A Stochastic Differential Game in the Orthrant¹

Mrinal K. Ghosh and K. Suresh Kumar

Department of Mathematics, Indian Institute of Science, Bangalore 560 012, India
E-mail: mkg@math.iisc.ernet.in; suresh@math.iisc.ernet.in

Submitted by L. Berkovitz

Received October 20, 1999

We study a zero-sum stochastic differential game in the nonnegative orthrant. The state of the system is governed by controlled reflecting diffusions in the nonnegative orthrant. We consider discounted and average payoff evaluation criteria. We prove the existence of values and optimal strategies for both payoff criteria.

Key Words: reflecting diffusions; discounted and average payoff; optimal strategy.

1. INTRODUCTION

We study a zero-sum stochastic differential game in the nonnegative orthrant. This has a potential application in communication networks with heavy traffic. In communication networks we often encounter situations where different users may have different objectives. Each user wishes to optimize a certain performance measure related to his traffic parameters, e.g., minimizing delays, maximizing throughput, minimizing blocking probabilities, etc. Conflicting situations arise when different users cannot coordinate their actions, and hence the problem cannot be reduced to a single control optimization problem. This can happen due to the nature of the network, or some user may be tempted to benefit at the expense of others by choosing an individual “good” policy. This motivates us to treat the problem in a game theoretic framework.

We first present a heuristic description of the network model. A more general problem will be discussed in the next section. Consider a sequence

¹ Research supported by Department of Science and Technology Grant DST/MS/III-045/96.
of open queueing networks consisting of $d$ service stations with increasing traffic intensity. Let the sequence be parameterized by $n$ so that in the heavy traffic limit $n \to \infty$. Let $Q^n_i(t)$ denote the number of customers at the service station $i$ at time $t$ of the $n$th network, and let $X^n_i(t) = \frac{1}{\sqrt{n}} Q^n_i(t)$ be the corresponding normalized process. Let $A^n_i(t)$ be the number of customers arriving at the $i$th server from outside by time $t$. Assume there are $M$ users (i.e., sources which send customers) for the networks. Each user controls the arrival process in an implicit manner. Thus the arrival process is a controlled counting process, the intensity of which is given by

$$
\lambda^n_i(x, u_1, \ldots, u_M) = \sqrt{n} \lambda_i \left( \frac{x_i}{\sqrt{n}}, u_1, \ldots, u_M \right) + n \lambda_i \left( \frac{x_i}{\sqrt{n}} \right),
$$

where $\lambda_i$ and $\lambda_i$ are nonnegative measurable functions, $x = (x_1, x_2, \ldots, x_d)$ is the state of the system, and $x_i$ denotes the number of customers at the $i$th server. The intensity of the arrival process is split into two additive terms with appropriate scaling. In the limiting situation as $n \to \infty$, the function $\lambda_i$ contributes to the drift of the system which is being controlled by the users by the appropriate choices of their actions $u_1, u_2, \ldots, u_M$ from their respective action spaces. The function $\lambda_i$ contributes to the volatility (i.e., the uncertainty) of the system which is not controlled by the users. In other words the users control the mean of the state but are not equipped to control the sudden fluctuations of the state process. For the $n$th network, let $N^n_i(t)$ be the number of customers moving from server $i$ to $j$ by time $t$ and let $N^n_i(t)$ denote the number of customers who complete service at server $i$ by time $t$. Let $N^n_i(t)$ be the number of customers who leave the network from station $i$ by time $t$. Then

$$
N^n_i(t) = \sum_{j=0}^{d} N^n_{ij}(t).
$$

We allow the users to control the service time as well, since by paying additional charges they can opt for better quality of the service. As a result, $N^n_i(t)$ is a controlled counting process with intensity given as follows: when the state is at $x = (x_1, x_2, \ldots, x_d)$ and the users choose their actions $u_1, u_2, \ldots, u_M$ from their respective action spaces, then the intensity of the process is given by

$$
\mu^n_i(x, u_1, \ldots, u_M) = \sqrt{n} \mu_i \left( \frac{x_i}{\sqrt{n}}, u_1, \ldots, u_M \right) + n \mu_i \left( \frac{x_i}{\sqrt{n}} \right),
$$

where $\mu_i, \mu_i$ are nonnegative measurable functions with suitable scaling.
Here again, in the limiting process the function $\Pi_i$ contributes to the drift of the system which is controlled by the users, and $\mu_i$ contributes to the volatility of the system. Let $p_{ij}$, $1 \leq i \leq d$, $0 \leq j \leq d$, represent the probability that a customer goes to server $j$ after completing the service from server $i$. We designate the routing matrix by $P = (p_{ij})_{1 \leq i \leq d}$. To ensure the stability of the system we assume that $P$ has a spectral radius strictly less than 1. We look at the limiting network—more precisely, the dynamics of the limiting network when the traffic is heavy. We say that the traffic in a network is heavy, if at each server the traffic intensity, i.e., the ratio of the mean service time to the mean interarrival time, is nearly 1. Hence as $n \to \infty$ we have to assume that the traffic intensity of the $n$th network tends to 1. To this end we assume that

$$\lambda_i(x_i) + \sum_{j=1}^{d} p_{ij} \mu_j(x_j) - \mu_i(x_i) = 0, \quad x = (x_1, \ldots, x_d). \quad (1.3)$$

This ensures that the traffic intensity tends to 1 as $n \to \infty$. Note that for a single server queue (1.3) reflects the fact that the intensity of the arrival process and the intensity of the service time process are equal, which is precisely the heavy traffic condition. Hence the limiting process represents the dynamics of a network with $d$ servers under heavy traffic assumption. Following [16] we can show, under suitable assumptions, that $X^n(\cdot) = (X^n_1(\cdot), X^n_2(\cdot), \ldots, X^n_d(\cdot))$ converges weakly to the process $X(\cdot)$, which is represented by the stochastic differential equation given by

$$dX_i(t) = \tilde{b}_i(X(t), u_1(t), \ldots, u_M(t)) dt + \int_0^t \sqrt{\lambda_i(X_i(s))} dW_i(s)$$

$$+ \sum_{j=1}^{d} \int_0^t \sqrt{p_{ij}\mu_j(X_j(s))} dB_j(s)$$

$$- \sum_{j=0}^{d} \int_0^t \sqrt{p_{ij}\mu_i(X_i(s))} dB_j(s) + \xi(t) - \sum_{j=1}^{d} p_{ij}\xi_j(s),$$

$$d\xi_i(t) = I\{X_i = 0\} d\xi_i(t), \quad t \geq 0, \ i = 1, 2, \ldots, d,$$

where $X(t) = (X_1(t), \ldots, X_d(t))$, and $W_i(\cdot), B_j(\cdot)$ are the standard independent Wiener process in $\mathbb{R}$, $u_i(t), i = 1, 2, \ldots, M$, $t \geq 0$, are the actions taken by the users at time $t$, and

$$\tilde{b}_i(x, u_1, \ldots, u_M) = \overline{\lambda}_i(x, u_1, \ldots, u_M) + \sum_{j=1}^{d} p_{ij}\overline{\mu}_j(x, u_1, \ldots, u_M)$$

$$- \overline{\mu}_i(x, u_2, \ldots, u_M).$$

14 GHOSH AND KUMAR
The equation (1.4) represents a controlled diffusion process in the non-negative orthrant. By a solution to (1.4) we mean a pair of continuous time processes \((X(\cdot), \xi(\cdot))\), where \(X(\cdot)\) takes values in the nonnegative orthrant. The process \(\xi(\cdot)\) is a continuous time nondecreasing process which increases only at the boundary of the orthrant. When \(X(\cdot)\) hits the boundary it is reflected instantaneously along a vector field which we describe below. The direction of reflection is given by the matrix \(I - P\). When \(X(\cdot)\) hits the boundary point in \(\{x_i = 0\}\), it is reflected along \((I - P)_{ii}\), the \(i\)th row of \(I - P\). On the boundary points \(\{x_i = x_{i_2} = \cdots = x_{i_k} = 0\}\), the direction of reflection is \(\frac{1}{k} [(I - P)_{i_1} + \cdots + (I - P)_{i_k}]\), \(1 \leq i_1, \ldots, i_k, k \geq 2\). Under suitable assumptions, it can be shown that at the corner, the process is never absorbed.

We now describe the differential game problem. Each user (referred to as a player) considers the rest of the players as a single superplayer and tries to find a minimax equilibrium. This gives him an “optimal” strategy against the worst case scenario, i.e., the aim of each player is to guarantee the best performance under the worst case behavior of the superplayer. We can view the situation as follows: each player takes the rest of the players as his adversary. Since the actions of the superplayer are not completely known to the particular player, to achieve his security strategy against the worst case scenario, he assumes that he controls the arrival process, and the superplayer tries to block him by controlling the service time. Thus, the particular player, say player 1, controls the arrival process of the network, and the superplayer controls the service time process through their actions. Hence the drift \(\bar{b}\) takes the following form:

\[
\bar{b}_i(x, u_1, u_2) = \bar{\lambda}_i(x_i, u_1) + \sum_{j=1}^d p_{ji} \bar{\mu}_j(x_j, u_2) - \bar{\mu}_i(x_i, u_2),
\]

where \(u_1\) denotes the action of player 1 and \(u_2\) denotes the actions of the superplayer. (Here \(u_2\) represents the actions chosen by the players 2, 3, \ldots, \(M\). Thus \(u_2\) replaces \((u_2, \ldots, u_M)\) in the previous notation.) We assign a cost to the particular player, i.e., player 1 against the other players (superplayer), as

\[
r(x, u_1, u_2) = \gamma u_1 - \theta u_2 - c(x),
\]

where \(c(\cdot)\) typically represents the holding cost, and \(\gamma > 0\) and \(\theta > 0\) are constants, \(u_1, u_2, \in [0, 1]\). When the state is \(x\) and for the actions \(u_i\) of the player \(i\), the player 1 incurs a cost \(r(x, u_1, u_2)\). Naturally, player 1 tries to minimize the cost through his actions, whereas player 2 (superplayer) tries to maximize the same through his actions. Thus we have reduced the \(M\)-player game to a two-player game. An analogous model in discrete time has been studied by Altman [2]. Motivated by this problem, we study a
more general stochastic differential game with reflecting diffusion in the orthrant, which we describe in the next section. In [5] Borkar and Ghosh have studied stochastic differential games with nondegenerate diffusions for various payoff criteria. Ghosh and Kumar [7] have studied zero-sum stochastic differential games with reflecting (nondegenerate) diffusions in a smooth bounded domain. In this paper we first discretize the orthrant by a sequence of bounded smooth domains. We then use the results from [7] and limiting arguments to derive the corresponding results in the orthrant.

Our paper is organized as follows. In Section 2 we describe a more general problem which subsumes the game problem arising in the network. In Section 3 we discuss the stochastic differential game with the discounted payoff criteria. We show the existence of the value function and optimal Markov strategies for both players. In Section 4, we study the ergodic payoff criterion. Using a Lyapunov-type stability assumption, we prove the existence of the value of the game and optimal Markov strategies for both players. In Section 5 we apply our theory to a simplified model. Section 6 concludes the paper with a few remarks.

2. PROBLEM DESCRIPTION AND PRELIMINARIES

Let \( U_i, i = 1, 2, \) be compact metric spaces and let \( V_i := \mathcal{P}(U_i) \) be the space of probability measures on \( U_i \) endowed with the Prohorov topology. Let \( D = \{ x \in \mathbb{R}^d | x_i > 0, \text{ for all } i = 1, 2, \ldots, d \} \), and let \( \overline{D} \) be the closure of \( D \).

Let \( \tilde{b} : \overline{D} \times U_1 \times U_2 \to \mathbb{R}^d, \sigma : \overline{D} \to \mathbb{R}^{d \times d}, \) and let \( \gamma \) be an \( \mathbb{R}^d \)-valued function defined in a neighborhood of \( \partial D \). Define \( b : \overline{D} \times V_1 \times V_2 \to \mathbb{R}^d \) by

\[
\begin{align*}
    b(x, v_1, v_2) &= \int_{U_1} \int_{U_2} \tilde{b}(x, u_1, u_2)v_1(du_1)v_2(du_2), \\
    &\quad \text{for } x \in \overline{D}, v_1 \in V_1, v_2 \in V_2.
\end{align*}
\]

We consider a stochastic differential game with the state of the game evolving according to a controlled reflecting diffusion in the orthrant \( \overline{D} \). It is represented by the following stochastic differential equation:

\[
\begin{align*}
    dX(t) &= b(X(t), v_1(t), v_2(t)) \, dt + \sigma(X(t)) \, dW(t) - \gamma(X(t)) \, d\xi(t), \\
    d\xi(t) &= I \{ X(t) \in \partial D \} \, d\xi(t), \quad t \geq 0, \\
    X(0) &= x_0 \in \overline{D}, \quad \xi(0) = 0,
\end{align*}
\]

where \( W(\cdot) \) is the standard Wiener process in \( \mathbb{R}^d \) and \( v_i(\cdot) \) is a \( V_i \)-valued process which is progressively measurable with respect to the \( \sigma \)-field generated by \( X(t) \). The process \( v_i(\cdot) \) is called an admissible strategy for the
player $i$, $i = 1, 2$. By a solution to (2.1) we mean a pair of continuous time processes $(X(\cdot), \xi(\cdot))$ satisfying (2.1). The process $X(\cdot)$ is $\bar{D}$ valued. In $D$ it evolves like a controlled diffusion process, and when it hits the boundary, it is reflected instantaneously in the interior of $D$ along the vector field governed by $\gamma$. The process $\xi(\cdot)$ is a nondecreasing process which increases only when $X(\cdot)$ hits the boundary $\partial D$. We call $A_i = \{v_i, v_i(1) = v_i(2) = \cdots = v_i(d)\}$ the set of all admissible strategies for player $i$. An admissible strategy $v_i \in A_i$ is said to be a Markov strategy if

$$v_i(t) = \bar{v}_i(X(t))$$

for some $\bar{v}_i : \bar{D} \rightarrow V_i$ measurable; i.e., the choice of the action by player $i$ at time $t$ depends only on the state of the dynamics at time $t$. By an abuse of notation, we denote the map $\bar{v}_i$ as the Markov strategy for player $i$. We denote by $M_i$ the set of all Markov strategies for the player $i$. The existence of a solution of (2.1) is usually proved by the well-known penalization method [12, 13]. But note that the domain $\bar{D}$ has a corner at the origin. This creates a technical problem. If the direction of reflection $\gamma$ is not chosen properly then the process $X(t)$ may be absorbed at the origin [15], a situation that is unrealistic for application to network problems. To avoid the absorption of $X(t)$ at the origin, we impose appropriate conditions on $\gamma$. The existence of a solution of (2.1) is usually achieved in three steps: (i) approximate $\bar{D}$ by appropriate smooth domains, (ii) establish the existence of a solution to (2.1) in these smooth domains, (iii) use convergence arguments to obtain a solution of (2.1) in $\bar{D}$. To this end we first approximate the nonnegative orthrant $\bar{D}$ in the following way. For $i = 1, 2, \ldots, d$, $m \geq 1$, define

$$C_{im} = \left\{ x \in D \mid \left( x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d \right) - \left( \frac{1}{m}, \frac{1}{m}, \ldots, \frac{1}{m} \right) < \frac{1}{m}, \frac{1}{m} \leq x_i < \infty \right\}$$

and

$$C_{0m} = \left\{ x \in D \mid \left( x_1, x_2, \ldots, x_d \right) - \left( \frac{1}{m}, \frac{1}{m}, \ldots, \frac{1}{m} \right) < \frac{1}{m} \right\}.$$

Set

$$I_{im} = \left\{ x \in D \mid 0 \leq x_j \leq \frac{1}{m}, j \neq i, i, j = 1, \ldots, d, 0 \leq x_i < \infty \right\},$$

$$I_{0m} = \left\{ x \in D \mid 0 \leq x_j \leq \frac{1}{m}, \text{ for all } j = 1, \ldots, d \right\}.$$

Define

$$D_{im} = (D \setminus I_{im}) \cup C_{im}, \quad i \geq 0, \quad D_m = \bigcup_{i=1}^{d} D_{im}.$$
Then $D_m$ satisfies the following properties.

(i) $\overline{D_m} \uparrow \overline{D}$.

(ii) $\partial D_m$, the boundary of $D_m$, is $C^2$.

(iii) For any $K$ compact and $K \subseteq D$, we have $K \subseteq D_m$ for $m$ sufficiently large.

(iv) $\partial D_m \cap C_{im}^c = \partial D \cap C_{im}^c$, $i = 0, 1, \ldots, d$, $m \geq 1$.

To ensure the existence of a unique solution of (2.1), we make the following assumptions on $b$, $\sigma$, and $\gamma$.

(A1) (i) The function $\bar{f}$ is continuous, Lipschitz continuous in its first argument uniformly with respect to the rest.

(ii) The function $\sigma$ is Lipschitz continuous.

(iii) The function defined by $a = \sigma \cdot \sigma'$ is uniformly elliptic, i.e., there exists $\delta_0 > 0$ such that

$$xa(\cdot)x' \geq \delta_0 \|x\|^2,$$

for all $x \in \overline{D}$.

(A2) (i) The function $\gamma$ is such that all of the partial derivatives exist and are continuous and there exists $\delta_1 > 0$ such that

$$\gamma(x) \cdot n(x) \geq \delta_1,$$

for all $x \in \Sigma_0$,

$$\gamma(x) \cdot n_m(x) \geq \delta_1,$$

for all $x \in \partial D_m$,

where

$$\Sigma_0 = \{x \in \partial D \mid x_i = 0 \text{ for at most one } i = 1, \ldots, d\},$$

and $n(\cdot), n_m(\cdot)$ denote, respectively, the outward normal to $\partial D$ (on $\Sigma_0$) and $\partial D_m$. Note that for $m$ large enough $\gamma$ is defined on $\partial D_m$. So we can assume, without loss of generality, that $\gamma$ is defined on $\partial D_m$ for all $m \geq 1$.

(ii) There exists a symmetric matrix valued map $M: \mathbb{R}^d \to \mathbb{R}^{d \times d}$ such that $M = (m_{ij})$, $m_{ij} \in C_0(\mathbb{R}^d) \cap W^{2, \infty}(\mathbb{R}^d)$, for $i, j = 1, 2, \ldots, d$, and satisfies the following:

(a) There exists $\delta_2 > 0$ such that

$$M(x) \geq \delta_2 I_d, \quad \text{for all } x \in \mathbb{R}^d,$$

where $I_d$ denotes the $d \times d$ identity matrix.

(b) There exists $C_0 \geq 0$ such that

$$C_0 \|x - x_0\|^2 + \sum_{i, j=1}^d m_{ij}(x)(x_i - x_i')\gamma_j(x) \geq 0, \quad \text{for all } x \in \partial D, x' \in \overline{D}.$$
(c) Let \( z \in \overline{D}, \; x \in \partial D \). If for some \( C_0 \geq 0 \)
\[
C_0 \|x - x_0\|^2 + \sum_{i,j=1}^{d} m_{ij}(x)(x_i - x'_i)z_j \geq 0, \quad \text{for all } x' \in \overline{D},
\]
then \( z = \theta \gamma(x) \) for some \( \theta \geq 0 \).

For a given \( (v_1, v_2) \in A_1 \times A_2 \), under the assumptions (A1) and (A2), (2.1) has a unique weak solution \cite{12}. Moreover, if the players use Markov strategies then the process \( X(\cdot) \) is strong Markov.

Some comments are in order now.

Remark 2.1. When \( d = 2 \), \( D_m \) has the following simple form:
\[
D_m = \left\{ \left[ \frac{1}{m}, \infty \right) \times [0, \infty) \cup \left\{ [0, \infty) \times \left[ \frac{1}{m}, \infty \right) \right\} \cup B\left( \left( \frac{1}{m}, \frac{1}{m} \right), \frac{1}{m} \right) \right\}.
\]

Remark 2.2. (a) The assumptions (A1)(i), (ii) are the usual Lipschitzian assumptions. The assumption (A1)(iii) is the nondegeneracy condition. In other words we are studying controlled nondegenerate reflecting diffusion in the nonnegative orthrant.

(b) The assumption (A2)(i) ensures that when the process \( X(t) \) hits the boundary \( \partial D \), it is reflected inward instantaneously along a direction governed by \( \gamma \), and it does not “slip” along the boundary.

(c) The assumption (A2)(ii) is rather technical. The condition is used to get suitable estimates on the process \( X(t) \) in appropriate seminorm, which leads to the uniqueness of the solution of (2.1). For a smooth domain, (A2)(ii) follows from (A2)(i); see \cite{12} for details.

(d) Following \cite{5} we give an interpretation of the admissible strategies of the players. Note that an admissible strategy \( v_i(\cdot) \) for player \( i, \; i = 1, 2 \), is a nonanticipative functional of the process \( X(\cdot) \), i.e., \( v_i(t) = f_i(t, X(\cdot)) \) for a measurable, adapted (w.r.t. the \( \sigma \)-field generated by \( X(t) \)) \( f_i(\cdot, \cdot) \). The idea behind this is that whatever extraneous randomization the players might want to incorporate into their controls is already subsumed in the fact that they are choosing \( V_i \)-valued processes rather than \( U_i \)-valued ones. One consequence of this is that the conditional law of \( X(\cdot) \), given \( X(0) = x \), is a.s. the law of a process \( X(x, \cdot) \) controlled by strategies \( v_i(\cdot) = f_i(\cdot, X(x, \cdot)) \), with \( X(x, 0) = x \). Thus we may prescribe the strategies \( \{v_i(\cdot)\} \) for arbitrary initial data by prescribing the \( f_i \)'s. Therefore player 1 chooses the function \( f_1 \), whereas player 2 chooses \( f_2 \). These choices are made independently of each other. This is how the strict noncooperative nature of the game is maintained at all times.
Remark 2.3. We now analyze the network problem discussed in Section 1 as a special case of (2.1). By renaming \((u_2, \ldots, u_M)\) as \(u_2\), the equation (1.4) can be rewritten as

\begin{align*}
\frac{dX(t)}{dt} &= \tilde{b}(X(t), u_1(t), u_2(t)) \, dt + \Sigma(X(t)) \, d\tilde{W}(t) - (I_d - P) \, d\xi(t), \\
\frac{d\xi(t)}{dt} &= I\{X(t) \in \partial D\} \, d\xi(t), \quad t \geq 0, \\
X(0) &= x_0 \in D, \quad \xi(0) = 0,
\end{align*}

where for \(x = (x_1, \ldots, x_d) \in \overline{D}, \, u_i \in U_i, \, i = 1, 2\), \(P\) is the routing matrix,

\[ \tilde{b}_i(x, u_1, u_2) = \lambda_i(x_i, u_1) + \sum_{j=1}^d p_{ij} \mu_j(x_j, u_2) - \tilde{\mu}_i(x_i, u_2), \]

\[ \Sigma(x) = (A(x), B_1(x), \ldots, B_d(x)), \]

with

\[ A(x) = \text{diag} \left( \sqrt{\lambda_1(x_1)}, \ldots, \sqrt{\lambda_d(x_d)} \right), \]

\[ B_i(x) = (B_i^0(x), B_i^1(x), \ldots, B_i^d(x)), \]

and

\[ B_i^j(x) = \left( 0, \ldots, 0, -\sqrt{\mu_j(x_i)}, 0, \ldots, 0 \right)'. \]

Here \(\tilde{W}(\cdot) = (\tilde{W}_i(\cdot), \tilde{B}_i(\cdot))_{1 \leq i \leq d, 0 \leq j \leq d}\). We make the following assumptions for the network problem.

(A1)’

(i) \(\inf_{0 \leq s < \infty} \min \{\lambda_i(x), \mu_i(x)\} > 0\).

(ii) The spectral radius of \(P\) is strictly less than \(1\).

(iii) For \(x = (x_1, \ldots, x_d) \in \overline{D}, \)

\[ (\lambda_1(x_1), \ldots, \lambda_d(x_d))(I - P)^{-1} > 0, \quad (\mu_1(x_1), \ldots, \mu_d(x_d)) > 0; \]

i.e., each element of the vectors is strictly positive.
(iv) For \( x \in \bar{D} \),

\[
\lambda_i(x_i) + \sum_{j=1}^{d} p_{ij} \mu_j(x_j) - \mu_i(x_i) = 0, \quad x = (x_1, \ldots, x_d).
\]

Using (A1)'(ii) we can show that \( \Sigma \Sigma' \) is positive definite. Note that \( \Sigma \) is not a square matrix. So by a martingale representation theorem [10, pp. 382–390] there exists a \( d \times d \) matrix valued function \( \sigma \) satisfying \( \sigma(\cdot)\sigma(\cdot)' = \Sigma(\cdot)\Sigma(\cdot)' \) such that (2.2) is equivalent (in law) to

\[
\begin{align*}
dX(t) &= \tilde{b}(X(t), u_1(t), u_2(t)) dt + \sigma(X(t)) dW(t) - (I_d - P) d\xi(t), \\
d\xi(t) &= I\{X(t) \in \partial D\} d\xi(t), \quad t \geq 0, \\
X(0) &= x_0 \in \bar{D}, \quad \xi(0) = 0.
\end{align*}
\]

Note that in (2.3), the direction of reflection \( \gamma \) does not satisfy the smoothness property. In fact \( \gamma \) is piecewise constant. With the additional assumptions (A1)', the equation (2.3) has a unique weak solution [16].

Remark 2.4. The assumption (A1)'(i) makes the diffusion matrix \( \Sigma \) real-valued. The condition (A1)'(ii) guarantees that \( \Sigma \Sigma' \) is positive definite and the assumption (A1)'(iii) is the heavy traffic condition.

We now describe the zero-sum stochastic differential game. Let \( \bar{r} : \bar{D} \times U_1 \times U_2 \to \mathbb{R} \) be the cost function. When the state of the system is \( x \in \bar{D} \) and player \( i \) chooses his action \( u_i, i = 1, 2 \), player 2 receives a payoff \( \bar{r}(x, u_1, u_2) \) from player 1. Naturally, player 2 tries to maximize the cumulative payoff and player 1 tries to minimize the same. We make the following assumption for the cost function.

(A3) The function \( \bar{r} \) is bounded, continuous, and Lipschitz in its first argument uniformly with respect to the rest.

Define the function \( r : \bar{D} \times V_1 \times V_2 \to \mathbb{R} \) by

\[
r(x, v_1, v_2) = \int_{U_2} \int_{U_1} \bar{r}(x, u_1, u_2)v_1(du_1)v_2(du_2).
\]

The planning horizon is infinite, and we study two different types of payoff criteria: discounted payoff and (long-run) average payoff.

**Discounted Payoff**

Let \( \alpha > 0 \) be the discount factor. Let \((v_1, v_2) \in A_1 \times A_2\). The \( \alpha \)-discounted payoff to player 2 for the initial condition \( x \in \bar{D} \) is defined by

\[
R_\alpha[v_1, v_2](x) := \mathbb{E}\left[ \int_0^\infty e^{-\alpha t} r(X(t), v_1(t), v_2(t)) dt \left| X(0) = x \right. \right].
\]
A strategy \( v^*_1 \in A_1 \) is said to be \( \alpha \)-discounted optimal for player 1 for initial condition \( x \) if

\[
R_\alpha [v^*_1, \tilde{v}_2](x) \leq \sup_{v_2 \in A_2} \inf_{v_1 \in A_1} R_\alpha [v_1, v_2](x) := R_\alpha(x) \tag{2.5}
\]

for any \( \tilde{v}_2 \in A_2 \). The function \( R_\alpha: \overline{D} \to \mathbb{R} \) is called the \( \alpha \)-discounted lower value function of the game. Similarly, a strategy \( v^*_2 \in A_2 \) is said to be \( \alpha \)-discounted optimal for player 2 for initial condition \( x \) if

\[
R_\alpha [\tilde{v}_1, v^*_2](x) \geq \inf_{v_1 \in A_1} \sup_{v_2 \in A_2} R_\alpha [v_1, v_2](x) := \overline{R}_\alpha(x) \tag{2.6}
\]

for any \( \tilde{v}_1 \in A_1 \). The function \( \overline{R}_\alpha: \overline{D} \to \mathbb{R} \) is called the \( \alpha \)-discounted upper value function of the game. If \( \overline{R}_\alpha \equiv R_\alpha \), then the game is said to admit a value for discounted criterion, and the common function is denoted by \( R_\alpha \) and is called an \( \alpha \)-discounted value function. If a Markov strategy \( v_i \in M_i \) is \( \alpha \)-discounted optimal for all initial conditions, then it is said to be \( \alpha \)-discounted optimal for player \( i \). Clearly the existence of a pair of \( \alpha \)-discounted optimal strategies for both players ensures that the \( \alpha \)-discounted value function exists.

**Average Payoff**

Let \( (v_1, v_2) \in A_1 \times A_2 \). Then average payoff to player 1 for the initial condition \( x \) is defined as

\[
L[v_1, v_2](x) := \lim_{T \to \infty} \frac{1}{T} E \left[ \int_0^T r(X(t), v_1(t), v_2(t)) \, dt \bigg| X(0) = x \right]. \tag{2.7}
\]

The definitions of average optimal strategies and average value are similar.

For \( (u_1, u_2) \in U_1 \times U_2 \) and for a suitable function \( f: \overline{D} \to \mathbb{R} \) write

\[
L^{u_1, u_2}[f](x) = \frac{1}{2} \sum_{i,j=1}^d a_{ij}(x) \frac{\partial^2 f(x)}{\partial x_i \partial x_j} + \sum_{i=1}^d b_i(x, u_1, u_2) \frac{\partial f(x)}{\partial x_i}. \tag{2.8}
\]

More generally, for \( (v_1, v_2) \in V_1 \times V_2 \), we write

\[
L^{v_1, v_2}[f](x) = \int_{U_2} \int_{U_1} L^{u_1, u_2}[f](x) v_2(du_2) v_1(du_1). \tag{2.9}
\]

**Remark 2.5.** Note that the domain \( D_m \) has a smooth boundary. It is unbounded, however. In the next section we use results from reflecting diffusions in a smooth bounded domain. To do this we approximate the domain \( D_m \) by an increasing sequence of smooth bounded domains \( D_{mn} \) in
the following manner. Set \( I_{n+1/m} = [0, n + 1/m]^d \). Define for \( n \geq 1 \), \( C_1^n, i = 1, \ldots, d - 1 \), as follows.

\[
C_1^n = \left\{ x \in \mathbb{R}^d \mid 0 \leq x_i \leq n + \frac{1}{m}, 0 \leq x_j \leq \frac{1}{m}, i \neq j, \text{ for } i = 1, \ldots, d \right\}
\]

\[
C_2^n = \left\{ x \in \mathbb{R}^d \mid 0 \leq x_i \leq n + \frac{1}{m}, n \leq x_j \leq n + \frac{1}{m}, 0 \leq x_k \leq \frac{1}{m}, k \neq i, j, \text{ for } i, j = 1, 2, \ldots, d, i \neq j \right\}
\]

The sets \( C_1^n, i \geq 3 \), are defined in a similar fashion. Set for \( x = (x_1, \ldots, x_d) \),

\[
(x^k, x') = (x_1, \ldots, x_{k-1}, x', x_{k+1}, \ldots, x_d)
\]

and

\[
(x^{k_1: k_2}, x^{',''}) = (x_1, \ldots, x_{k_1-1}, x', x_{k_1+1}, \ldots, x_{k_2-1}, x'', x_{k_2+1}, \ldots, x_d).
\]

Now define

\[
E_1^n = \left\{ x \in \mathbb{R}^d \mid 0 < x_i < n + \frac{1}{m}, \left\| x - \left( \left( \frac{1}{m}, \ldots, \frac{1}{m}, x_i \right) \right) \right\| < \frac{1}{m}, i = 1, \ldots, d \right\}
\]

\[
E_2^n = \left\{ x \in \mathbb{R}^d \mid 0 < x_i < n + \frac{1}{m}, \left\| x - \left( \left( \frac{1}{m}, \ldots, \frac{1}{m}, x_i, n \right) \right) \right\| < \frac{1}{m}, i, j = 1, \ldots, d, i \neq j \right\}
\]

The set \( E_1^n, i \geq 3 \), is defined similarly. Define

\[
D_{m_n} = \left( (I_{n+1/m} \cap D_m) \right) \cup_{i=1}^{d-1} C_i^n \cup_{i=1}^{d-1} E_i^n.
\]

Then \( D_{m_n} \) is an increasing sequence of bounded open sets in \( D_m \) with smooth boundary \( \partial D_{m_n} \) such that \( D_{m_n} \uparrow D_m \) as \( n \to \infty \). Moreover, for any open connected set \( I \subseteq \partial D_m \), for sufficiently large \( n \), \( I \subseteq D_{m_n} \). When \( d = 2 \) we have a simple representation for \( D_{m_n} \):

\[
D_{m_n} = \left\{ D_m \cap \left[ 0, n + \frac{1}{m} \right] \times \left[ 0, n + \frac{1}{m} \right] \right\} \setminus \tilde{I}_{n+1/m} \cup B((1, n), \frac{1}{m}) \cup B((n, n), \frac{1}{m}) \cup B((n, 1), \frac{1}{m}).
\]
where

\[ \tilde{I}_{n+1/m} = \left\{ \left[ 0, \frac{1}{m} \right] \times \left[ n, n + \frac{1}{m} \right] \right\} \cup \left\{ \left[ n, n + \frac{1}{m} \right] \times \left[ 0, \frac{1}{m} \right] \right\} \]

In the next sections we consider stochastic differential equations on domains \( \overline{D}_{mn} \). So we need the assumption (A2)(i) to be satisfied for \( \overline{D}_{mn} \) also. This can be achieved by extending \( \gamma \) to the whole of \( \mathbb{R}^d \) such that \( \gamma \) satisfies (A2)(i) on \( \overline{D}_{mn} \). We also consider a stochastic differential game with state evolving according to (2.1) in \( D_m \) (or \( \overline{D}_{mn} \)) with reflection along \( \gamma \). The definitions of the admissible and Markov strategies for these games are analogous to the definitions in Section 2. By an abuse of notation we continue to denote the set of all admissible and Markov strategies for player \( i \) of these games by \( A_i \) and \( M_i \), respectively.

3. DISCOUNTED PAYOFF CRITERION

The value function of a differential game is usually associated with the solution of a nonlinear partial differential equation referred to as an Isaacs equation. The Isaacs equation for an \( \alpha \)-discounted payoff is given by

\[
\alpha \phi(x) = \inf_{v_1 \in V_1} \sup_{v_2 \in V_2} \left[ L^{v_1,v_2} \phi(x) + r(x, v_1, v_2) \right] \quad \text{in } D,
\]

\[
= \sup_{v_2 \in V_2} \inf_{v_1 \in V_1} \left[ L^{v_1,v_2} \phi(x) + r(x, v_1, v_2) \right] \quad \text{in } D, \tag{3.1}
\]

\[
\frac{\partial \phi}{\partial \gamma}(x) = 0 \quad \text{on } \partial D.
\]

Note that in (3.1) \( \inf \sup \) need not be equal to \( \sup \inf \) in general. In our set-up we use a minimax theorem of Fan [6] to ensure the equality. Indeed we show that (3.1) has a unique solution.

**Theorem 3.1.** Assume (A1)–(A3). Then the \( \alpha \)-discounted value function \( R_\alpha(\cdot) \) exists and is the unique bounded solution in \( C^2, \gamma(D) \cap C^1, \gamma(\overline{D}) \) of (3.1) for any \( \gamma \in (0, 1) \).

**Proof.** Fix \( m \geq 1 \). Consider the partial differential equation

\[
\alpha \phi_m(x) = \inf_{v_1 \in V_1} \sup_{v_2 \in V_2} \left[ L^{v_1,v_2} \phi_m(x) + r(x, v_1, v_2) \right] \quad \text{in } D_m,
\]

\[
\frac{\partial \phi_m}{\partial \gamma}(x) = 0 \quad \text{on } \partial D_m. \tag{3.2}
\]
First we show that (3.2) has a solution in an appropriate sense.

Consider the partial differential equation

\[
\alpha \frac{\partial \phi_{mn}}{\partial x}(x) = \inf_{v_1 \in V_1} \sup_{v_2 \in V_2} [L^{v_1, v_2} \phi_{mn}(x) + r(x, v_1, v_2)] \quad \text{in } D_{mn},
\]

(3.3)

\[
\frac{\partial \phi_{mn}}{\partial \gamma}(x) = 0 \quad \text{on } \partial D_{mn}.
\]

Note that (3.3) is the Isaacs equation for the same stochastic differential game with \(D_{mn}\) as the state space and with reflection on \(\partial D_{mn}\) along \(\gamma\). Hence from the result of [7] it follows that (3.3) has a unique solution \(\phi_{mn} \in C^2(D_{mn}) \cap C^1(\overline{D_{mn}})\), and \(\phi_{mn}\) is the \(\alpha\)-discounted value function of this game. Thus \(\phi_{mn}\) is given by

\[
\phi_{mn}(x) = \inf_{v_1 \in A_1} \sup_{v_2 \in A_2} E_x \left[ \int_0^\infty e^{-\alpha t} r(X(t), v_1(X(t)), v_2(X(t))) \, dt \right],
\]

(3.4)

where \(X(\cdot)\) is the process given by (2.1) in \(\overline{D_{mn}}\) with reflection along \(\gamma\), corresponding to \((v_1, v_2) \in A_1 \times A_2\) (see Remark 2.5). From (3.4) it is clear that

\[
|\phi_{mn}(x)| \leq \frac{C}{\alpha}, \quad x \in D_{mn},
\]

where \(C > 0\) is a bound for the cost function \(\bar{r}\). Let \(Q_k \subset D_m, k \geq 1\), be an increasing sequence of compact sets such that \(\bigcup_{k=1}^\infty Q_k = D_m\). Let \(N_k\) be a positive integer such that

\[
Q_k \subseteq D_{mn}, \quad \text{for all } n \geq N_k, \ k \geq 1.
\]

Using the estimate from [11, Lemma 1.1, p. 247] there exists a constant \(C_1 > 0\) which depends only on \(C, \alpha, b, \sigma\), and \(\text{dist}(Q_k, D_{mn})\) such that

\[
\|\nabla \phi_{mn}\|_{L^2(Q_k)} \leq C_1, \quad \text{for all } n \geq N_k.
\]

Then using the uniform Lipschitzian assumption in (A1), it follows from (3.3) that

\[
\|\phi_{mn}\|_{W^{2,2}(Q_k)} \leq C_2, \quad \text{for all } n \geq N_k,
\]

where \(C_2 > 0\) is a constant depending on \(C, \alpha, b, \sigma\), and \(\text{dist}(Q_k, D_{mn})\).

We next use the estimates from [8, p. 177], for \(2 \leq p < \infty\), to find a constant \(C_3 > 0\) depending on \(C, \alpha, b, \sigma\), and \(\text{dist}(Q_k, D_{mn})\) such that

\[
\|\phi_{mn}\|_{W^{2, p}(Q_k)} \leq C_3, \quad \text{for all } n \geq N_k.
\]

Since \(W^{2, p}(Q_k)\) is compactly imbedded in \(W^{1, p}(Q_k)\), for each \(k \geq 1\), there exists a subsequence, say \(\{\phi_{mn}\}\), which converges to some \(\phi_m\) in
Now by a routine diagonalization argument, we have, along a suitable subsequence,
\[ \phi_{mn} \to \phi_m \text{ in } W^{1,p}(Q_k) \quad \text{for all } k \geq 1. \]

Hence along a suitable subsequence (for simplicity of notation we denote this by the same sequence)
\[ \phi_{mn} \to \phi_m \text{ in } W^{1,p}_{\text{loc}}(D_m) \text{ as } n \to \infty. \quad (3.5) \]

Hence, using the uniform Lipschitzian condition in (A1), we have
\[
\inf_{v_1 \in V_1} \sup_{v_2 \in V_2} [b(x, v_1, v_2) \cdot \nabla \phi_{mn}(x) + r(x, v_1, v_2)] \\
\to \inf_{v_1 \in V_1} \sup_{v_2 \in V_2} [b(x, v_1, v_2) \cdot \nabla \phi_m(x) + r(x, v_1, v_2)] \\
\quad \in L^p_{\text{loc}}(D_m), \ 2 \leq p < \infty. \quad (3.6)
\]

Now using (3.5) and (3.6), it follows from (3.3), by letting \( n \to \infty \), that
\[
\alpha \phi_m(x) = \inf_{v_1 \in V_1} \sup_{v_2 \in V_2} [L^{v_1, v_2} \phi_m(x) + r(x, v_1, v_2)] \quad \text{in } D \quad (3.7)
\]

in the sense of distribution and \( \phi_m \in W^{1,p}_{\text{loc}}(D_m) \). Using (A1)(iii), it follows from elliptic regularity results [8, p. 175] that \( \phi_m \in W^{1,p}_{\text{loc}}(D_m) \cap W^{2,1}_{\text{loc}}(\overline{D}_m), \ 2 \leq p < \infty \). Clearly
\[
|\phi_m(x)| \leq \frac{C}{\alpha}, \quad x \in \overline{D}_m.
\]

Let \( \widetilde{Q}_k \subseteq D_k \), \( k \geq 1 \) be an increasing sequence of compact sets such that \( \bigcup_{k=1}^{\infty} \widetilde{Q}_k = D \) and let \( R_k \) be a positive integer such that
\[
\widetilde{Q}_k \subseteq D_m \cap B(0, R) \quad \text{for all } R \geq R_k, \ m \geq R_k, \ k \geq 1.
\]

Then using the arguments as above, for \( 2 \leq p < \infty \), we can find a constant \( C_4 \) depending only on \( C, \alpha, b, \sigma \), and \( \text{dist}(\overline{Q}_k, D_{R_k} \cap B(0, R_k)) \) such that
\[
\|\phi_m\|_{W^{2, r}(\widetilde{Q}_k)} \leq C_4 \quad \text{for all } m \geq R_k.
\]

Now repeating the arguments described above, we can show that along a subsequence
\[
\phi_m \to \phi \quad \text{in } W^{1,p}_{\text{loc}}(D),
\]

and
\[
\alpha \phi(x) = \inf_{v_1 \in V_1} \sup_{v_2 \in V_2} [L^{v_1, v_2} \phi(x) + r(x, v_1, v_2)] \quad \text{in } D \quad (3.8)
\]
in the sense of distribution and \( \phi \in W^{1,p}_{\text{loc}}(D) \). As before by [8, p. 175], \( \phi \in W^{2,p}_{\text{loc}}(D), 2 \leq p < \infty \). Hence by the Sobolev imbedding theorem [1, pp. 97–98] \( \phi \in C^{2,\gamma}(D) \cap C^{1,\gamma}(\bar{D}) \) for any \( \gamma \in (0, 1) \).

Now we show that \( \partial \phi / \partial \gamma = 0 \) on \( \partial D \). Let \( O \subseteq \partial D \) be open bounded. Set \( O_m = O \cap \partial D_m \), then \( O_m \uparrow O \). Fix \( m \geq 1 \). Since \( D_m \subseteq D_k \) for all \( k \geq m \), we can see that \( \phi_k, \phi \in W^{1,p}_{\text{loc}}(D_m) \), for all \( k \geq m, p \geq 2 \). Using (A2)(i) and the trace result from [9, pp. 63–64], we have

\[
\frac{\partial \phi_k}{\partial \gamma}, \frac{\partial \phi}{\partial \gamma} \in W^{1-1/p, p}_{\text{loc}}(\partial D_{mn}), \quad \text{for all } n \geq 1, k \geq m.
\]

By the continuity of the trace map, \( \psi \rightarrow \partial \psi / \partial \gamma \) from \( W^{1,p}_{\text{loc}}(D_{mn}) \rightarrow W^{1-1/p, p}_{\text{loc}}(\partial D_{mn}) \), we have for a fixed \( m \),

\[
\frac{\partial \phi_k}{\partial \gamma} \rightarrow \frac{\partial \phi}{\partial \gamma} \quad \text{in } W^{1-1/p, p}_{\text{loc}}(\partial D_{mn}) \text{ as } k \rightarrow \infty, \quad \text{for all } n \geq 1.
\]

Since \( O_m \subset \partial D_{mn} \) is bounded, for \( n \) large enough (say, \( n_0 \)), \( O_m \subset \partial D_{mn_0} \). Therefore \( \partial \phi_k / \partial \gamma = 0 \) on \( O_m \). Since

\[
\frac{\partial \phi_k}{\partial \gamma} \rightarrow \frac{\partial \phi}{\partial \gamma} \quad \text{in } W^{1-1/p, p}_{\text{loc}}(\partial D_{mn}) \text{ as } k \rightarrow \infty,
\]

we have \( \partial \phi / \partial \gamma = 0 \) a.e. in \( O_m \). Therefore,

\[
\frac{\partial \phi}{\partial \gamma} = 0 \quad \text{a.e. in } O.
\]

This in turn implies that

\[
\frac{\partial \phi}{\partial \gamma} = 0 \quad \text{a.e. on } \partial D.
\]

Hence \( \phi \) satisfies (3.2). Thus from Fan’s minimax theorem [6], it follows that

\[
\inf_{v_1 \in V_1} \sup_{v_2 \in V_2} [L^{v_1, v_2} \phi(x) + r(x, v_1, v_2)]
\]

\[
= \sup_{v_2 \in V_2} \inf_{v_1 \in V_1} [L^{v_1, v_2} \phi(x) + r(x, v_1, v_2)], \quad \text{a.e. } x \in D.
\]

Hence \( \phi \) is a solution of (3.1). By a measurable selection theorem [4], there exist \( v_1^* \in M_1 \) and \( v_2^* \in M_2 \) such that

\[
\alpha \phi(x) = \sup_{v_2 \in V_2} [L^{v_1^*, v_2} \phi(x) + r(x, v_1^*(x), v_2)] \quad \text{a.e. in } D, \quad (3.9)
\]

\[
\alpha \phi(x) = \inf_{v_1 \in V_1} [L^{v_1, v_2^*} \phi(x) + r(x, v_1, v_2^*(x))] \quad \text{a.e. in } D. \quad (3.10)
\]
Let \( v_2 \in A_2 \) and let \( v_1^* \in M_1 \) be as in (3.9). Let \( X(\cdot) \) be the process given by (2.1) in \( \bar{D} \) corresponding to \( (v_1^*, v_2) \) with the initial condition \( X(0) = x \). Then by standard arguments involving Ito's formula for the function \( e^{-at}\phi(\cdot) \) and for the process \( X(\cdot) \), it follows as in [7] that
\[
\phi(x) \geq E \left[ \int_0^\infty e^{-at}r(X(t), v_1^*(X(t)), v_2(t)) \, dt \right] = R_\alpha[v_1^*, v_2](x), \quad \text{for all } v_2 \in M_2.
\]
Therefore,
\[
\phi(x) \geq \sup_{v_2 \in A_2} R_\alpha[v_1^*, v_2](x).
\]
Similarly,
\[
\phi(x) \leq \inf_{v_1 \in A_1} R_\alpha[v_1, v_2^*](x).
\]
Thus it follows that
\[
\phi(x) = \bar{R}_\alpha(x) = R_\alpha(x), \quad x \in \bar{D}.
\]
Hence the value function \( R_\alpha(\cdot) \) exists and is a bounded solution to (3.1) in the desired class of functions. Now to complete the proof we have to establish the uniqueness of the solution. For this, suppose \( \psi \) to be another bounded solution to (3.1) in the same class of functions. Let \( K > 0 \) be a common bound for \( |\bar{R}_\alpha(\cdot)| \) and \( |\psi(\cdot)| \). Using Ito's formula, it can be shown as in [7] that
\[
|R_\alpha(x) - \psi(x)| \leq 2e^{-at}K, \quad t \geq 0, \ x \in \bar{D}.
\]
Letting \( t \to \infty \) we have \( R_\alpha \equiv \psi \). 

By closely mimicking the arguments in [7] we can prove the following theorem concerning the existence of optimal strategies. We omit the details.

**THEOREM 3.2.** Let \( v_1^* \in M_1 \) be such that
\[
\sup_{v_2 \in \bar{V}_2} \left[ b(x, v_1^*(x), v_2) \cdot \nabla R_\alpha(x) + r(x, v_1^*(x), v_2) \right] = \inf_{v_1 \in V_1} \sup_{v_2 \in \bar{V}_2} \left[ b(x, v_1, v_2) \cdot \nabla R_\alpha(x) + r(x, v_1, v_2) \right] \quad \text{a.e. in } D, \tag{3.11}
\]
Then \( v_1^* \) is \( \alpha \)-discounted optimal for player 1. Similarly, let \( v_2^* \in M_2 \) be such that
\[
\inf_{v_1 \in \bar{V}_1} \left[ b(x, v_1, v_2^*(x)) \cdot \nabla R_\alpha(x) + r(x, v_1, v_2^*(x)) \right] = \sup_{v_2 \in \bar{V}_2} \inf_{v_1 \in \bar{V}_1} \left[ b(x, v_1, v_2) \cdot \nabla R_\alpha(x) + r(x, v_1, v_2) \right] \quad \text{a.e. in } D, \tag{3.12}
\]
Then \( v_2^* \) is \( \alpha \)-discounted optimal for player 2.
Remark 3.1. The existence of the outer minimizing selector $v^*_1 \in M_1$ in (3.11) and the outer maximizing selector $v^*_2 \in M_2$ in (3.12) is guaranteed by a measurable selection theorem [4]. Hence Theorem 3.2 proves the existence of $\alpha$-discounted optimal Markov strategies for both players.

Remark 3.2. From the proof of Theorem 3.1, we can see that, by invoking Fan’s minimax theorem and an appropriate application of Ito’s formula, the $\alpha$-discounted value function of the stochastic game with state given by (2.1) in $\bar{D}_m$ with reflection on $\partial D_m$ along $\gamma$ exists and is the unique bounded solution in $W^{2,p}(\bar{D}_m) \cap W^{1,p}(D_m)$, $2 \leq p < \infty$, of the Isaacs equation

\[
\alpha \phi^m(x) = \inf_{v_1 \in V_1} \sup_{v_2 \in V_2} [L_{v_1,v_2}^m \phi^m(x) + r(x, v_1, v_2)] \quad \text{in } D_m,
\]

\[
\frac{\partial \phi^m}{\partial \gamma}(x) = 0 \quad \text{on } \partial D_m.
\]

This fact will be used in the next section.

4. ERGODIC PAYOFF CRITERION

In the ergodic payoff, the asymptotic behavior of the system plays a crucial role. Hence we make a further assumption to ensure the stability of the system.

(A4) There exist a symmetric positive definite matrix $Q$ and a scalar $\lambda > 0$ such that, for all $x, y \in \bar{D}$, $(u_1, u_2) \in U_1 \times U_2$,

\[
2\tilde{b}(x, y, u_1, u_2) \cdot (Q(x - y)) - \frac{(Q(x - y))'a(x, y)(Q(x - y))}{(x - y)'Q(x - y)} + \text{trace}(a(x, y)Q) \leq -\lambda \|x - y\|^2,
\]

where $a(x, y) = (\sigma(x) - \sigma(y))(\sigma(x) - \sigma(y))'$, and $\tilde{b}(x, y, u_1, u_2) = \tilde{b}(x, u_1, u_2) - \tilde{b}(x, u_1, u_2)$.

First we give a couple of examples where the assumption (A4) is satisfied.

Example 4.1. Let $U_i = [0, 1]^d$, $i = 1, 2$, $b(x, u_1, u_2) = Bx + C_1 u_1 + C_2 u_2$, $\sigma(x) = I_d$, $x \in \mathbb{R}^d$, and $u_i \in U_i$; let $C_1$, $C_2$ be real matrices of order
Hence for \( \lambda > \), \( (A4) \) is satisfied for any \( \lambda \). Let \( U_i = [0, 1]^2 \), \( i = 1, 2 \), and for \( x = (x_1, x_2) \in \mathbb{R}^2 \), let \( u_i = (u_{i1}, u_{i2}) \in U_i \) define
\[
\tilde{b}(x, u_1, u_2) = \begin{pmatrix} \sin x_1 & 0 \\ 0 & \cos x_2 \end{pmatrix} \begin{pmatrix} u_{11} \\ u_{12} \end{pmatrix} + \begin{pmatrix} \sin x_1 & -\cos x_1 \\ \cos x_2 & \sin x_2 \end{pmatrix} \begin{pmatrix} u_{21} \\ u_{22} \end{pmatrix} - \lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},
\]
where \( \lambda \) is a suitable constant. Set \( \sigma(x) = I_d \). Then with \( Q = \frac{1}{2} I_d \), we have
\[
2\tilde{b}(x, y, u_1, u_2) \cdot (Q(x - y)) - \frac{(Q(x - y))' a(x, y)(Q(x - y))}{(x - y)' Q(x - y)} + \text{trace}(a(x, y)Q)
\]
\[
= (x - y)' (B' Q + QB)(x - y)
\]
\[
= (x - y)' (-I_d)(x - y) = -\|x - y\|^2.
\]
Hence \( (A4) \) is satisfied for any \( \lambda \in (0, 1) \).

**Example 4.2.** Let \( U_i = [0, 1]^2 \), \( i = 1, 2 \), and for \( x = (x_1, x_2) \in \mathbb{R}^2 \), let \( u_i = (u_{i1}, u_{i2}) \in U_i \) define
\[
\tilde{b}(x, u_1, u_2) = \begin{pmatrix} \sin x_1 & 0 \\ 0 & \cos x_2 \end{pmatrix} \begin{pmatrix} u_{11} \\ u_{12} \end{pmatrix} + \begin{pmatrix} \sin x_1 & -\cos x_1 \\ \cos x_2 & \sin x_2 \end{pmatrix} \begin{pmatrix} u_{21} \\ u_{22} \end{pmatrix} - \lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},
\]
where \( \lambda \) is a suitable constant. Set \( \sigma(x) = I_d \). Then with \( Q = \frac{1}{2} I_d \), we have
\[
2\tilde{b}(x, y, u_1, u_2) \cdot (Q(x - y)) - \frac{(Q(x - y))' a(x, y)(Q(x - y))}{(x - y)' Q(x - y)} + \text{trace}(a(x, y)Q)
\]
\[
\leq \left( \frac{5}{\lambda} - 1 \right) \|x - y\|.
\]
Hence for \( \lambda > 5 \), \( (A4) \) is satisfied.

**Remark 4.1.** Let \( (v_1, v_2) \in M_1 \times M_2 \). Let \( X(\cdot) \) be the corresponding solution of (2.1). By \( (A1), (A2), \) and \( (A4) \), it follows from [14, pp. 19–21] that \( X(t) \) has a unique invariant measure denoted by \( \eta[v_1, v_2] \). Indeed, using the Lyapunov function \( (x' Q x)^{1/2} \), we can prove the following additional result on the flow governed by (2.1). An analogous result was obtained in [3].

**Lemma 4.1.** Let \( (v_1, v_2) \in A_1 \times A_2 \) be given. For the initial condition \( x \in \mathcal{D}_m \), let \( X(x, \cdot) \) denote the corresponding process given by (2.1) with \( \mathcal{D}_m \) as the state space and the reflecting boundary condition along \( \gamma \) in \( \mathcal{D}_m \). Then there exists positive constants \( C_5, C_6 \) independent of \( (v_1, v_2) \) and \( m \geq 1 \), such that
\[
E \|X(x, t) - X(y, t)\| \leq C_6 e^{-C_5 \|x - y\|}, \quad x, y \in \mathcal{D}_m.
\]
Proof. Let \( w(x) = (x'Qx)^{1/2}, x \in \overline{D}_m \). Let \( \partial_i \) denote the partial derivative with respect to the \( i \)th coordinate. Set
\[
\widehat{L}^{\bar{v}_1, \bar{v}_2} = \sum_{i=1}^{d} (b_i(x, \bar{v}_1, \bar{v}_2) - b_i(y, \bar{v}_1, \bar{v}_2)) \partial_i \\
+ \frac{1}{2} \sum_{i,j=1}^{d} a_{ij}(x, y) \partial_i \partial_j, \quad (\bar{v}_1, \bar{v}_2) \in V_1 \times V_2.
\]

Now using (A4), for \( x \neq y \), we have
\[
\widehat{L}^{\bar{v}_1, \bar{v}_2} w(x - y) \leq -\frac{\lambda}{2} \{(x - y)'Q(x - y)\}^{1/2} \|x - y\|^2,
\]
where \( \lambda \) depends only on \( \lambda > 0 \). Hence for some \( C_3 > 0 \)
\[
\widehat{L}^{\bar{v}_1, \bar{v}_2} w(x - y) \leq -C_3 w(x - y), \quad x \neq y, \ x, y \in \overline{D}_m.
\]

Let \( \tau = \inf \{t \geq 0 \mid X(x, t) = X(y, t) \} \). Using Ito’s formula and the above inequality, we have
\[
Ew(X(x, t \wedge \tau) - X(y, t \wedge \tau)) - w(x - y) \\
\leq -C_3E \int_0^{t \wedge \tau} w(X(x, s) - X(y, s)) \, ds.
\]

Since (2.1) evolving in \( \overline{D}_m \) has a pathwise unique solution (see [11]), we have for \( t \geq \tau, X(x, t) = X(y, t) \) a.s. Hence
\[
Ew(X(x, t) - X(y, t)) \leq w(x - y) - C_3E \int_0^t w(X(x, s) - X(y, s)) \, ds.
\]
Using Grownwall’s lemma, we have
\[
Ew(X(x, t) - X(y, t)) \leq e^{-C_3 t} w(x - y), \quad x, y \in \overline{D}_m.
\]

Since \( Q \) is positive definite, we have
\[
E\|X(x, t) - X(y, t)\| \leq C_6 e^{-C_3 t}\|x - y\|, \quad x, y \in \overline{D}_m,
\]
where \( C_6 > 0 \) depends only on the smallest and largest eigenvalues of \( Q \).

As in the discounted payoff, we prove the existence of the value of the game and the optimal Markov strategies by analytic methods. The Isaacs equations for the average payoff criterion are given by,
\[
\rho = \inf_{\nu_1 \in V_1} \sup_{\nu_2 \in V_2} [L^{\nu_1, \nu_2} \phi(x) + r(x, \nu_1, \nu_2)] \quad \text{in } D,
\]
\[
\rho = \sup_{\nu_2 \in V_2} \inf_{\nu_1 \in V_1} [L^{\nu_1, \nu_2} \phi(x) + r(x, \nu_1, \nu_2)] \quad \text{in } D, \quad (4.1)
\]
\[
\frac{\partial \phi}{\partial \gamma} (x) = 0 \quad \text{on } \partial D,
\]
where $\rho$ is a scalar and $\phi$ is a suitable function. By a solution we mean a pair $(\rho, \phi)$ satisfying (4.1) in an appropriate sense. Let

$$H = \left\{ \phi \in \bigcap_{2 \leq p < \infty} \left( W^{2, p}_{\text{loc}}(D) \cap W^{1, p}_{\text{loc}}(\overline{D}) \right) \mid \|\phi(x)\| \leq C'(1 + \|x\|), \right. \text{for some } C' > 0 \right\}. $$

**Theorem 4.1.** Under assumptions (A1), (A2), (A3), and (A4), the equation (4.1) has a unique solution in $H \times \mathbb{R}$, satisfying $\phi(x_0) = 0, 0 \neq x_0 \in \overline{D}$, fixed.

**Proof.** Consider the partial differential equation

$$\rho = \inf_{v_1 \in \mathcal{V}_1, v_2 \in \mathcal{V}_2} \left\{ \mathcal{L}^{v_1, v_2} \phi_m(x) + r(x, v_1, v_2) \right\} \text{ in } D_m, $$

$$\frac{\partial \phi_m}{\partial \gamma}(x) = 0 \text{ on } \partial \overline{D}_m. $$

Since $x_0 \neq 0$, we choose $m$ large enough to ensure that $x_0 \in \overline{D}_m$. First we prove that (4.2) has a solution in an appropriate class of functions. Let $\phi_m^\alpha(\cdot)$ denote the $\alpha$-discounted value function of the stochastic differential game where the state is given by (2.1) and with $\overline{D}_m$ as the state space with reflecting boundary condition along $\gamma$ on $\overline{D}_m$. Then by Remark 3.2, $\phi_m^\alpha(\cdot)$ is the unique bounded solution in $W^{2, p}_{\text{loc}}(D_m) \cap W^{1, p}_{\text{loc}}(\overline{D}_m)$, $2 \leq p < \infty$ of (3.13). For $x, y \in \overline{D}_m$,

$$|\phi_m^\alpha(x) - \phi_m^\alpha(y)| = \inf_{v_1 \in \mathcal{V}_1, v_2 \in \mathcal{V}_2} \sup \left\{ E \left[ \int_0^\infty e^{-at} r(X(x, t), v_1(t), v_2(t)) \, dt \right] \right\}$$

$$\leq \sup_{v_1 \in \mathcal{V}_1, v_2 \in \mathcal{V}_2} \left\{ E \left[ \int_0^\infty e^{-at} \left| r(X(x, t), v_1(t), v_2(t)) - r(X(y, t), v_1(t), v_2(t)) \right| \, dt \right] \right\}$$

$$\leq C_7 C_6 \int_0^\infty e^{-at} e^{-C_7 t} \|x - y\| \, dt$$

$$= \frac{C_7 C_6}{\alpha + C_5} \|x - y\|, $$

where the last inequality follows by Lemma 4.1 and (A3). Hence

$$|\phi_m^\alpha(x) - \phi_m^\alpha(y)| \leq C_8 \|x - y\|, $$

(4.3)
where $C_8 > 0$ is independent of $\alpha$. Let $\tilde{\phi}_n^\alpha(x) = \phi_n^\alpha(x) - \phi_n^\alpha(x_0)$. Then $\tilde{\phi}_n^\alpha(\cdot)$ is the unique bounded solution in $W^{1,p}_\text{loc}(D_m) \cap W^{1,p}_{\text{loc}}(\overline{D_m})$, $2 \leq p < \infty$, of
\[
\alpha \phi_n(x_0) + \alpha \tilde{\phi}_n(x) = \inf_{v_1, v_2} \sup_{v_1, v_2} \left[ L^{\nu_1, \nu_2} \tilde{\phi}_n^\alpha(x) + r(x, v_1, v_2) \right] \quad \text{in } D_m,
\]
\[
\frac{\partial \tilde{\phi}_n^\alpha}{\partial \gamma}(x) = 0 \quad \text{on } \partial D_m.
\]
(4.4)

Let $Q \subseteq D_m$ be a compact subset. From (4.3) we can see that $\tilde{\phi}_n^\alpha(\cdot)$ is bounded by a constant which depends only on $C_8$ and the diameter($Q$), uniformly in $\alpha \in (0, 1)$. By the estimate from [11, Lemma 1.1, p. 247], there exists a constant $C_9 > 0$ which depends only on $C, b, \sigma, C_8$, diameter($Q$), and dist($Q, D_m$) and is independent of $0 < \alpha < 1$, such that
\[
\|\nabla \tilde{\phi}_n^\alpha\|_{L^\infty(Q)} \leq C_9.
\]

Now using arguments similar to those used in the proof of Theorem 3.1, we can show that, along a suitable sequence $\alpha_n \to 0$,
\[
\tilde{\phi}_n^\alpha \to \phi_m \quad \text{in } W^{1,p}_{\text{loc}}(D_m), \quad 2 \leq p < \infty.
\]

Let $\rho_m$ be a limit point of $\alpha_n \tilde{\phi}_n^\alpha(x_0)$. By letting $n \to \infty$ in (4.4), we can show as in the proof of Theorem 3.1 that $(\phi_m, \rho_m) \in W^{1,p}_{\text{loc}}(D_m) \cap W^{1,p}_{\text{loc}}(\overline{D_m}) \times \mathbb{R}$, $2 \leq p < \infty$, satisfying (4.2). Moreover, it is clear that
\[
|\phi_m| \leq C_{11}(1 + \|x\|), \quad x \in \overline{D_m},
\]

where $C_{11} > 0$ is independent of $m$.

Let $Q_k \subseteq D, k \geq 1$, be an increasing sequence of compact sets so that $\bigcup_{k=1}^\infty Q_k = D$. Let $R_k$ be a positive integer such that
\[
Q_k \subseteq D_m \cap B(0, R) \quad \text{for all } R \geq R_k, m \geq R_k, \quad k \geq 1.
\]

Now repeating the arguments as in the proof of Theorem 3.1, there exists a constant $C_{12} > 0$ which depends only on $C, b, \sigma$, and dist($Q_k, D \cap B(0, R_k)$) such that
\[
\|\phi_m\|_{W^{1,\nu}(Q_k)} \leq C_{12}, \quad \text{for all } m \geq R_k.
\]

Also, $|\rho_m| \leq C$, where $C > 0$ is a bound for $\tilde{\rho}$. Hence along a suitable subsequence,
\[
(\phi_m, \rho_m) \to (\phi, \rho^*) \quad \text{in } W^{1,p}(D) \times \mathbb{R}.
\]

Now by a routine diagonalization argument as in the proof of Theorem 3.1, we can show that along a suitable subsequence,
\[
(\phi_m, \rho_m) \to (\phi, \rho^*) \quad \text{in } W^{1,p}_{\text{loc}}(D) \times \mathbb{R}. \quad (4.5)
\]
Now using arguments as in Theorem 3.1, we can show that \((\phi, \rho^*)\) is a solution to (4.1) in \(W^{1,\infty}_0(D) \cap W^{1,\infty}_0(\overline{D}) \times \mathbb{R}\). Clearly,
\[
|\phi(x)| \leq C_1(1 + \|x\|).
\]

Let \(v_1^* \in M_1\) and \(v_2^* \in M_2\) be such that
\[
\rho^* = \inf_{v_1 \in V_1} \{ L^{v_1}(x, v_1, v_2) \phi(x) + r(x, v_1, v_2(x)) \} \quad \text{a.e. (4.6)}
\]
\[
\rho^* = \sup_{v_2 \in V_2} \{ L^{v_2}(x, v_1, v_2) \phi(x) + r(x, v_1, v_2(x)) \} \quad \text{a.e. (4.7)}
\]

Let \(v_2 \in A_2\) and let \(v_1^* \in M_1\) be as in (4.6). Let \(X(\cdot)\) be the process given by (2.1) in \(\overline{D}\) corresponding to \((v_1^*, v_2)\) with the initial condition \(X(0) = x\). Then by standard arguments involving Ito’s formula for the function \(e^{\alpha t}\phi(\cdot)\) and for the process \(X(\cdot)\), as in [7], it follows that
\[
\rho^* \geq \lim\inf_{T \to \infty} E_x \left[ \int_0^T r(X(t), v_1^*(X(t)), v_2(t)) \, dt \right]
\]
\[
= \rho[v_1^*, v_2] \quad \text{for all } v_2 \in M_2.
\]

Therefore,
\[
\rho^* \geq \sup_{v_2 \in A_2} \rho[v_1^*, v_2].
\]

Similarly,
\[
\rho^* \leq \inf_{v_1 \in A_1} \rho[v_1, v_2^*].
\]

Thus, \(\rho^* = \rho[v_1^*, v_2^*]\). Moreover,
\[
\rho^* = \tilde{\rho} = \underline{\rho}.
\]

Hence \(\rho^*\) is the value of the game. The uniqueness of \(\phi\) can be obtained by closely mimicking the arguments in [5]. We omit the details. \(\square\)

Finally, using Theorem 4.1, we can obtain the analog of Theorem 3.2. We again omit the details.

**Remark 4.2.** As before, by the Sobolev imbedding theorem, \(\phi \in C^{2, \gamma}(D) \cap C^{1, \gamma}(\overline{D})\) for any \(\gamma \in (0, 1)\).

**Theorem 4.2.** Let \((\phi, \rho^*)\) be the unique solution to (4.1). Let \(v_1^* \in M_1\) be such that
\[
\sup_{v_2 \in V_2} \{ b(x, v_1^*(x), v_2) \cdot \nabla \phi(x) + r(x, v_1^*(x), v_2) \}
\]
\[
= \inf_{v_1 \in V_1} \sup_{v_2 \in V_2} \{ b(x, v_1, v_2) \cdot \nabla \phi(x) + r(x, v_1, v_2) \} \quad \text{a.e. in } D. \quad (4.8)
\]
Then \( v_1^* \) is \( \alpha \)-discounted optimal for player 1. Similarly, let \( v_2^* \in M_2 \) be such that

\[
\inf_{v_1 \in V_1} \left[ b(x, v_1, v_2^*(x)) \cdot \nabla \phi(x) + r(x, v_1, v_2^*(x)) \right] = \sup_{v_2 \in V_2} \inf_{v_1 \in V_1} \left[ b(x, v_1, v_2) \cdot \nabla \phi(x) + r(x, v_1, v_2) \right] \quad \text{a.e. in } D. \quad (4.9)
\]

Then \( v_2^* \) is \( \alpha \)-discounted optimal for player 2.

5. AN APPLICATION TO A QUEUING MODEL

We first consider a simplified version of the model discussed in Section 1. Consider a heavy traffic queue which is the limit of a sequence of queues with one server. Let the intensity of the arrival process be

\[
\lambda_n(x, u_1, u_2) = \sqrt{n\lambda x u_1} + n\lambda, \quad x \geq 0,
\]

where \( \lambda, \bar{\lambda} \) are suitable positive constants and \( u_1 \) denotes the action of the user (player 1) from the action space \([0, 1]\). The intensity of the service time process is

\[
\mu_n(x, u_1, u_2) = \sqrt{n\bar{\mu} x u_2} + n\mu, \quad x \geq 0,
\]

where \( \mu, \bar{\mu} \) are suitable positive constants and \( u_2 \) denotes the action of the superuser (player 2) from the action space \([0, 1]\). The state of the queue is governed by the stochastic differential equation

\[
\begin{align*}
\dot{X}(t) &= (\bar{\lambda} u_1(t) - \bar{\mu} u_2(t))X(t) \, dt + (\sqrt{\bar{\lambda}}, \sqrt{\bar{\mu}}) \cdot dW(t) + d\xi(t), \\
\dot{\xi}(t) &= I\{X(t) = 0\} \, d\xi(t), \quad t \geq 0, \\
X(0) &= x, \quad \xi(0) = 0, \quad x \geq 0,
\end{align*}
\]

(5.1)

where \( W = (W_1, W_2) \) is a standard 2-dimensional Wiener process and \( u_i(\cdot) \) is a \( \mathcal{F}([0, 1]) \) valued nonanticipative process, which is the admissible strategy for player \( i \). The cost function is given by

\[
\bar{r}(x, u_1, u_2) = \gamma u_1 - \theta u_2 - x^2, \quad \gamma, \theta > 0,
\]

where \( \gamma, \theta \) are suitable positive constants described in Section 1. For this payoff function and the discount factor \( \alpha > 0 \), the \( \alpha \)-discounted Isaacs
equation is given by
\[ \alpha \phi(x) = \inf_{v_1 \in \mathcal{D}([0,1])} \sup_{v_2 \in \mathcal{D}([0,1])} \left[ (\tilde{\lambda} x v_1 - \tilde{\mu} x v_2) \phi'(x) - x^2 + \gamma v_1 - \theta v_2 \right] \\
+ \frac{1}{2} (\lambda + \mu) \phi''(x), \quad x > 0, \]
\[ = \sup_{v_1 \in \mathcal{D}([0,1])} \inf_{v_2 \in \mathcal{D}([0,1])} \left[ (\tilde{\lambda} x v_1 - \tilde{\mu} x v_2) \phi'(x) - x^2 + \gamma v_1 - \theta v_2 \right] \quad (5.2) \\
+ \frac{1}{2} (\lambda + \mu) \phi''(x), \quad x > 0, \]
\[ \phi'(0) = 0. \]

Clearly the assumptions (A1), (A1)′ are satisfied. Note that the function \( \tilde{\phi} \) is unbounded. We can modify the arguments in Section 3, however, to show that the \( \alpha \)-discounted value function \( R_\alpha \) is the unique solution of (5.2) in the class of functions belonging to \( C^2(0, \infty) \cap C^1[0, \infty) \) with quadratic growth condition. Note that (5.2) can be written as
\[ \alpha \phi(x) = \inf_{v_1 \in \mathcal{D}([0,1])} \left[ (\tilde{\lambda} x v_1 + \gamma v_1) + \sup_{v_2 \in \mathcal{D}([0,1])} \left[ -\theta v_2 - \tilde{\mu} x v_2 \phi'(x) \right] \right] \\
- x^2 + \frac{1}{2} (\lambda + \mu) \phi''(x), \quad x > 0, \quad (5.3) \]
\[ \phi'(0) = 0. \]

Since \( \mathcal{D}([0,1]) \) is convex and compact, the infimum and the supremum are attained at the extremum points. Hence (5.3) becomes
\[ \alpha \phi(x) = \inf_{u_1 \in [0,1]} \left[ (\tilde{\lambda} x u_1 + \gamma u_1) + \sup_{u_2 \in [0,1]} \left[ -\theta u_2 - \tilde{\mu} x u_2 \phi'(x) \right] \right] \\
- x^2 + \frac{1}{2} (\lambda + \mu) \phi''(x), \quad x > 0, \quad (5.4) \]
\[ \phi'(0) = 0. \]

Then by Theorem 3.2, the optimal strategy for player 1 is given by
\[ v_1^*(x) = \begin{cases} 
1 & \text{if } (\tilde{\lambda} x R_\alpha^*(x) + \gamma) < 0 \\
p & \text{if } (\tilde{\lambda} x R_\alpha^*(x) + \gamma) = 0 \\
0 & \text{if } (\tilde{\lambda} x R_\alpha^*(x) + \gamma) > 0, 
\end{cases} \quad (5.5) \]
where \( p \) is any value in \([0,1]\). Similarly, the optimal strategy for player 2 is
\[ v_2^*(x) = \begin{cases} 
1 & \text{if } (\tilde{\mu} x R_\alpha^*(x) + \theta) < 0 \\
p & \text{if } (\tilde{\mu} x R_\alpha^*(x) + \theta) = 0 \\
0 & \text{if } (\tilde{\mu} x R_\alpha^*(x) + \theta) > 0. 
\end{cases} \quad (5.6) \]

Thus both players have nonrandomized Markov strategies, and the structure of the optimal strategies of both players are explicitly determined in
terms of the value function. For the ergodic payoff criterion, we can obtain analogous results. For the queueing network with $d$ servers as described in the Introduction, we can derive analogous results. It is difficult, however, to get the explicit structure of the optimal strategies as in (5.5), (5.6).

6. CONCLUSIONS

In this paper, we study an abstract stochastic game problem where the state is given by the controlled reflecting diffusion in the nonnegative orthrant. This abstract problem generalizes a differential game problem arising in communication networks with heavy traffic. For the abstract problem, we establish the existence of the value and optimal Markov strategies for both players. For the network problem, we establish the existence of nonrandomized optimal Markov strategies for both players.

REFERENCES