Example. In this section, we give an example of a controlled birth and death system to illustrate our assumptions.

Example 4.1 (A controlled birth and death system). Consider an animals birth-and-death system with the state space \( S := \{0, 1, 2, 3, \ldots\} \). The state variable denotes the total animals size at time \( t \geq 0 \). There are natural death and birth rates functions, say \( \mu(x) \) and \( \lambda(x) \), when the state of the system is \( x \in S \). Depending on the number of animals \( x \in S \) in the system, player 1 can modify the death rate by choosing some action \( u_1 \), which will result in an increased death rate equal to \( \mu(x) + h(x, u_1) \), with \( h \) being a function on the action space of player 1 and state of the system. But this action will also result in a cost rate given by \( c(x, u_1) \). On the other hand, player 2 can modify the birth rate by choosing some action \( u_2 \), which will result in an increased birth rate given by \( \lambda(x) + g(x, u_2) \), with \( g \) being a function on the action space of player 2 and state of the system. The action of player 2 results in a reward rate given by \( c_2(x, u_2) \). If at any given time there are \( x \) animals in the system, then it generates a cost at the rate \( c(x) \) for player 1. Now this system can easily be modelled via the game model considered in this paper. So the corresponding transition rate \( q(y|x, u_1, u_2) \) is given as follows. For \( x = 0 \),
\[
q(1|0, u_1, u_2) = -q(0|0, u_1, u_2) := \lambda(0) + g(0, u_2)
\]
and for \( x \geq 1 \)
\[
q(y|x, u_1, u_2) = \begin{cases} 
\mu(x) + h(x, u_1) & \text{if } y = x - 1 \\
\lambda(x) + g(x, u_2) & \text{if } y = x + 1 \\
-(\mu(x) + \lambda(x) + h(x, u_1) + g(x, u_2)) & \text{if } y = x \\
0 & \text{otherwise}
\end{cases}
\]
The cost rate for player 1 is given by \( r(x, u_1, u_2) = c(x) + c_1(x, u_1) - c_2(x, u_2) \).

For a given discounted factor \( \alpha > 0 \), our aim is to find conditions that ensure the existence of saddle point equilibrium. In order to ensure the assumptions (A0) and (A1) are satisfied, we impose the following conditions:

1. \( \mu(x) + h(x, u_1) > 0 \) for all \( x \in S \) and \( u_1 \in U_1 \).
2. \( \lambda(x) + g(x, u_2) > 0 \) for all \( x \in S \) and \( u_2 \in U_2 \).
3. \( c(x) + c_1(x, u_1) \geq c_2(x, u_2) \) for all \( x \in S \) and \( u_1, u_2 \in U_1 \times U_2 \).
(4) The functions $\mu(i), \lambda(x), h(x, u_1), g(x, u_2), c_1(x, u_1), c_2(x, u_2)$ are bounded.
(5) The cost rate $r(x, u_1, u_2)$ and the transition rate $q(y|x, u_1, u_2)$ are continuous in $u_1, u_2$ for all $x, y \in S$.

Acknowledgments. We wish to thank anonymous referees for their suggestions, which improved the presentation of this paper. The research of Somnath Pradhan is supported in part by DST-SERB, India, grant EMR/2016/004810.

REFERENCES
[1] R. Bellman, Dynamic Programming, Princeton University Press, Princeton, NJ, 1957.
[2] V. E. Beneš, Existence of optimal strategies based on specified information, for a class of stochastic decision problems, SIAM J. Control, 8 (1970), 179–188.
[3] K. Fan, Fixed-point and minimax theorems in locally convex topological linear spaces, Proc. Nat. Acad. Sci. U.S.A., 38 (1952), 121–126.
[4] M. K. Ghosh, K. S. Kumar and C. Pal, Zero-sum risk-sensitive stochastic games for continuous time Markov chains, Stoch. Anal. Appl., 34 (2016), 835–851.
[5] M. K. Ghosh and S. Saha, Risk-sensitive control of continuous time Markov chains, Stochastics, 86 (2014), 655–675.
[6] X. Guo, Continuous-time Markov decision processes with discounted rewards: The case of Polish spaces, Math. Oper. Res., 32 (2007), 73–87.
[7] X. Guo and O. Hernández-Lerma, Nonzero-sum games for continuous-time Markov chains with unbounded discounted payoffs, J. Appl. Probab., 42 (2005), 303–320.
[8] X. Guo and O. Hernández-Lerma, Zero-sum games for continuous time jump Markov processes in Polish spaces: Discounted payoffs, Adv. in Appl. Probab., 39 (2007), 645–668.
[9] X. Guo and O. Hernández-Lerma, Zero-sum games for continuous-time Markov chains with unbounded transition and average payoff rates, J. Appl. Probab., 40 (2003), 327–345.
[10] X. Guo and Z.-W. Liao, Risk-sensitive discounted continuous-time Markov decision processes with unbounded rates, SIAM J. Control Optim., 57 (2019), 3857–3883.
[11] X. Guo and Y. Zhang, On risk-sensitive piecewise deterministic Markov decision processes, Appl. Math. Optim., 81 (2020), 685–710.
[12] O. Hernández-Lerma and J. B. Lasserre, Further Topics on Discrete-Time Markov Control Processes, Applications of Mathematics (New York), 42, Springer-Verlag, New York, 1999.
[13] A. S. Nowak, Notes on risk-sensitive Nash equilibria, in Advances in Dynamic Games, Ann. Internat. Soc. Dynam. Games, 7, Birkhäuser Boston, Boston, MA, 2005, 95–109.
[14] C. Pal and S. Pradhan, Risk-sensitive control of pure jump processes on a general state space, Stochastics, 91 (2019), 155–174.
[15] K. Suresh Kumar and C. Pal, Risk-sensitive control of pure jump process on countable space with near monotone cost, Appl. Math. Optim., 68 (2013), 311–331.
[16] K. Suresh Kumar and C. Pal, Risk-sensitive ergodic control of continuous time Markov processes with denumerable state space, Stoch. Anal. Appl., 33 (2015), 863–881.
[17] Q. Wei, Zero-sum games for continuous-time Markov jump processes with risk-sensitive finite-horizon cost criterion, Oper. Res. Lett., 46 (2018), 69–75.
[18] Q. Wei and X. Chen, Stochastic games for continuous-time jump processes under finite-horizon payoff criterion, Appl. Math. Optim., 74 (2016), 273–301.
[19] P. Whittle, Risk-Sensitive Optimal Control, Wiley-Interscience Series in Systems and Optimization, John Wiley & Sons, Ltd., Chichester, 1990.
[20] W. Zhang, Continuous-time constrained stochastic games under the discounted cost criteria, Appl. Math. Optim., 77 (2018), 275–296.
[21] Y. Zhang, Continuous-time Markov decision processes with exponential utility, SIAM J. Control Optim., 55 (2017), 2636–2660.

Received April 2021; revised June 2021.
E-mail address: cpal@iitg.ac.in
E-mail address: somnath@iiserpune.ac.in