Numerical Solutions for Stochastic Differential Games With Regime Switching

Qingshuo Song, G. George Yin, *Fellow, IEEE*, and Zhimin Zhang

***Abstract*—This paper is concerned with numerical methods for stochastic differential games of regime-switching diffusions. Numerical methods using Markov chain approximation techniques are developed. A new proof of the existence of a saddle point for the stochastic differential game is provided. This new proof enables us to treat certain systems with nonseparable (in controls) structure. Convergence of the algorithms is derived by means of weak convergence methods. In addition, examples are also provided for demonstration purposes.**

***Index Terms*—Controlled regime-switching diffusion, convergence, Markov chain approximation, saddle point, stochastic differential game.**

I. Introduction

THIS PAPER is concerned with numerical solutions for stochastic differential games with regime switching. The merge of differential games and regime-switching models stems from a wide range of applications in network problems, complex systems, and financial engineering. Many problems arising in, for example, pursuit-evasion games, queueing systems in heavy traffic, risk-sensitive control, and constrained optimization problems, can be formulated as two-player stochastic differential games [1], [7], [8]. In most of these problems, closed-form solutions are difficult to obtain. As a viable alternative, one is contended with numerical approximations [16]. While the convergence of approximations with various forms of continuous-state and continuous-time dynamic games was developed, a systematic approach of numerical approximation for stochastic models was provided in [12]. Part of the difficulties in dealing with such game problems is that the existence of the value of a game is rather difficult to establish. Naturally, one may proceed with discrete strategies to approximate its continuous counterpart. In fact, using limits of discrete strategies to define the upper and lower values of games can be traced back to [4] and [5]. The classical definition [7] does not use discretization, but many applications based on discretization are shown to yield the same values as in these references. Using a Markov chain approximation technique, numerical methods are developed in [12] via a probabilistic approach that require that both the objective function and the drift of the diffusion be separable with respect to the controls. It would be nice to be able to relax the separability. Along another line, recent applications in finance and in wireless communications demand the consideration of systems that better describe the random environment. One such modeling point is the use of the so-called regime-switching models (see [3], [6], [20], [21], [25], [26] and also [2], [9], [14], [15] and references therein). In light of these developments, it is natural to consider stochastic differential games in which both the drift and the diffusion terms are modulated by another random process.

In this paper, we aim to develop numerical methods for approximation of stochastic differential games for switching diffusions. Our contributions in this paper include: 1) first, compared with the numerical methods developed in the literature, where only diffusion models were treated, we consider a regime-switching model that includes both continuous dynamics and discrete events. The modulating stochastic process is assumed to be a continuous-time Markov chain representing the random environment and other random factors not included in the usual diffusion formulation; 2) compared to the most recent development in [12], we expand the applicability of the numerical methods to include dynamics and objective functions being not necessarily separable with respect to controls; 3) using Markov chain approximation techniques, we develop numerical schemes to solve the underlying game problems. In contrast to [12], we prove the existence of the saddle point for a game problem with a finite state space without using delayed actions. In the proof of existence of saddle points, we assume that either the functions are separable or that the cost is of convex–concave type, and the controls in the drift are bilinear with an additional constant term (to be specified later). The essence is to be able to interchange the supremum and infimum by using the results of von Neumann [19] and Sion [17], which enable us to treat certain nonseparable functions. In addition, our formulation allows us to handle regime-switching systems; 4) We prove the convergence of the numerical schemes, and provide numerical experimental results for the stochastic differential game problems.

The rest of the paper is arranged as follows. Section II begins with the formulation of the problem, in which we give definitions of the differential game problems, as well as values and saddle points of a game. Moreover, the notation of the relaxed control is recalled briefly. Section III presents Markov chain

Manuscript received November 11, 2005; revised November 15, 2006. Recommended by Associate Editor A. Lim. The work of Q. S. Song was supported in part by the U.S. Army Research Office Multidisciplinary University Research Initiative (MURI) under Grant W911NF-06-1-0094 at the University of Southern California and in part by Wayne State University Research Enhancement Program. The work of G. Yin was supported in part by the National Science Foundation under Grant DMS-0603287 and Grant DMS-0624849, and in part by the National Security Agency under Grant MSPF-068-029. The work of Z. Zhang was supported in part by the National Science Foundation under Grant DMS-0311807 and Grant DMS-0612908, and in part by Michigan Life Science Corridor.

Q. S. Song is with the Department of Mathematics, University of Southern California, Los Angeles, CA 90089-1113, USA (e-mail: qingshus@usc.edu).

G. Yin and Z. Zhang are with the Department of Mathematics, Wayne State University, Detroit, MI 48202, USA (e-mail: gyin@math.wayne.edu; zzhang@math.wayne.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2007.915169

approximation procedures. Section IV proceeds with the convergence of the procedure. It establishes a weak convergence result for any constructed approximating Markov chain as long as it is locally consistent. Section V proves the existence of a saddle point of the game either for systems having separable forms or for functions being convex–concave type. As was alluded to, the proof is interesting in its own right for people working in control and optimization in general, and in game theory in particular. In Section VI, we present a numerical example for demonstration purposes. Section VII issues some further remarks. Finally, an Appendix is provided, collecting the proofs of results.

II. Formulation

Consider a two-player stochastic game of regime-switching diffusions. For a finite set $\mathcal{M} = \{1, \ldots, m_0\}$, $x \in \mathbb{R}^{l_0}$, $b(\cdot,\cdot,\cdot,\cdot) : \mathbb{R}^{l_0} \times \mathcal{M} \times \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}^{l_0}$, $\sigma(\cdot,\cdot) : \mathbb{R}^{l_0} \times \mathcal{M} \mapsto \mathbb{R}^{l_0} \times \mathbb{R}^{l_0}$, the dynamic system is given by

$$x(t) = x(0) + \int_0^t b(x(s), \alpha(s), u_1(s), u_2(s))ds + \int_0^t \sigma(x(s), \alpha(s))\, dw(s) \tag{1}$$

where, for each $i = 1, 2$, $u_i(\cdot)$ is a control for player i, $w(\cdot)$ is a standard $\mathbb{R}^{l_0}$-valued Brownian motion, and $\alpha(\cdot)$ is a continuous-time Markov chain having state space $\mathcal{M}$ with generator $Q = (q_{\iota,\ell}) \in \mathbb{R}^{m_0 \times m_0}$ and state space $\mathcal{M}$. Let $\{\mathcal{F}_t : 0 \le t\}$ be a filtration, which might depend on controls, and which measures at least $\{(w(s), \alpha(s)) : s \le t\}$. We suppose that, for each $i = 1, 2$, $u_i(\cdot)$ is $\mathcal{F}_t$-adapted and takes values in a compact subset $U_i \subset \mathbb{R}$, which is called *admissible control*. Let $A(x, \iota) = \sigma(x, \iota)\sigma'(x, \iota) = (a_{j_0 k_0}(x, \iota)) \in \mathbb{R}^{l_0} \times \mathbb{R}^{l_0}$, which is symmetric and positive definite.

Note that the Markov chain is used to represent the possible regimes of the environment. For any controls with the correct information structures, there will be a filtration with respect to which $w(\cdot)$ is a standard vector-valued Brownian motion and $\alpha(\cdot)$ is a continuous-time Markov chain, and to which the controls are adapted [12].

Let $G \subset \mathbb{R}^{l_0}$ be a compact set, that is, the closure of its interior G^0 and τ be the first exit time of $x(t)$ from G^0 with

$$\tau = \min\{t : x(t) \notin G^o\}. \tag{2}$$

Using a real number $\beta > 0$ to denote the discount factor, let the cost function be

$$W(x, \iota, u_1, u_2) = E^u_{x,\iota}\left[\int_0^\tau e^{-\beta s}\widetilde{k}(x(s), \alpha(s), u_1(s), u_2(s))ds + \widetilde{g}(x(\tau), \alpha(\tau))\right] \tag{3}$$

where $\widetilde{k}(\cdot)$ and $\widetilde{g}(\cdot)$ are appropriate functions representing the running cost and the terminal cost, respectively, and $E^u_{x,\iota}$ denotes the expectation taken with the initial data $x(0) = x$ and $\alpha(0) = \iota$, and the given control process $u(\cdot) = (u_1(\cdot), u_2(\cdot))$ used. To facilitate the proof of weak convergence, we introduce the relaxed control representation [12], [13].

Definition 2.1: Let $\mathcal{B}(U \times [0,\infty))$ be the σ-algebra of Borel subsets of $U \times [0,\infty)$. An *admissible deterministic relaxed control* or simply a *relaxed control* $m(\cdot)$ is a measure on $\mathcal{B}(U \times [0,\infty))$ such that $m(U \times [0,t]) = t$ for each $t \ge 0$. Given a relaxed control $m(\cdot)$, there is an $m_t(\cdot)$ such that $m(drdt) = m_t(dr)dt$. In fact, we can define $m_t(B) = \lim_{\delta\to 0} \frac{m(B\times[t-\delta,t])}{\delta}$ for $B \in \mathcal{B}(U)$, a Borel subset of U.

Let the probability space $(\Omega, \mathcal{F}, P)$ be given together with a filtration $\mathcal{F}_t$, an $\mathcal{F}_t$-Brownian motion $w(\cdot)$, and an $\mathcal{F}_t$-switching process $\alpha(\cdot)$. We say $m(\cdot)$ is an admissible relaxed stochastic control for $(w(\cdot), \alpha(\cdot))$ or $(m(\cdot), w(\cdot), \alpha(\cdot))$ is admissible, if $m(\cdot, \omega)$ is a deterministic relaxed control with probability one and if $m(B \times [0,t])$ is $\mathcal{F}_t$-adapted for all $B \in \mathcal{B}(U)$. There is a derivative $m_t(\cdot)$ such that $m_t(\cdot)$ is $\mathcal{F}_t$-adapted for all $B \in \mathcal{B}(U)$ and that $m(B) = \int_{U\times[0;\infty)} I_{\{(u,t)\in B\}} m_t(dc)dt$ for all $B \in \mathcal{B}(U \times [0,\infty))$ w.p.1.

Using the notion of relaxed controls, the dynamic system (1) and the cost function (3) can be rewritten as

$$x(t) = x(0) + \int_0^t \int_{U_1\times U_2} b(x(s), \alpha(s), r_1, r_2) \times m_s(dr_1 \times dr_2)ds + \int_0^t \sigma(x(s), \alpha(s))dw(s) \tag{4}$$

$$W(x, \iota, m) = E^m_{x,\iota}\left[\int_0^\tau e^{-\beta s}\int_{U_1\times U_2} \widetilde{k}(x(s), \alpha(s), r_1, r_2)m_s \times (dr_1 \times dr_2)ds + \widetilde{g}(x(\tau), \alpha(\tau))\right]. \tag{5}$$

To proceed, we need the following assumptions.

A1) For each $\iota \in \mathcal{M}$, $\widetilde{k}(\cdot, \iota, \cdot, \cdot)$ and $b(\cdot, \iota, \cdot, \cdot)$ are continuous functions on the compact set $G \times U_1 \times U_2$.

A2) For each $\iota \in \mathcal{M}$, the functions $\sigma(\cdot, \iota)$ and $\widetilde{g}(\cdot, \iota)$ are continuous on G.

A3) Equation (4) has a unique weak sense solution (i.e., unique in the sense of distribution) for each admissible triple $(w(\cdot), \alpha(\cdot), m(\cdot))$, where $m(\cdot) = (m_1(\cdot), m_2(\cdot))$.

A4) For any $\iota \in \mathcal{M}$, $j_0, k_0 \in \{1, 2, \ldots, l_0\}, j_0 \ne k_0$, $a_{j_0 j_0}(x, \iota) > \sum_{k_0 \ne j_0} |a_{j_0 k_0}(x, \iota)|$.

A5) Let

$$\hat{\tau}(\phi) = \begin{cases} \infty, \text{ if } \phi(t) \in G^0 & \text{for all } \ t < \infty \\ \inf\{t : \phi(t) \notin G^0\}, & \text{otherwise.}\end{cases}$$

The function $\hat{\tau}(\cdot)$ is continuous as a mapping from $D[0,\infty)$ to $[0,\infty]$ with probability 1 relative to the measure induced by any solution to (4) with an initial condition (x, ι), where $D[0,\infty)$ denotes the space of functions that are right continuous and have left limits endowed with the Skorohod topology, and $[0,\infty]$ is the interval $[0,\infty)$ compactified (see [13, p. 259]).

A6) The functions $b(\cdot)$ and $\widetilde{k}(\cdot)$ are separable in r_1 and r_2 for every $(x, \iota) \in G \times \mathcal{M}$. That is, $b(x, \iota, r_1, r_2) = \sum_{i=1}^2 b^i(x, \iota, r_i)$ and $\widetilde{k}(x, \iota, r_1, r_2) = \sum_{i=1}^2 \widetilde{k}^i(x, \iota, r_i)$.

A7) There exist $\mathbb{R}^{l_0}$-valued continuous functions $b^i(x,\iota)$ $(i=0,1,2)$ and $c(x,\iota)$ such that $b(x,\iota,r_1,r_2) = r_1 r_2 b^0(x,\iota) + r_1 b^1(x,\iota) + r_2 b^2(x,\iota) + c(x,\iota)$. The cost $\widetilde{k}(\cdot)$ is convex–concave with respect to (r_1,r_2). That is, $\widetilde{k}(\cdot)$ is convex with respect to r_1 for every (x,ι,r_2) and concave with respect to r_2 for every (x,ι,r_1).

Assumption A4) is used for the construction of transition probabilities of the approximating Markov chain. It requires that the diffusion matrix be diagonally dominated. If the given dynamic system does not satisfy A4), then we can adjust the coordinate system to satisfy assumption A4) (see [13, p. 110]). Assumption A5) is a broad condition that is satisfied in most applications. The main purpose is to avoid the *tangency* problem as discussed in [p. 278]. In Section V, we will establish the existence of *saddle points* using either A6) or A7) in addition to A1)–A5). Assumption A7) requires the drift to be bilinear with an addition of the constant term $c(x,\iota)$ with respect to control variables. As can be seen, in this case, the function is nonseparable for the controls. Assumption A7) is a sufficient condition to have convex–concave property of transition probability used in approximating the Markov chain by a central difference scheme, and it is an important condition that leads to the existence of saddle points.

Now, we are ready to define upper values, lower values, and *saddle points* of differential games. In a zero-sum differential game, player 1 wants to minimize and player 2 wants to maximize the objective function (3). If, at each time, player 1 goes first and player 2 goes last, then the value of the objective function is said to be an upper value. If the order of the actions of the two players is reversed, the corresponding value is termed a lower value. Nevertheless, since the real time elapsed is continuous, the order of actions for a given instance is not transparent. Therefore, we use a piecewise constant control space on one player who goes first, denoted by $\mathcal{L}_i(\Delta)$ (see the definition given later, and the corresponding part of [12] without regime switching).

Let $\mathcal{U}_i = \{u_i(\cdot) :$ admissible ordinary control with respect to $\mathcal{F}_t\}$. For $\Delta > 0$, let $\mathcal{U}_i(\Delta) \subset \mathcal{U}_i$ such that $u_i(\cdot)$ are piecewise constant on the intervals $[k\Delta, k\Delta+\Delta), k = 0,1,2,\ldots,$ and $u_i(k\Delta)$ is $\mathcal{F}_{k\Delta}$-measurable. Let B be a Borel subset of U_1. Let $\mathcal{L}_1(\Delta)$supset$\mathcal{U}_1(\Delta)$ denote the set of such piecewise constant controls for player 1 that are represented by functions $Q_{1,n}(B,\cdot)$, $n = 0,1,\ldots,$ of the conditional probability type

$$P\{u_1(n\Delta) \in B | w(s), \alpha(s), u_2(s), s < n\Delta; u_1(k\Delta), k < n\}$$
$$= Q_{1,n}(B; w(s), \alpha(s), u(s), s < n\Delta) \quad (6)$$

where $Q_{1,n}(B;\cdot)$ is a measurable function for $B \in \mathcal{B}(U_1)$. Controls determined by (6) may be called strategies, owing to their explicit dependence on the past actions of both players. If a rule for player 1 is given by (6), in the arguments of the cost functions, we will sometimes write it as $u_1(u_2)$ to emphasize its dependence on u_2. Define $\mathcal{L}_2(\Delta)$ and the associated rule $u_2(u_1)$ for player 2 analogously.

Note that $\mathcal{L}_i(\Delta)$ differs from $\mathcal{U}_i(\Delta)$ in that the control in $\mathcal{L}_i(\Delta)$ is determined by a series of measurable functions in (6). For example, for $u_1(\cdot) \in \mathcal{L}_1(\Delta)$, it depends on the past $u_2(\cdot) \in \mathcal{U}_2$, where u_2 is normally not in $\mathcal{U}_2(\Delta)$. However, the uniqueness assumption A3) implies that it is only the probability law of $(w(\cdot), \alpha(\cdot), u(\cdot))$, or more generally of $(w(\cdot), \alpha(\cdot), m(\cdot))$, which determines the law of the solution, and hence, the value of the cost. Thus, we can always suppose that, if the control of, for example, Player 1 is determined by a form such as (6), then (in relaxed control terminology) the law of $(w(\cdot), \alpha(\cdot), m_2(\cdot))$ is determined recursively by a conditional probability law given by

$$P(\{w(t), \alpha(t), m_2(t), k\Delta \le t < k\Delta + \Delta\} \in \cdot$$
$$| w(s), \alpha(s), m_2(s), s < t, m_1(s), s \le k\Delta). \quad (7)$$

Now, we are in a position to introduce the definition of *upper and lower values*.

Definition 2.2: For initial condition $x(0) = x, \alpha(0) = \iota$, define the upper and lower values for the game as

$$V^+(x,\iota) = \lim_{\Delta\to 0} \inf_{u_1 \in \mathcal{L}_1(\Delta)} \sup_{u_2 \in \mathcal{U}_2} W(x,\iota,u_1(u_2),u_2) \quad (8)$$

$$V^-(x,\iota) = \lim_{\Delta\to 0} \sup_{u_2 \in \mathcal{L}_2(\Delta)} \inf_{u_1 \in \mathcal{U}_1} W(x,\iota,u_1,u_2(u_1)). \quad (9)$$

If the lower value and the upper value are equal, then we say that there exists a saddle point for the game, and its value is

$$V^+(x,\iota) = V^-(x,\iota) = V(x,\iota)\, \forall x \in G, \quad \iota \in \mathcal{M}. \quad (10)$$

Remark 2.3: Let us give the interpretation of (8). For a fixed $\Delta > 0$, consider the right-hand side of (8). For each k, at time $k\Delta$, player 1 uses a rule of the form (6) with a constant action taken on $[k\Delta, k\Delta+\Delta)$. That is, player 1 "goes first." Accordingly, player 2 selects its strategy simply to be admissible. This operation yields an admissible $u(\cdot) = (u_1(\cdot), u_2(\cdot))$. As $\Delta \to 0$, the $\inf\sup$ is monotonically decreasing since player 1 can make decisions more often. The analogous comments hold for (9).

III. Markov Chain Approximation Methods

The Markov chain approximation methods for stochastic control problems without regime switching were considered in [11] and [13]. The Markov chain approximations were also used in numerical solutions for stochastic games without regime switching [12]. Using the Markov chain approximation method for stochastic control with regime switching was considered in [18]. Here, we will construct a two-component Markov chain. The approximation is of finite difference type. The basis of the approximation is a discrete-time, finite-state, controlled Markov chain $\{(\xi_n^h, \alpha_n^h) : n < \infty\}$ whose properties are *locally consistent* with that of (1). In addition, the Markov chain in this section is designed for computational purposes. An alternative approximation will be given in Section V for establishing the existence of saddle points.

For each $h > 0$, let G_h be a finite subset of G such that $d(G_h, G) \to 0$ as $h \to 0$, where $d(\cdot)$ is a metric defined by $d(G_h, G) = \max_{p\in G} \min_{q \in G_h} d(p,q)$. Let $\{(\xi_n^h, \alpha_n^h) : n < \infty\}$ be a controlled discrete-time Markov chain on a discrete state space $G_h \times \mathcal{M}$ with transition probabilities denoted by $p^h((x,\iota),(y,\ell)|r)$, where $r = (r_1, r_2) \in U_1 \times U_2$. We use

$(u^h_{1,n}, u^h_{2,n})$ to denote the actual control action for the chain at discrete time n. Suppose that we have a positive function $\Delta t^h(\cdot)$ on $G_h \times \mathcal{M} \times U_1 \times U_2$ such that $\sup_{x,\iota,r} \Delta t^h(x,\iota,r) \to 0$ as $h \to 0$, but $\inf_{x,\iota,r} \Delta t^h(x,\iota,r) > 0$ for each $h > 0$. We take an interpolation of the discrete Markov chain $\{(\xi^h_n, \alpha^h_n)\}$ by using the interpolation interval $\Delta t^h_n = \Delta t^h_n(\xi^h_n, \alpha^h_n, u^h_{1,n}, u^h_{2,n})$. Now, we give the definition of local consistency.

Definition 3.1: Let $\{p^h((x,\iota),(y,\ell)|r)\}$, for $(x,\iota),(y,\ell) \in G_h \times \mathcal{M}$ and $r \in U_1 \times U_2$, be a collection of well-defined transition probabilities for the two-component Markov chain $\{(\xi^h_n, \alpha^h_n)\}$, approximation to $(x(\cdot),\alpha(\cdot))$. Define the difference $\Delta\xi^h_n = \xi^h_{n+1} - \xi^h_n$. Assume that $\lim_{h\to 0}\sup_{x,\iota,r}\Delta t^h(x,\iota,r) = 0$. Denote by $E^{r,h}_{x,\iota,n}$, $\mathrm{cov}^{r,h}_{x,\iota,n}$, and $p^{r,h}_{x,\iota,n}$ the conditional expectation, covariance, and probability, given $\{\xi^h_k, \alpha^h_k, u^h_{1,k}, u^h_{2,k}, k \le n, \xi^h_n = x, \alpha^h_n = \iota, (u^h_{1,n}, u^h_{2,n}) = r\}$. The sequence $\{(\xi^h_n, \alpha^h_n)\}$ is said to be *locally consistent* with (1), if

$$
\begin{aligned}
&E^{r,h}_{x,\iota,n}\Delta\xi^h_n = b(x,\iota,r)\Delta t^h(x,\iota,r) + o(\Delta t^h(x,\iota,r))\\
&\mathrm{cov}^{r,h}_{x,\iota,n}\Delta\xi^h_n = A(x,\iota)\Delta t^h(x,\iota,r) + o(\Delta t^h(x,\iota,r))\\
&p^{r,h}_{x,\iota,n}\{\alpha^h_{n+1} = \ell\} = \Delta t^h(x,\iota,r)q_{\iota\ell} + o(\Delta t^h(x,\iota,r)), \quad \text{for } \ell \ne \iota\\
&p^{r,h}_{x,\iota,n}\{\alpha^h_{n+1} = \iota\} = \Delta t^h(x,\iota,r)(1 + q_{\iota\iota}) + o(\Delta t^h(x,\iota,r))\\
&\sup_{n,\omega\in\Omega}|\Delta\xi^h_n| \to 0 \quad \text{as } h \to 0.
\end{aligned}
\tag{11}
$$

To approximate the expected cost given in (3), we define a cost function using the aforementioned Markov chain. Let

$$
t^h_n = \sum_{j=0}^{n-1}\Delta t^h_j \quad \text{and} \quad N_h = \inf\{n : \xi^h_n \notin G^o_h\}.
$$

Then, define the cost function with respect to $u^h = \{(u^h_{1,n}, u^h_{2,n}) : n = 1,2,\ldots,N_h\}$ as

$$
W^h(x,\iota,u^h) = E_{x,\iota}\left[\sum_{n=0}^{N_h-1} e^{-\beta t^h_n}\widetilde{k}(\xi^h_n, \alpha^h_n, u^h_{1,n}, u^h_{2,n})\Delta t^h_n + \widetilde{g}(\xi^h_{N_h}, \alpha^h_{N_h})\right] \tag{12}
$$

where $(\xi^h_0, \alpha^h_0) = (x,\iota)$.

To approximate the value function in (8) and (9) in the discrete case with parameter h, we need to define admissible controls with respect to $\mathcal{F}_{k\Delta}$ (i.e., the σ-algebra $\mathcal{F}_t$ with $t = k\Delta$), so that the value can be computed. Using $\mathcal{U}^h_i(1)$ to denote the space of the ordinary controls that Player i goes first, and its strategy is defined by measurable functions of the type similar to (6). That is, for $u^h_i \in \mathcal{U}^h_i(1)$, it can be written by the conditional probability law

$$
P\{u^h_{i,n} \in \cdot \,|\, \xi^h_k, \alpha^h_k, k \le n; u^h_{1,k}, u^h_{2,k}, k < n\}. \tag{13}
$$

Using $\mathcal{U}^h_i(2)$ to denote the collection of the ordinary controls that player i goes last, the strategy is defined by the conditional probability law

$$
P\{u^h_{i,n} \in \cdot|\xi^h_k, \alpha^h_k, k \le n; u^h_{i,k}, k < n; u^h_{j,k}, k \le n, j \ne i\}. \tag{14}
$$

Now, we are ready to define the upper and lower values using $W^h(\cdot)$, $\mathcal{U}^h_i(1)$, and $\mathcal{U}^h_j(2)$ as

$$
V^{h,+}(x,\iota) = \inf_{u^h_1 \in \mathcal{U}^h_1(1)} \sup_{u^h_2 \in \mathcal{U}^h_2(2)} W^h(x,\iota,u^h_1,u^h_2) \tag{15}
$$

$$
V^{h,-}(x,\iota) = \sup_{u^h_2 \in \mathcal{U}^h_2(1)} \inf_{u^h_1 \in \mathcal{U}^h_1(2)} W^h(x,\iota,u^h_1,u^h_2) \tag{16}
$$

respectively. Note that $V^{h,+}$ is not necessarily decreasing as $h \to 0$. Similarly, $V^{h,-}$ is not necessarily increasing as $h \to 0$. Furthermore, the corresponding $V^{h,+}(\cdot)$ and $V^{h,-}(\cdot)$ are not necessarily equal to each other.

The corresponding dynamic programming equations are

$$
V^{h,+}(x,\iota) = \min_{r_1\in U_1}\{\max_{r_2\in U_2} E_{x,\iota}[e^{-\beta\Delta t^h(x,\iota,r)}V^{h,+}(\xi^h_1, \alpha^h_1) + \widetilde{k}(x,\iota,r)\Delta t^h(x,\iota,r)]\} \tag{17}
$$

$$
V^{h,-}(x,\iota) = \max_{r_2\in U_2}\{\min_{r_1\in U_1} E_{x,\iota}[e^{-\beta\Delta t^h(x,\iota,r)}V^{h,-}(\xi^h_1, \alpha^h_1) + \widetilde{k}(x,\iota,r)\Delta t^h(x,\iota,r)]\} \tag{18}
$$

where $(\xi^h_0, \alpha^h_0) = (x,\iota)$. Owing to the contraction implied by the discounting, there is a unique solution to both (17) and (18).

Theoretically, once we have locally consistent Markov chain $\{(\xi^h_n, \alpha^h_n)\}$, we can compute the cost $W^h(x,\iota,u^h)$ for any admissible $u^h = (u^h_1, u^h_2)$ by (12). Then, $V^{h,+}(x,\iota)$ and $V^{h,-}(x,\iota)$ could be found by (15) and (16). Practically, we can find $V^{h,+}(\cdot)$ and $V^{h,-}(\cdot)$ by solving (17) and (18) using an iteration method. This is possible owing to the following lemma.

Lemma 3.2: For any initial value $\{V^{h,+}_0(x,\iota) : x \in G_h, \iota \in \mathcal{M}\}$, the sequence

$$
V^{h,+}_{n+1}(x,\iota) = \min_{r_1\in U_1}\{\max_{r_2\in U_2} E_{x,\iota}[e^{-\beta\Delta t^h(x,\iota,r)}V^{h,+}_n(\xi^h_1, \alpha^h_1) + \widetilde{k}(x,\iota,r)\Delta t^h(x,\iota,r)]\} \tag{19}
$$

converges to $V^{h,+}(x,\iota)$, the unique solution to (17) as $n \to \infty$. Analogously, for any initial value $\{V^{h,-}_0(x,\iota) : x \in G_h, \iota \in \mathcal{M}\}$, the sequence

$$
V^{h,-}_{n+1}(x,\iota) = \max_{r_2\in U_2}\{\min_{r_1\in U_1} E_{x,\iota}[e^{-\beta\Delta t^h(x,\iota,r)}V^{h,-}_n(\xi^h_1, \alpha^h_1) + \widetilde{k}(x,\iota,r)\Delta t^h(x,\iota,r)]\} \tag{20}
$$

converges to $V^{h,-}(x,\iota)$, the unique solution to (18) as n goes to infinity.

Proof: The proof can be obtained in a way similar to that of [10, Lemma 2]. □

We proceed to specify the transition probabilities of the Markov chain $\{(\xi^h_n, \alpha^h_n)\}$. There might be many alternative Markov chains satisfying local consistency. We generalize the method introduced in [13] to construct a Markov chain for the switching diffusions. In this procedure, we compare the dynamic programming equation with the discrete Hamiltonian–Jacobi–Issacs equation (HJI) to get a Markov chain. Note that the subsequent results in this paper depend only on the final form of the constructed Markov chain and its properties, not on the HJI equations.

We begin with a special case, in which the control space has only one admissible feedback control $u^h(\cdot)$. In this case, the minimum and the maximum in (17) and (18) can be ignored, and the lower value and the upper value are exactly the same as $W^h(x,\iota,u^h)$, denoted by $V^h(x,\iota)$ for simplicity. Approximating $e^{-\beta\Delta t^h(x,\iota,r)}$ by $1-\beta\Delta t^h(x,\iota,r)$, we rewrite (17) and (18) as

$$V^h(x,\iota) = \widetilde{k}(x,\iota,r)\Delta t^h(x,\iota,r) + (1-\beta\Delta t^h(x,\iota,r)) \sum_{(y,\ell)\in G_h\times\mathcal{M}} p^h((x,\iota)(y,\ell)|r)V^h(y,\ell). \quad (21)$$

For fixed $r\in U$, define the differential operator for (1) as

$$L^r\phi(x,\iota) = \phi_x' b(x,\iota,r) + \frac{1}{2}\text{tr}(\phi_{xx}(x,\iota)A(x,\iota)) + Q\phi(x,\cdot)(\iota) \quad (22)$$

where $\iota\in\mathcal{M}$ and $Q\phi(x,\cdot)(\iota) = \sum_{\ell\neq\iota} q_{\iota\ell}(\phi(x,\ell)-\phi(x,\iota)) = \sum_{\ell=1}^{m_0} q_{\iota\ell}\phi(x,\ell)$. Formally, $V(x,\iota)$ satisfies the system of equations

$$\begin{aligned} 0 &= \inf_{r_1\in U_1}\sup_{r_2\in U_2}(L^rV(x,\iota) - \beta V(x,\iota) + \widetilde{k}(x,\iota,r)) \\ &= \sup_{r_2\in U_2}\inf_{r_1\in U_1}(L^rV(x,\iota) - \beta V(x,\iota) + \widetilde{k}(x,\iota,r)). \end{aligned} \quad (23)$$

Assume temporarily that $u(\cdot)$ is the only admissible control in $\mathcal{U}_1\times\mathcal{U}_2$. Drop inf sup, we obtain

$$\begin{aligned} &L^rV(x,\iota) - \beta V(x,\iota) + \widetilde{k}(x,\iota,r) = 0,\ x\in G^0 \\ &V(x,\iota) = \widetilde{g}(x,\iota),\ x\in\partial G. \end{aligned} \quad (24)$$

Let $b_{j_0}(\cdot)$ be the j_0th component of $b(\cdot)$, $V_{x_{j_0}} = (\partial/\partial x_{j_0})V$, and $V_{x_{j_0}x_{k_0}} = (\partial^2/\partial x_{j_0}\partial x_{k_0})V$. Then,

$$\sum_{j_0=1}^{l_0} V_{x_{j_0}}(x,\iota)b_{j_0}(x,\iota,r) + \frac{1}{2}\sum_{j_0,k_0=1}^{l_0} V_{x_{j_0}x_{k_0}}(x,\iota)a_{j_0k_0}(x,\iota) + \sum_{\ell=1}^{m_0} q_{\iota\ell}V(x,\ell) - \beta V(x,\iota) + \widetilde{k}(x,\iota,r) = 0. \quad (25)$$

Denote the standard unit vector by e_{j_0} for $j_0 = 1,\ldots,l_0$. Using the modified upwind finite-difference method, (25) discretizes to

$$\begin{aligned} &V(x,\iota)\to V^h(x,\iota) \\ &V_{x_{j_0}}(x,\iota)\to\frac{V^h(x+e_{j_0}h,\iota)-V^h(x,\iota)}{h},\ \text{if } b_{j_0}(x,\iota,r)\geq 0 \\ &V_{x_{j_0}}(x,\iota)\to\frac{V^h(x,\iota)-V^h(x-e_{j_0}h,\iota)}{h},\ \text{if } b_{j_0}(x,\iota,r)< 0 \\ &V_{x_{j_0}x_{j_0}}(x,\iota)\to[V^h(x+e_{j_0}h,\iota)-2V^h(x,\iota) \\ &\quad + V^h(x-e_{j_0}h,\iota)]/h^2 \\ &V_{x_{j_0}x_{k_0}}(x,\iota)\to[2V^h(x,\iota)+V^h(x+e_{j_0}h+e_{k_0}h,\iota) \\ &\quad + V^h(x-e_{j_0}h-e_{k_0}h,\iota) - V^h(x+e_{j_0}h,\iota) \\ &\quad - V^h(x-e_{j_0}h,\iota) - V^h(x+e_{k_0}h,\iota) \\ &\quad + V^h(x-e_{k_0}h,\iota)]/(2h^2),\quad \text{if } j_0\neq k_0, a_{j_0k_0}(x,\iota)\geq 0 \\ &V_{x_{j_0}x_{k_0}}(x,\iota)\to[-2V^h(x,\iota) - V^h(x+e_{j_0}h-e_{k_0}h,\iota) \\ &\quad - V^h(x-e_{j_0}h+e_{k_0}h,\iota) + V^h(x+e_{j_0}h,\iota) \\ &\quad + V^h(x-e_{j_0}h,\iota) + V^h(x+e_{k_0}h,\iota) \\ &\quad + V^h(x-e_{k_0}h,\iota)]/(2h^2),\quad \text{if } j_0\neq k_0, a_{j_0k_0}(x,\iota)< 0. \end{aligned} \quad (26)$$

Let $b^+_{j_0}(x,\iota,r) = \max\{b_{j_0}(x,\iota,r),0\}$ and $b^-_{j_0}(x,\iota,r) = \max\{-b_{j_0}(x,\iota,r),0\}$. Analogously, define $a^+_{j_0k_0}(\cdot)$ and $a^-_{j_0k_0}(\cdot)$. Note that $A(x,\iota)$ is a symmetric matrix. Let

$$D^h(x,\iota,r_1,r_2) = \sum_{j_0=1}^{l_0}[(|b_{j_0}(x,\iota,r)|h) + a_{j_0j_0}(x,\iota)] - \sum_{j_0<k_0}|a_{j_0k_0}(x,\iota)| - q_{\iota\iota}h^2 + \beta h^2$$

and set

$$\Delta t^h(x,\iota,r_1,r_2) = h^2/D^h(x,\iota,r_1,r_2). \quad (27)$$

Comparing the upwind difference version of (25) to (21), we obtain the transition probabilities as

$$\begin{aligned} &p^h((x,\iota),(x\pm e_{j_0}h,\iota)|r) = \\ &\frac{hb^{\pm}_{j_0}(x,\iota,r) + (1/2)a_{j_0j_0}(x,\iota) - (1/2)\sum_{k_0\neq j_0}|a_{j_0k_0}(x,\iota)|}{D^h(x,\iota,r_1,r_2)-\beta h^2}, \\ &\qquad j_0 = 1,2,\ldots,l_0 \\ &p^h((x,\iota),(x+e_{j_0}h+e_{k_0}h,\iota)|r) = \frac{(1/2)a^+_{j_0k_0}(x,\iota)}{D^h(x,\iota,r_1,r_2)-\beta h^2} \\ &p^h((x,\iota),(x-e_{j_0}h-e_{k_0}h,\iota)|r) = \frac{(1/2)a^+_{j_0k_0}(x,\iota)}{D^h(x,\iota,r_1,r_2)-\beta h^2}, \\ &\qquad j_0<k_0 \\ &p^h((x,\iota),(x+e_{j_0}h-e_{k_0}h,\iota)|r) = \frac{(1/2)a^-_{j_0k_0}(x,\iota)}{D^h(x,\iota,r_1,r_2)-\beta h^2}, \\ &\qquad j_0\neq k_0 \\ &p^h((x,\iota),(x,\ell)|r) = \frac{q_{\iota\ell}h^2}{D^h(x,\iota,r_1,r_2)-\beta h^2},\ \ell\neq\iota \\ &p^h((x,\iota),(y,\ell)|r) = 0,\quad \text{otherwise}. \end{aligned} \quad (28)$$

Assumption A4) guarantees $D^h(\cdot)-\beta h^2>0$ and $p^h(\cdot)\geq 0$. Also, $\sum_{(y,\ell)}p^h((x,\iota),(y,\ell)|r) = 1$ for any (x,ι). So, the transition probabilities are well defined. Using $\Delta t^h\cdot$ given in (27) and by taking the interpolation of the Markov chain constructed with transition probability (28), we obtain a continuous-time process. Next, we verify the local consistency for our approximation sequences. The proof of the following lemma is relegated to the Appendix.

Lemma 3.3: Assume A1), A2), and A4). The Markov chain (ξ^h_n,α^h_n), together with transition probabilities $\{p^h(\cdot)\}$ and the interpolation function $\Delta t^h(\cdot)$ defined in (28) and (27), is locally consistent with (1).

IV. Convergence of Approximating Markov Chains Under Local Consistency

In this section, we prove a convergence result. For any approximating Markov chain, we show that the upper and lower values $V^{h,+}(\cdot)$ in (15) and $V^{h,-}(\cdot)$ in (16) converge to $V^+(\cdot)$ and $V^-(\cdot)$, respectively, as long as the approximating Markov chain is locally consistent (Definition 3.1). The essence is a careful use of "ε-optimal" control to approximate its value function.

Lemma 4.1: Suppose that the two-component Markov chain $\{\xi_n^h, \alpha_n^h\}$ is locally consistent in the sense of Definition 3.1. Then, as $h \to 0$, the interpolated process $\alpha^h(\cdot)$ of the Markov chain $\{\alpha_n^h\}$ converges weakly to $\alpha(\cdot)$, a Markov chain with generator $Q = (q_{\iota\ell})$.

Proof: The proof can be obtained in a way similar to that of [24, Theorem 3.1]. □

To proceed, we take a continuous-time interpolation of the locally consistent Markov chain. Let $(\xi^h(\cdot), \alpha^h(\cdot))$ be the piecewise constant interpolation of (ξ_n^h, α_n^h) and $u^h(\cdot) = (u_1^h(\cdot), u_2^h(\cdot))$ be the interpolation of $u_n^h = (u_{1,n}^h, u_{2,n}^h)$ with an interpolation interval $\Delta t_n^h(\xi_n^h, \alpha_n^h, u_{1,n}^h, u_{2,n}^h)$. Let $\tau_h = t_{N_h}^h$. In this way, $\xi^h(\cdot)$ has a representation close to (1) by [13, Sec. 10.4.1]. That is,

$$\xi^h(t) = x + \int_0^t b(\xi^h(s), \alpha^h(s), u^h(s))dt + \int_0^t \sigma(\xi^h(s), \alpha^h(s))dw^h(s) + \varepsilon^h(t) \quad (29)$$

where $\xi^h(t) \in G_h$ and $\varepsilon^h(\cdot)$ is negligible in that $\lim_h \sup_{x,\iota,u^h} \sup_{s\le T} E|\varepsilon^h(s)|^2 = 0$. The process $w^h(\cdot)$, a martingale with respect to the filtration induced by $(\xi^h(\cdot), \alpha^h(\cdot), u^h(\cdot), w^h(\cdot))$, converges weakly to a standard vector-valued Brownian motion. All processes in (29) are constants over the intervals $[t_n^h, t_{n+1}^h)$. Let $m^h(\cdot)$ be the relaxed control representation of u^h. Accordingly, the cost function can be written as

$$W^h(x,\alpha,m^h) = E_{x,\alpha}^{m^h}\left[\int_0^{\tau_h} e^{-\beta s}\int_U \widetilde{k}(\xi^h(s), \alpha^h(s), r)m_s^h(dr)ds + \widetilde{g}(\xi^h(\tau_h), \alpha^h(\tau_h))\right] + \varepsilon^h. \quad (30)$$

The proof of the following theorem is in the Appendix.

Theorem 4.2: Assume A1), A2), and A3). Let the approximating chain $\{\xi_n^h, \alpha_n^h, n < \infty\}$ be locally consistent in the sense of Definition 3.1, $\{u_n^h, n<\infty\}$ be the sequence of admissible controls, $\xi^h(\cdot), \alpha^h(\cdot), u^h(\cdot)$ be the corresponding continuous-time interpolations, and $m^h = (m_1^h, m_2^h)$ be the relaxed control representation of $u^h = (u_1^h, u_2^h)$ for $\xi^h(\cdot)$. Let $\mathcal{F}_t^h$ denote the minimal σ-algebra that measures $\{\xi^h(s), \alpha^h(s), m_s^h(\cdot), w^h(s), s \le t\}$ and $\{\tilde{\tau}_h\}$ be a sequence of $\mathcal{F}_t^h$-stopping times. Then, $\{\xi^h(\cdot), \alpha^h(\cdot), m^h(\cdot), w^h(\cdot), \tilde{\tau}_h\}$ is tight. Denote the limit of a weakly convergent subsequence by $(x(\cdot), \alpha(\cdot), m(\cdot), w(\cdot), \tilde{\tau})$ and denote by $\mathcal{F}_t$ the σ-algebra generated by $\{x(s), \alpha(s), m(s), w(s), s\le t, \tilde{\tau}I_{\{\tilde{\tau}\le t\}}\}$. Then, $w(\cdot)$ is a standard $\mathcal{F}_t$-Brownian motion, $\alpha(\cdot)$ is a continuous-time Markov chain generated by Q, $\tilde{\tau}$ is an $\mathcal{F}_t$-stopping time, and $m(\cdot)$ is an admissible control. Moreover, (4) is satisfied.

The next theorem presents the convergence of the numerical method. Its proof is also given in the Appendix.

Theorem 4.3: Assume A1)–A5). For $x \in G_h$, $\iota \in \mathcal{M}$, let $\{(\xi_n^h, \alpha_n^h)\}$ be a locally consistent Markov chain. $V^{h,+}(x,\iota), V^{h,-}(x,\iota)$, $V^+(x,\iota)$, and $V^-(x,\iota)$ are defined in (8), (9), (15), and (16). Then, we have $\lim_{h\to 0} V^{h,+}(x,\iota) = V^+(x,\iota)$ and $\lim_{h\to 0} V^{h,-}(x,\iota) = V^-(x,\iota)$.

V. Existence of Saddle Points

In this section, we show that there exists a saddle point for the stochastic differential game [i.e., $V^+(\cdot) = V^-(\cdot)$] under assumptions A1)–A5) together with either A6) or A7). The proof is carried out in the following steps.

1) Design a Markov chain $\{(\xi_n^h, \alpha_n^h)\}$ in which $\Delta t^h(\cdot)$ and the denominator of $p^h(\cdot)$ do not depend on controls.
2) Show that the constructed Markov chain is locally consistent (see Lemma 5.1), and thus, $V^{h,+}(\cdot) \to V^+(\cdot)$ and $V^{h,-}(\cdot) \to V^-(\cdot)$ by virtue of Theorem 4.3.
3) Using the earlier Markov chain, we consider $V^{h,+}(\cdot)$ and $V^{h,-}(\cdot)$ as (15) and (16) in a piecewise constant ordinary control space. Under assumption A6) or A7), we prove that $V^{h,+}(\cdot) = V^{h,-}(\cdot)$ for all h.

Note that the aforementioned Step 1 enables us to interchange inf and sup for an appropriate function. To begin, we need to design a locally consistent Markov chain different from the setup of (27) and (28). This Markov chain is generated by a central finite-difference scheme and is presented for analysis purpose. Let $D^h(x,\iota) = \sum_{j_0=1}^{l_0} a_{j_0 j_0}(x,\iota) - \sum_{j_0<k_0} |a_{j_0k_0}(x,\iota)| - q_{\iota\iota}h^2 + \beta h^2$. Set the interpolation interval as

$$\Delta t^h(x,\iota) = \frac{h^2}{D^h(x,\iota)}. \quad (31)$$

The transition probabilities are

$$\begin{aligned}
p^h((x,\iota),(x\pm e_{j_0}h,\iota)|r) &= [\pm(1/2)hb_{j_0}(x,\iota,r) + (1/2)a_{j_0j_0}(x,\iota) - (1/2)\sum_{k_0\ne j_0}|a_{j_0k_0}(x,\iota)|] \\
&\quad /[D^h(x,\iota)-\beta h^2], \quad j_0 = 1,2,\dots,l_0 \\
p^h((x,\iota),(x+e_{j_0}h+e_{k_0}h,\iota)|r) &= \frac{(1/2)a_{j_0k_0}^+(x,\iota)}{D^h(x,\iota)-\beta h^2} \\
p^h((x,\iota),(x-e_{j_0}h-e_{k_0}h,\iota)|r) &= \frac{(1/2)a_{j_0k_0}^+(x,\iota)}{D^h(x,\iota)-\beta h^2}, \quad j_0<k_0 \\
p^h((x,\iota),(x+e_{j_0}h-e_{k_0}h,\iota)|r) &= \frac{1/2\cdot a_{j_0k_0}^-(x,\iota)}{D^h(x,\iota)-\beta h^2}, \quad j_0\ne k_0 \\
p^h((x,\iota),(x,\ell)|r) &= \frac{q_{\iota\ell}h^2}{D^h(x,\iota)-\beta h^2}, \quad \ell\ne\iota \\
p^h((x,\iota),(y,\ell)|r) &= 0, \quad \text{otherwise.}
\end{aligned} \quad (32)$$

By A4), $D^h(x,\iota) - \beta h^2 > 0$. Also, we have $\sum_{(y,\ell)} p^h((x,\iota),(y,\ell)|r) = 1$. To ensure that $p^h(\cdot)$ is always nonnegative, we require

$$h \le \frac{\min_{j_0}\{a_{j_0 j_0}(x,\iota) - \sum_{k_0 \neq j_0} |a_{j_0 k_0}(x,\iota)|\}}{\max_r |b_{j_0}(x,\iota,r_1,r_2)|}. \tag{33}$$

Lemma 5.1: Assume A1), A2), A4), and h satisfy (33). The Markov chain (ξ_n^h, α_n^h) with transition probabilities $\{p^h(\cdot)\}$ and the interpolation $\Delta t^h(\cdot)$ defined in (31) and (32) are locally consistent with (1).

Proof: The proof can be obtained in a way similar to that of Lemma 3.3, and is thus, omitted. □

Next, we present a minimax principle in the static game. It is a generalization of von Neumann's result of [19] for convex–concave functions that was obtained by Sion [17].

Lemma 5.2: Let M_1 and M_2 be compact spaces, $\phi(\cdot,\cdot)$ be a continuous function on $M_1 \times M_2$ such that $\phi(\cdot,\nu)$ is convex in M_1 for every $\nu \in M_2$, and that $\phi(\mu,\cdot)$ is concave in M_2 for every $\mu \in M_1$. Then, $\inf_{\mu\in M_1}\sup_{\nu\in M_2}\phi(\mu,\nu) = \sup_{\nu\in M_2}\inf_{\mu\in M_1}\phi(\mu,\nu)$.

Proof: This is a special case of [17, Th. 4.2]. □

Theorem 5.3: Assume A1)–A5), either A6) or A7), and G_h is a finite set defined as in the paragraph above Definition 3.1. For $x \in G_h$ and $\iota \in \mathcal{M}$, a Markov chain is defined by (31) and (32). Let $V^{h,+}(x,\iota)$ and $V^{h,-}(x,\iota)$ be the associated upper and lower values defined in (15) and (16) in the control spaces $\mathcal{U}_i^h(1)$ and $\mathcal{U}_j^h(2)$ of (13) and (14). Then, there always exists a saddle point

$$V^{h,+}(x,\iota) = V^{h,-}(x,\iota) \tag{34}$$

provided h satisfies (33).

Remark 5.4: Theorem 5.3 is one of our main results, and the proof works for every $\beta \ge 0$ under appropriate assumptions. If we set the Markov chain game in the relaxed control space, there always exists a saddle point without assumptions A6) and A7), since the system is linear with respect to the relaxed controls. In this case, the existence of the saddle point can be obtained by using von Neumann's result [19] together with similar techniques, as the proof of Theorem 5.3.

Similar to the proof of Theorem 4.3, for the Markov chain constructed Lemma 5.1, we obtain the following convergence result.

Lemma 5.5: Assume that the conditions of Theorem 4.3 are satisfied. Then for the approximating Markov chain, the limits in Theorem 4.3 hold.

The following theorem is a direct consequence of Theorem 4.3 and Lemma 5.5. The proof is omitted.

Theorem 5.6: Assume that the conditions of Theorem 5.3 are satisfied. Then, the differential game has a saddle point in the sense

$$V^+(x,\iota) = V^-(x,\iota). \tag{35}$$

VI. Examples

In this section, we provide a couple of examples for demonstration. The numerical experiments were obtained by using MATLAB on a WinXP machine. Both separable and nonseparable cost functions are considered.

Consider a pursuit-evasion game. Suppose that the pursuer $x_1(\cdot)$ and evader $x_2(\cdot)$ are governed by the stochastic differential equations (see [10])

$$\begin{aligned} dx_1(t) &= b_1(x_1(t),\alpha(t),u_1(t))dt + \sigma_1(x_1(t),\alpha(t))dw_1(t)\\ dx_2(t) &= b_2(x_2(t),\alpha(t),u_2(t))dt + \sigma_2(x_2(t),\alpha(t))dw_2(t) \end{aligned} \tag{36}$$

where $x_i(t) \in \mathbb{R}$, $w_1(\cdot)$, and $w_2(\cdot)$ are mutually independent standard Brownian motions, and the controls $u_1(\cdot)$ and $u_2(\cdot)$ are determined by the strategy of the pursuer and evader, respectively. Set $x(t) = (x_1(t),x_2(t))'$, $w(t) = (w_1(t),w_2(t))'$, $b(\cdot) = (b_1(\cdot),b_2(\cdot))'$, and

$$\sigma(\cdot) = \begin{pmatrix} \sigma_1(\cdot) & 0 \\ 0 & \sigma_2(\cdot) \end{pmatrix}.$$

Rewrite (36) as

$$dx(t) = b(x(t),\alpha(t),u(t))dt + \sigma(x(t),\alpha(t))dw(t). \tag{37}$$

Let G_1 and G_2 be compact subsets of $\mathbb{R}$, $G = G_1 \times G_2$, and $\tau = \min\{t : x(t) \notin G^o\}$ be a stopping time of the game. Denoting $u(\cdot) = \{u_1(\cdot),u_2(\cdot)\}$, we define the cost function as

$$W(x,\iota,u) = E_{x,\iota}^u\left[\int_0^\tau e^{-\beta s}\widetilde{k}(x(s),\alpha(s),u(s))ds + \widetilde{g}(x(\tau),\alpha(\tau))\right]. \tag{38}$$

The upper and lower values $V^+(x,\iota)$ and $V^-(x,\iota)$ are defined as in Definition 2.2.

By solving dynamic programming equations (17) and (18), we obtain $V^{h,+}(\cdot)$ and $V^{h,-}(\cdot)$ separately for any $h > 0$. Using policy iterations, we obtain $V_n^{h,+}(\cdot) \to V^{h,+}(\cdot)$ as $n \to \infty$. The procedure is outlined as follows.

1) Set $n = 0$. For all $(x,\iota) \in G_h \times \mathcal{M}$, take initial policy $u_i^{h,0}(x,\iota) = 0, i = 1,2$;
2) Find an improved value, that is,

$$u^{h,n+1}(x,\iota) = \left(u_1^{h,n+1}(x,\iota), u_2^{h,n+1}(x,\iota)\right).$$

   a) Solve $V_n^{h,+}(x,\iota)$ from

$$V_n^{h,+}(x,\iota) = \sum_{y,\ell} p^h((x,\iota),(y,\ell)|r)e^{-\beta\Delta t^h(x,\iota,r)}V_n^{h,+}(y,\ell) + \widetilde{k}(x,\iota,r)\Delta t^h(x,\iota,r). \tag{39}$$

   b) Find an improved value $u^{h,n+1}(x,\iota)$ by

$$u^{h,n+1}(x,\iota) = \text{argmin}_{r_1\in U_1}\text{argmax}_{r_2\in U_2}\left[\sum_{y,\ell} p^h((x,\iota),(y,\ell)|r)e^{-\beta\Delta t^h(x,\iota,r)}V_n^{h,+}(y,\ell) + \widetilde{k}(x,\iota,r)\Delta t^h(x,\iota,r)\right]. \tag{40}$$

TABLE I
UPPER AND LOWER VALUES FOR SELECTED INITIAL STATES WITH $\widetilde{k}(x,\iota,r) = r_1^2 + r_2^2 + (x_1+x_2)^2$

step size h	2^{-2}	2^{-3}	2^{-4}
$V^{h,+}(1,1,2)$	0.474087	0.478232	0.479248
$V^{h,-}(1,1,2)$	0.474087	0.478232	0.479248

TABLE II
UPPER AND LOWER VALUES FOR SELECTED INITIAL STATES WITH $\widetilde{k}(x,\iota,r) = r_1 r_2 + (x_1+x_2)^2$

step size h	2^{-2}	2^{-3}	2^{-4}
$V^{h,+}(1,1,2)$	0.240899	0.244391	0.244391
$V^{h,-}(1,1,2)$	0.240523	0.244311	0.244311

3) If $|V_{n+1}^{h,+}(\cdot) - V_n^{h,+}(\cdot)| >$ tolerance, then $n \leftarrow n+1$ and go to step 2.

Similarly, we can also find the approximation of $V^{h,-}(\cdot)$.

Use tolerance 10^{-3}, and let $\mathcal{M} = \{1,2\}$, $G = [0,2]\times[0,2]$, and $U_i = [0,2]$, for $i = 1,2$. The generator of $\alpha(\cdot)$ is

$$Q = \begin{pmatrix} -.5 & .5 \\ .5 & -.5 \end{pmatrix}.$$

For $x_i \in [0,2]$ and $r_i \in [0,2]$, the set $b_i(x_i,\iota,r_i) = (i-1.5)(3-2\iota)(x_i+r_i)$ and $\sigma_i(x_i,\iota) = i\iota$.

First, we consider a separable cost function. For $x \in G$ and $r \in U_1 \times U_2$, the running cost is defined as $\widetilde{k}(x,\iota,r) = r_1^2 + r_2^2 + (x_1+x_2)^2$, and the terminal cost is $\widetilde{g}(x,\iota) = 0$. Table I gives selected computed results with initial value $x = (1,1)$ and $\iota = 2$ for mesh size $h = 2^{-2}$, 2^{-3}, 2^{-4}. The difference between the lower and the upper values are within given tolerance, which confirms our findings in Theorem 5.3 under A6). Next, we consider a nonseparable cost function defined as $\widetilde{k}(x,\iota,r) = r_1 r_2 + (x_1+x_2)^2$, and the terminal cost is $\widetilde{g}(x,\iota) = 0$. Table II gives selected computed results with initial value $x = (1,1)$ and $\iota = 2$ for the mesh size $h = 2^{-2}$, 2^{-3}, 2^{-4}. The difference between the lower and the upper values are within given tolerance, which confirms our findings in Theorem 5.3 under A7). For demonstration, Fig. 1 plots several graphs corresponding to $h = 2^{-4}$.

VII. FURTHER REMARKS

This paper is devoted to numerical methods of stochastic games for regime-switching diffusions. Markov chain approximation techniques are used. In addition to the convergence of the approximating Markov and its associated upper and lower values, a new proof for the existence of saddle points is provided.

For a regime-switching system in which the Markov chain has a large state space, we may use the ideas of the two-time-scale approach presented in [22] (see also [23] and references therein) to first reduce the complexity of the underlying system, and then, construct numerical solutions for the limit systems. Optimal strategies of the limit systems can be used for the construction of strategies of players of the original systems leading to near optimality.

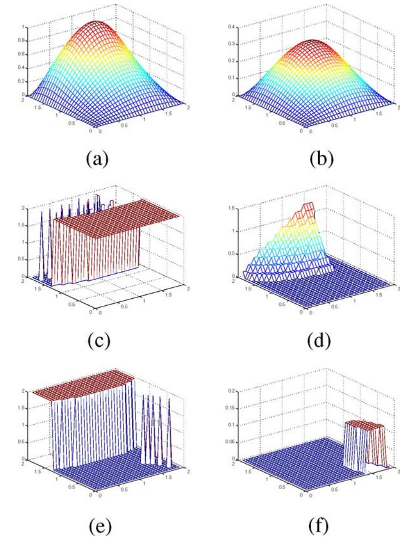


Fig. 1. Iteration in value space, $h = 2^{-4}$. (a) and (b) Graphs of upper values $V^{h,+}(x,1)$ and $V^{h,+}(x,2)$ for $x \in G_h$.(c) and (d) Graphs of player 1's optimal feedback controls when player 1 goes first. (e) and (f) Graphs of player 2's optimal feedback controls when player 1 goes first.

APPENDIX

Proof of Lemma 3.3: We need to verify the consistency criteria given in Definition 3.1. First,

$$\begin{aligned} E_{x,\iota,n}^{r,h}\Delta\xi_n^h &= E_{x,\iota,n}^{r,h}\left(\xi_{n+1}^h - x\right) \\ &= \sum_{j_0=1}^{l_0} e_{j_0} hp^h((x,\iota),(x+e_{j_0}h,\iota)|r) \\ &\quad + \sum_{j_0=1}^{l_0}(-e_{j_0})hp^h((x,\iota),(x-e_{j_0}h,\iota)|r) \\ &\quad + \sum_{j_0<k_0}(e_{j_0}+e_{k_0})hp^h((x,\iota),(x+e_{j_0}h+e_{k_0}h,\iota)|r) \\ &\quad + \sum_{j_0<k_0}(-e_{j_0}-e_{k_0})hp^h((x,\iota),(x-e_{j_0}h-e_{k_0}h,\iota)|r) \\ &\quad + \sum_{j_0\neq k_0}(e_{j_0}-e_{k_0})hp^h((x,\iota),(x+e_{j_0}h-e_{k_0}h,\iota)|r). \end{aligned} \tag{41}$$

Using the transition probability given in (28), we find that the sum of the third and the fourth terms in (41) is zero, and also the last term is zero. Hence,

$$\begin{aligned} E_{x,\iota,n}^{r,h}\Delta\xi_n^h &= \sum_{j_0=1}^{l_0}\frac{hb_{j_0}(x,\iota,r)}{D^h-\beta h^2}e_{j_0}h \\ &= b(x,\iota,r)\frac{h^2}{D^h-\beta h^2} \\ &= b(x,\iota,r)\Delta t^h(x,\iota,r) + b(x,\iota,r)\Delta t^h(x,\iota,r)\frac{\beta h^2}{D^h-\beta h^2} \\ &= b(x,\iota,r)\Delta t^h(x,\iota,r) + o(\Delta t^h(x,\iota,r)). \end{aligned} \tag{42}$$

The last line in (42) holds by the assumption A4). Use $e_{j_0 k_0}$ to denote the $l_0 \times l_0$ matrix with all entries 0, but 1 at the j_0th row and the k_0th column only. Then, we have

$$\begin{aligned}E^{r,h}_{x,\iota,n}\left(\Delta\xi^h_n\right)\left(\Delta\xi^h_n\right)' &= \sum_{j_0=1}^{l_0} e_{j_0 j_0} h^2 p^h((x,\iota),(x+e_{j_0}h,\iota)|r)\\ &+ \sum_{j_0=1}^{l_0} e_{j_0 j_0} h^2 p^h((x,\iota),(x-e_{j_0}h,\iota)|r)\\ &+ \sum_{j_0<k_0}(e_{j_0 j_0}+e_{k_0 j_0}+e_{j_0 k_0}+e_{k_0 k_0})h^2\\ &\quad p^h((x,\iota),(x+e_{j_0}h+e_{k_0}h,\iota)|r)\\ &+ \sum_{j_0<k_0}(e_{j_0 j_0}+e_{k_0 j_0}+e_{j_0 k_0}+e_{k_0 k_0})h^2\\ &\quad p^h((x,\iota),(x-e_{j_0}h-e_{k_0}h,\iota)|r)\\ &+ \sum_{j_0\neq k_0}(e_{j_0 j_0}-e_{k_0 j_0}-e_{j_0 k_0}+e_{k_0 k_0})h^2\\ &\quad p^h((x,\iota),(x+e_{j_0}h-e_{k_0}h,\iota)|r). \qquad (43)\end{aligned}$$

Using the symmetry of $A(x,\iota)$, we have

$$E^{r,h}_{x,\iota,n}\left(\Delta\xi^h_n\right)\left(\Delta\xi^h_n\right)' = A(x,\iota)\Delta t^h(x,\iota,r) + o(t^h(x,\iota,r)). \qquad (44)$$

By (28), the local consistency of $\alpha^h(\cdot)$ is straightforward to verify. Also, $|\xi^h_n| \le 2h$ with probability 1. This completes the proof. □

Proof of Theorem 4.2: We first note that, in view of Lemma 4.1, $\{\alpha^h(\cdot)\}$ is tight. Define a topology for the set $[0,\infty]$ such that $[0,\infty]$ (extended real numbers with one point compactification) is compact (see [13, pp. 258–259]). Then, the sequences $\{m^h, \tilde{\tau}_h\}$ are always tight, since their range spaces are compact (via the compactification). Similar to [13, Sec. 10.4.1], $\{w^h(\cdot)\}$ is tight, and converges weakly to the standard Brownian motion $w(\cdot)$. Now, we prove the tightness of $\{\xi^h(\cdot)\}$. To do this, we will verify that the two conditions of [13,Th 2.1 Ch. 9] hold for the process $\xi^h(\cdot)$. From (29), we have

$$\begin{aligned}&E_{x,\iota}|\xi^h(t)-x|^2\\ &= E_{x,\iota}\left|\int_0^t\int_U b(\xi^h(s),\alpha^h(s),r)m^h_s(dr)ds\right.\\ &\qquad \left.+\int_0^t \sigma(\xi^h(s),\alpha^h(s))dw^h(s)+\varepsilon^h(t)\right|^2\\ &= 3E_{x,\iota}\left|\int_0^t\int_U b(\xi^h(s),\alpha^h(s),r)m^h_s(dr)ds\right|^2\\ &\qquad + 3E_{x,\iota}\left|\int_0^t \sigma(\xi^h(s),\alpha^h(s))dw^h(s)\right|^2+\varepsilon^h(t)\\ &\le Kt^2+Kt+\varepsilon^h(t) \qquad (45)\end{aligned}$$

where K is a generic positive constant.

Let $\mathcal{T}^h_T$ be the set of $\mathcal{F}^h_t$-stopping times being less than or equal to T with probability 1. Then, for $\delta > 0$, $\tilde{\nu}_n \in \mathcal{T}^h_T$, by (45) and strong Markov property, we have

$$E_{x,\iota}|\xi^h(\tilde{\nu}_h+\delta)-\xi^h(\tilde{\nu}_h)|^2 = O(\delta)+\varepsilon^h(\delta) \text{ as } \delta\to 0. \quad (46)$$

So far, we have proved that $\{\xi^h(\cdot),\alpha^h(\cdot),m^h(\cdot),w^h(\cdot),\tilde{\tau}_h\}$ is tight.

For the rest of the proof, we assume that the probability space is chosen as required by the Skorokhod representation (see [13, Th. 1.7, Ch. 9]). With a slight abuse of notation, we may assume the convergence of the sequence $\{\xi^h(\cdot),\alpha^h(\cdot),m^h(\cdot),w^h(\cdot),\tilde{\tau}_h\}$ itself with a limit denoted by $(x(\cdot),\alpha(\cdot),m(\cdot),w(\cdot),\tilde{\tau})$, and the convergence is in the sense with probability 1 via Skorohod representation.

For $\delta > 0$ and any process $y(\cdot)$, define the process $y_\delta(\cdot)$ by $y_\delta(t) = y(n\delta)$, $t \in [n\delta, n\delta+\delta)$. Then, by the tightness of $\{\xi^h(\cdot),\alpha^h(\cdot)\}$, (29) can be written as

$$\begin{aligned}\xi^h(t) = x &+ \int_0^t\int_U b(\xi^h(s),\alpha^h(s),r)m^h_s(dr)ds\\ &+ \int_0^t \sigma\left(\xi^h_\delta(s),\alpha^h_\delta(s)\right)dw^h(s)+\varepsilon^{h,\delta}(t) \qquad (47)\end{aligned}$$

where $\lim_{\delta\to 0}\limsup_{h\to 0} E|\varepsilon^{h,\delta}(t)| = 0$. Taking limit as $h\to 0$, the convergence with probability 1 (through Skorohod representation) yields

$$\begin{aligned}E&\left|\int_0^t\int_U b(\xi^h(s),\alpha^h(s),r)m^h_s(dr)ds\right.\\ &\left.-\int_0^t\int_U b(x(s),\alpha(s),r)m^h_s(dr)ds\right| \to 0. \qquad (48)\end{aligned}$$

On the other hand, the sequence $\{m^h(\cdot)\}$ converges in the compact weak topology. In particular, for any bounded and continuous function $\phi(\cdot)$ with compact support, as $h\to 0$, we have

$$\int_0^\infty\int_U \phi(r,s)m^h(dr\,ds) \to \int_0^\infty\int_U \phi(r,s)m(dr\,ds).$$

The weak convergence and the Skorohod representation imply that, as $h\to 0$, we have

$$\begin{aligned}&\int_0^t\int_U b(x(s),\alpha(s),r)m^h_s(dr)ds\\ &\qquad \to \int_0^t\int_U b(x(s),\alpha(s),r)m_s(dr)ds \qquad (49)\end{aligned}$$

uniformly in t on any bounded interval with probability 1.

Since $\xi_\delta^h(\cdot)$ and $\alpha_\delta^h(\cdot)$ are piecewise constant functions, it follows from the probability 1 convergence, as $h \to 0$,

$$\begin{aligned}\int_0^t \sigma\big(\xi_\delta^h(s), \alpha_\delta^h(s)\big) dw^h(s) &= \sum_{l=1}^{t/\delta} \sigma\big(\xi_\delta^h(l\delta-\delta), \alpha_\delta^h(l\delta-\delta)\big)\big(w^h(l\delta) - w^h(l\delta-\delta)\big) \\ &\to \sum_{l=1}^{t/\delta} \sigma(x_\delta(l\delta-\delta), \alpha_\delta(l\delta-\delta))(w(l\delta) - w(l\delta-\delta)) \\ &= \int_0^t \sigma(x_\delta(s), \alpha_\delta(s)) dw(s). \end{aligned} \tag{50}$$

Combining (47)–(50), we can write

$$x(t) = x + \int_0^t \int_U b(x(s), \alpha(s), r) m_s(dr) ds + \int_0^t \sigma(x_\delta(s), \alpha_\delta(s)) dw(s) + \varepsilon_\delta(t) \tag{51}$$

where $\operatorname{limsup}_{\delta\to 0} E|\varepsilon_\delta(t)| = 0$. Finally, taking limits in the previous equation as $\delta \to 0$ yields the result.

Proof of Theorem 4.3: Suppose that player 1 goes first. Given $\varepsilon > 0$, there exist a small $\Delta > 0$ and an ε-optimal minimizing rule $u_1^{\Delta,\varepsilon}(\cdot|m_2) \in \mathcal{L}_1(\Delta)$ such that the following hold.

1) $u_1^{\Delta,\varepsilon}(\cdot|m_2)$ is defined by the conditional probability law: for small $\rho > 0$,

$$\begin{aligned} P\big\{u_1^{\Delta,\varepsilon}(n\Delta) = r_1 \big| u_1^{\Delta,\varepsilon}(k\Delta), k < n; w(s), \alpha(s), & \\ m_2(s), s < n\Delta\big\} & \\ = P\big\{u_1^{\Delta,\varepsilon}(n\Delta) = r_1 \big| w(k\rho), \alpha(k\rho), u_2^\rho(k\rho|m_2), & \\ k\rho < n\Delta; u_1^{\Delta,\varepsilon}(k\Delta), k < n\big\} & \\ = \widetilde{p}_{1,n}(r_1; w(k\rho), \alpha(k\rho), u_2^\rho(k\rho|m_2), l\rho < n\Delta; & \\ u_1^{\Delta,\varepsilon}(k\Delta), k < n), & \end{aligned} \tag{52}$$

where $u_2^\rho(\cdot|m_2)$ is the player 2's strategy that is a constant on $[k\rho, k\rho+\rho)$. Furthermore, $\widetilde{p}_{1,n}(\cdot)$ is a continuous function of its argument, otherwise, we can mollify $\widetilde{p}_{1,n}(\cdot)$ to be a smooth function (see [12, Th. 6.1]).

2) For small Δ and large Δ/ρ,

$$\sup_{m_2 \in \mathcal{U}_2} W\big(x, \iota, u_1^{\Delta,\varepsilon}(\cdot|m_2), m_2\big) \le V^+(x,\iota) + \varepsilon. \tag{53}$$

Note that $u_1^{\Delta,\varepsilon}(\cdot|m_2) \in \mathcal{L}_1(\Delta)$ needs to be modified or adapted for the Markov chain $\{\xi_n^h, \alpha_n^h\}$. Recall $\mathcal{U}_1^h(1)$ defined in (13). A control in $\mathcal{U}_1^h(1)$ is a constant on the interval $[t_n^h, t_{n+1}^h)$, and depends on the past information $u_2^h(s), s < t_n^h$. Let $w^h(\cdot)$ be given in (29), whose weak limit is a Brownian motion. For small ρ, the adaptation of $u_1^{\Delta,\varepsilon}(\cdot|m_2)$, denoted by $u_1^{h,\varepsilon}(\cdot|m_2^h) \in \mathcal{U}_1^h(1)$, for player 1 for the chain is represented by

$$\begin{aligned} P\{u_1^{h,\varepsilon}(t_n^h) = r_1 | u_1^{h,\varepsilon}(t_k^h), k < n;\ w^h(s), \alpha^h(s), m_{2,s}^h, s < t_n^h\} & \\ = P\big\{u_1^{h,\varepsilon}(t_n^h) = r_1 | w^h(k\rho), \alpha^h(k\rho), u_2^\rho\big(k\rho|m_2^h\big), k\rho < t_n^h; & \\ u_1^{h,\varepsilon}(t_k^h), k < n\} & \\ = \widetilde{p}_{1,[t_n^h/\Delta]}\big(r_1; w^h(k\rho), \alpha^h(k\rho), u_2^\rho\big(k\rho|m_2^h\big), k\rho < t_n^h; & \\ u_1^{h,\varepsilon}(k\Delta), k\Delta < t_n^h\big). & \end{aligned} \tag{54}$$

Given $u_1^{h,\varepsilon}(\cdot|m_2^h)$ for player 1, player 2 selects a maximizing control. Let $u_2^{h,*}(\cdot)$ denote player 2's maximizing choice in $\mathcal{U}_2^h(2)$ with its relaxed control representation $m_2^{h,*}(\cdot)$.

Let m^h be the relaxed control representation of $(u_1^{\Delta,\varepsilon}(\cdot|u_2^{h,*}), u_2^{h,*}\cdot)$. Choose an arbitrary weakly convergent subsequence of $\{\xi^h(\cdot), \alpha^h(\cdot), m^h(\cdot), w^h(\cdot)\}$ (abusing the terminology slightly, for simplicity, still index this subsequence by h), with limit $(x(\cdot), \alpha(\cdot), m(\cdot), w(\cdot))$, a solution of (4) by Theorem 4.2, where $m(\cdot) = (m_1(\cdot), m_2(\cdot))$. Then,

$$W^h(x, \iota, m^h(\cdot)) \to W(x, \iota, m(\cdot)) \text{ as } h \to 0. \tag{55}$$

Let us use the Skorohod representation theorem so that all processes are defined on the same probability space and the weak convergence becomes convergence with probability 1. Then, as $h \to 0$, we have

$$\begin{aligned} \widetilde{p}_{1,n}(r_1; w^h(k\rho), \alpha^h(k\rho), u_2^\rho\big(k\rho|m_2^{h,*}\big), k\rho < n\Delta; & \\ u_1^{h,\varepsilon}(k\Delta), k < n) & \\ \to \widetilde{p}_{1,n}(r_1; w(k\rho), \alpha(k\rho), u_2^\rho(k\rho|m_2), k\rho < n\Delta; & \\ u_1^{\Delta,\varepsilon}(k\Delta), k < n) \quad \text{with probability 1.} & \end{aligned} \tag{56}$$

This implies that $(u_1^{h,\varepsilon}(\cdot|m_2^{h,*}), m_2^{h,*}(\cdot)) \Rightarrow (u_1^{\Delta,\varepsilon}(\cdot|m_2), m_2(\cdot))$. Hence, as $h \to 0$, we have

$$W^h\big(x, \iota, u_1^{h,\varepsilon}\big(\cdot\,|u_2^{h,*}\big), u_2^{h,*}\big) \to W(x, \iota, u_1^{\Delta,\varepsilon}(\cdot|m_2), m_2). \tag{57}$$

Therefore,

$$\begin{aligned} V^{h,+}(x,\iota) &= \inf_{u_1^h \in \mathcal{U}_1^h(1)} \sup_{u_2^h \in \mathcal{U}_2^h(2)} W^h\big(x, \iota, u_1^h, u_2^h\big) \\ &\le \sup_{u_2^h \in \mathcal{U}_2^h(2)} W^h\big(x, \iota, u_1^{h,\varepsilon}\big(\cdot\,|u_2^h\big), u_2^h\big) \\ &= W^h\big(x, \iota, u_1^{h,\varepsilon}\big(\cdot\,|u_2^{h,*}\big), u_2^{h,*}\big) \\ &\to W\big(x, \iota, u_1^{\varepsilon,\Delta}(\cdot|m_2), m_2\big) \text{ as } h \to 0. \end{aligned} \tag{58}$$

From (53) and (58), we conclude

$$\operatorname*{limsup}_{h\to 0} V^{h,+}(x,\iota) \le V^+(x,\iota) + \varepsilon. \tag{59}$$

On the other hand, note that, for $\Delta > 0$ and small $h > 0$, there exists an $\varepsilon^{\Delta,h}$-optimal minimizing control $\overline{u}_1^{h,\Delta}(\cdot|m_2^h) \in \mathcal{U}_1^h(1)$, with relaxed control representation $\overline{m}_1^{h,\Delta}(\cdot|m_2^h)$ such that the following hold.

1) $\overline{u}_1^{h,\Delta}(\cdot|m_2^h)$ is constant on $[l\Delta, l\Delta+\Delta)$.

2)

$$\sup_{m_2^h \in \mathcal{U}_2^h(2)} W^h\big(x, \iota, \overline{m}_1^{h,\Delta}\big(\cdot\,|m_2^h\big), m_2^h\big) \le V^{h,+}(x,\iota) + \varepsilon^{\Delta,h}. \tag{60}$$

3) $\lim_{\Delta\to 0} \operatorname{limsup}_{h\to 0} \varepsilon^{\Delta,h} = 0$.

Given $\overline{m}_1^{h,\Delta}(\cdot|m_2^h) \in \mathcal{U}_1^h(1)$, there exists a maximizing control $\overline{m}_2^{h,\Delta} \in \mathcal{U}_2^h(2)$ satisfying the following.

1) If $(\xi^h(\cdot), \alpha^h(\cdot), \overline{m}_1^{h,\Delta}, \overline{m}_2^{h,\Delta}) \Rightarrow (x(\cdot), \alpha(\cdot), \overline{m}_1^{\Delta}, \overline{m}_2^{\Delta})$, then it is the weak solution of (4) by Theorem 4.2 and $\overline{m}_1^{\Delta} \in \mathcal{L}_1(\Delta), \overline{m}_2^{\Delta} \in \mathcal{U}_2$, such that

$$W^h\big(x, \iota, \overline{m}_1^{h,\Delta}\big(\cdot\,|\overline{m}_2^{h,\Delta}\big), \overline{m}_2^{h,\Delta}\big) \to W\big(x, \iota, \overline{m}_1^{\Delta}\big(\cdot\,|\overline{m}_2^{\Delta}\big), \overline{m}_2^{\Delta}\big) \text{ as } h \to 0. \quad (61)$$

2) Furthermore, $\overline{m}_2^{\Delta}$ is an optimal maximizing control in $\mathcal{U}_2$, given $\overline{m}_1^{\Delta} \in \mathcal{L}_1(\Delta)$. That is,

$$\sup_{m_2 \in \mathcal{U}_2} W\big(x, \iota, \overline{m}_1^{\Delta}(\cdot|m_2), m_2\big) = W\big(x, \iota, \overline{m}_1^{\Delta}\big(\cdot\,|\overline{m}_2^{\Delta}\big), \overline{m}_2^{\Delta}\big). \quad (62)$$

Consider $m_2(\cdot)$ as a pure control of $W(x, \iota, \overline{m}_1^{\Delta}, \cdot)$. Then, the earlier statement can be derived from a result from pure control problem (see [18, Th. 5.3]).

Let $\varepsilon^{\Delta} = \limsup_{h\to 0} \varepsilon^{\Delta,h}$. By the controls defined before, we have

$$\begin{aligned}
V^{h,+}(x,\iota) &= \inf_{m_1^h \in \mathcal{U}_1^h(1)} \sup_{m_2^h \in \mathcal{U}_2^h(2)} W^h\big(x,\iota,m_1^h,m_2^h\big)\\
&\geq \sup_{m_2^h \in \mathcal{U}_2^h} W^h\big(x,\iota,\overline{m}_1^{h,\Delta}\big(\cdot\,|m_2^h\big),m_2^h\big) - \varepsilon^{\Delta,h} \text{ by (60)}\\
&= W^h\big(x,\iota,\overline{m}_1^{h,\Delta}\big(\cdot\,|\overline{m}_2^{h,\Delta}\big),\overline{m}_2^{h,\Delta}\big) - \varepsilon^{\Delta,h}\\
&\to W\big(x,\iota,\overline{m}_1^{\Delta}\big(\cdot\,|\overline{m}_2^{\Delta}\big),\overline{m}_2^{\Delta}\big) - \varepsilon^{\Delta} \text{ as } h \to 0\\
&= \sup_{m_2 \in \mathcal{U}_2} W(x,\iota,\overline{m}_1^{\Delta}(\cdot|m_2),m_2) - \varepsilon^{\Delta} \quad \text{by (62)}\\
&\geq \inf_{m_1 \in \mathcal{L}_1(\Delta)} \sup_{m_2 \in \mathcal{U}_2} W(x,\iota,m_1(\cdot|m_2),m_2) - \varepsilon^{\Delta}\\
&\to V^+(x,\iota) \text{ as } \Delta \to 0. \quad (63)
\end{aligned}$$

This proves $\liminf_{h\to 0} V^{h,+}(x,\iota) \geq V^+(x,\iota)$. Combining (59) and the previous limit, we obtain the first limit in Theorem 4.3. The proof of $\lim_{h\to 0} V^{h,-}(x,\iota) = V^-(x,\iota)$ is similar. □

Proof of Theorem 5.3: Rewrite the dynamic programming equations (17) and (18) as

$$V^{h,+}(x,\iota) = \inf_{r_1 \in U_1} \sup_{r_2 \in U_2} [\widetilde{k}(x,\iota,r_1,r_2)\Delta t^h(x,\iota) + \sum_{(y,\ell)\in G_h \times \mathcal{M}} p^h((x,\iota),(y,\ell)|r_1,r_2) e^{-\beta\Delta t^h(x,\iota)} V^{h,+}(y,\ell)] \quad (64)$$

$$V^{h,-}(x,\iota) = \sup_{r_2 \in U_2} \inf_{r_1 \in U_1} [\widetilde{k}(x,\iota,r_1,r_2)\Delta t^h(x,\iota) + \sum_{(y,\ell)\in G_h \times \mathcal{M}} p^h((x,\iota),(y,\ell)|r_1,r_2) e^{-\beta\Delta t^h(x,\iota)} V^{h,-}(y,\ell)]. \quad (65)$$

Define two functions

$$\phi^{h,+}(x,\iota,r_1,r_2) = \widetilde{k}(x,\iota,r_1,r_2)\Delta t^h(x,\iota) + \sum_{(y,\ell)\in G_h \times \mathcal{M}} p^h((x,\iota),(y,\ell)|r_1,r_2) e^{-\beta\Delta t^h(x,\iota)} V^{h,+}(y,\ell) \quad (66)$$

and

$$\phi^{h,-}(x,\iota,r_1,r_2) = \widetilde{k}(x,\iota,r_1,r_2)\Delta t^h(x,\iota) + \sum_{(y,\ell)\in G_h \times \mathcal{M}} p^h((x,\iota),(y,\ell)|r_1,r_2) e^{-\beta\Delta t^h(x,\iota)} V^{h,-}(y,\ell). \quad (67)$$

Then, the dynamic programming equations (64) and (65) are written as

$$V^{h,+}(x,\iota) = \inf_{r_1} \sup_{r_2} \phi^{h,+}(x,\iota,r_1,r_2) \quad (68)$$

$$V^{h,-}(x,\iota) = \sup_{r_2} \inf_{r_1} \phi^{h,-}(x,\iota,r_1,r_2). \quad (69)$$

To proceed, we first establish a lemma. This lemma enables us to interchange the supremum and infimum in $\phi^{h,+}(\cdot)$.

Lemma 1.1: Under A1)–A5), and A6) or A7), $\inf_{r_1} \sup_{r_2} \phi^{h,+}(x,\iota,r_1,r_2) = \sup_{r_2} \inf_{r_1} \phi^{h,+}(x,\iota,r_1,r_2)$.

Proof of Lemmma 1.1: Under A6), $b(x,\iota,\cdot,\cdot)$ is separable, so is $p^{h,+}((x,\iota),(y,\ell)|\cdot,\cdot)$ of (32). In addition, $\widetilde{k}(x,\iota,\cdot,\cdot)$ is separable implies that $\phi^{h,+}(x,\iota,\cdot,\cdot)$ is separable in r_1 and r_2 for every $(x,\iota) \in G_h \times \mathcal{M}$. This implies that inf and sup in (68) and (69) can be interchanged without affecting the equality. Under A7), every $p^h((x,\iota),(y,\ell)|r_1,r_2)$ satisfies the convex–concave condition, so

$$\sum_{(y,\ell)\in G_h \times \mathcal{M}} p^h((x,\iota),(y,\ell)|r_1,r_2) e^{-\beta\Delta t^h(x,\iota)} V^{h,+}(y,\ell)$$

is convex–concave. In addition, $\widetilde{k}(x,\iota,\cdot,\cdot)$ is convex–concave. This implies that $\phi^{h,+}(x,\iota,\cdot,\cdot)$ is convex–concave. By Lemma 5.2, we can interchange sup and inf in (68) and (69). □

Completion of Proof of Theorem 5.3: To complete the proof of the theorem, let $\rho = \max_{(x,\iota)\in G_h \times \mathcal{M}}(V^{h,+}(x,\iota) - V^{h,-}(x,\iota))$. It is plain that $\rho \geq 0$. So, it is enough to show that $\rho = 0$. Note that there exists $(\hat{x},\hat{\iota}) \in G_h \times \mathcal{M}$, which satisfies

$$\begin{aligned}
&V^{h,+}(\hat{x},\hat{\iota}) = V^{h,-}(\hat{x},\hat{\iota}) + \rho\\
&V^{h,+}(x,\iota) \leq V^{h,-}(x,\iota) + \rho \;\forall (x,\iota) \in G_h \times \mathcal{M}. \quad (70)
\end{aligned}$$

For $(\hat{x},\hat{\iota}) \in G_h \times \mathcal{M}$ given in (70), we have the following series of inequalities:

$$\begin{aligned}
V^{h,+}(\hat{x},\hat{\iota}) &= \inf_{r_1} \sup_{r_2} \{\phi^{h,+}(\hat{x},\hat{\iota},r_1,r_2)\}, \quad \text{by (68)}\\
&= \sup_{r_2} \inf_{r_1} \{\phi^{h,+}(\hat{x},\hat{\iota},r_1,r_2)\}, \quad \text{by Lema 1.1}\\
&= \sup_{r_2} \inf_{r_1} \{\widetilde{k}(\hat{x},\hat{\iota},r_1,r_2)\Delta t^h(\hat{x},\hat{\iota}) + \quad \text{by (66)}\\
&\sum_{(y,\ell)\in G_h \times \mathcal{M}} p^h((\hat{x},\hat{\iota}),(y,\ell)|r_1,r_2) e^{-\beta\Delta t^h(\hat{x},\hat{\iota})} V^{h,+}(y,\ell)\},
\end{aligned}$$

$$
\begin{aligned}
&\leq \sup_{r_2}\inf_{r_1}\{\widetilde{k}(\hat{x},\hat{\imath},r_1,r_2)\Delta t^h(\hat{x},\hat{\imath}) + \sum_{(y,\ell)\in G_h\times\mathcal{M}} e^{-\beta\Delta t^h(\hat{x},\hat{\imath})} \\
&\quad p^h((\hat{x},\hat{\imath}),(y,\ell)|r_1,r_2)(V^{h,-}(y,\ell)+\rho)\}, \quad \text{by (70)} \\
&= \sup_{r_2}\inf_{r_1}\{\phi^{h,-}(\hat{x},\hat{\imath},r_1,r_2) + e^{-\beta\Delta t^h(\hat{x},\hat{\imath})}\rho\} \\
&= \sup_{r_2}\inf_{r_1}\{\phi^{h,-}(\hat{x},\hat{\imath},r_1,r_2)\} + e^{-\beta\Delta t^h(\hat{x},\hat{\imath})}\rho \\
&= V^{h,-}(\hat{x},\hat{\imath}) + e^{-\beta\Delta t^h(\hat{x},\hat{\imath})}\rho, \quad \text{by (69)} \\
&\leq V^{h,-}(\hat{x},\hat{\imath}) + \rho = V^{h,+}(\hat{x},\hat{\imath}), \quad \text{since } \beta>0, \quad \rho\geq 0.
\end{aligned}
$$

Note that $\rho > 0$ implies that the last line in this equation is a strict inequality, which is a contradiction to (70). Therefore, $\rho = 0$. The existence of the saddle point is established. □

ACKNOWLEDGMENT

The authors would like to thank Prof. W. Cohn for pointing out [19] to them and for discussion on static games. They are also grateful to Prof. H. Kushner for discussion on stochastic differential game problems and for checking over an early version of their proof in the existence of a saddle point of the game. They are further grateful to the reviewers and editors for their comments and suggestions.

REFERENCES

[1] T. Basar and P. Bernhard, *H_∞-Optimal Control and Related Minimax Problems*. Boston, MA: Birkhäuser, 1991.

[2] W. P. Blair and D. D. Sworder, "Feedback control of a class of linear discrete systems with jump parameters and quadratic cost criteria," *Int. J. Control*, vol. 21, pp. 833–841, 1986.

[3] G. B. Di Masi, Y. M. Kabanov, and W. J. Runggaldier, "Mean variance hedging of options on stocks with Markov volatility," *Theory Probab. Appl.*, vol. 39, pp. 173–181, 1994.

[4] W. H. Fleming, "The convergence problem for differential games," *J. Math. Anal. Appl.*, vol. 3, pp. 102–116, 1961.

[5] W. H. Fleming, "The convergence problem for differential games," in *Advances in Game Theory, Annals of Mathematical Studies,* vol. 52, Princeton, NJ: Princeton Univ. Press, 1964, pp. 195–210.

[6] M. K. Ghosh, A. Arapostathis, and S. I. Marcus, "Optimal control of switching diffusions with application to flexible manufacturing systems," *SIAM J. Control Optim.*, vol. 31, pp. 1183–1204, 1993.

[7] R. J. Elliott and N. J. Kalton, "Existence of value in differential games," *Mem. Amer. Math. Soc.*, vol. 126, pp. 1–67, 1972.

[8] W. H. Fleming and W. M. McEneaney, "Risk-sensitive control on an infinite time horizon," *SIAM J. Control Optim.*, vol. 33, pp. 1881–1915, 1995.

[9] Y. Ji and H. J. Chizeck, "Controllability, stabilizability, and continuous-time Markovian jump linear quadratic control," *IEEE Trans. Autom. Control*, vol. 35, no. 7, pp. 777–788, Jul. 1990.

[10] H. J. Kushner and S. G. Chamberlain, "On stochastic differential games: sufficient conditions that a given strategy be a saddle point, and numerical procedures for the solution of the game," *J. Math. Anal. Appl.*, vol. 26, pp. 560–575, 1969.

[11] H. J. Kushner, "Numiercal methods for stochastic control problems in continuous time," *SIAM J. Control Optim.*, vol. 28, pp. 999–1048, 1990.

[12] H. J. Kushner, "Numerical approximations for stochastic differential games," *SIAM J. Control Optim.*, vol. 41, pp. 457–486, 2002.

[13] H. J. Kushner and P. Dupuis, *Numerical Methods for Stochastic Control Problems in Continuous Time*, 2nd ed. New York: Springer-Verlag, 2001.

[14] X. Mao, "Stability of stochastic differential equations with Markovian switching," *Stochastic Process. Appl.*, vol. 79, pp. 45–67, 1999.

[15] M. Mariton and P. Bertrand, "Robust jump linear quadratic control: A mode stabilizing solution," *IEEE Trans. Autom. Control*, vol. AC-30, no. 11, pp. 1145–1147, Nov. 1985.

[16] T. E. S. Raghvan and J. A. Filar, "Algorithms for stochastic games: A survey," *Z. Oper. Res.*, vol. 35, pp. 437–472, 1991.

[17] M. Sion, "On general minimax theorems," *Pacific J. Math.*, vol. 8, pp. 171–176, 1958.

[18] Q. S. Song, G. Yin, and Z. Zhang, "Numerical method for controlled regime-switching diffusions and regime-switching jump diffusions," *Automatica*, vol. 42, pp. 1147–1157, 2006.

[19] J. von Neumann, *Collected Works Volume VI, Theory of Games, Astrophysics, Hydrodynamics and Meteorology*. New York: Pergamon Press, 1962, pp. 82–88.

[20] G. Yin, V. Krishnamurthy, and C. Ion, "Regime switching stochastic approximation algorithms with application to adaptive discrete stochastic optimization," *SIAM J. Optim.,* vol. 14, pp. 1187–1215, 2004.

[21] G. Yin, R. H. Liu, and Q. Zhang, "Recursive algorithms for stock liquidation: A stochastic optimization approach," *SIAM J. Optim.,* vol. 13, pp. 240–263, 2002.

[22] G. Yin and Q. Zhang, *Continuous-Time Markov Chains and Applications: A Singular Perturbation Approach*. New York: Springer-Verlag, 1998.

[23] G. Yin and Q. Zhang, *Discrete-time Markov Chains: Two-time-scale Methods and Applications*. New York: Springer-Verlag, 2005.

[24] G. Yin, Q. Zhang, and G. Badowski, "Discrete-time singularly perturbed Markov chains: Aggregation occupation measures, and switching diffusion limit," *Adv. Appl. Probab.*, vol. 35, pp. 449–476, 2003.

[25] Q. Zhang, "Stock trading: An optimal selling rule," *SIAM J. Control Optim.*, vol. 40, pp. 64–87, 2001.

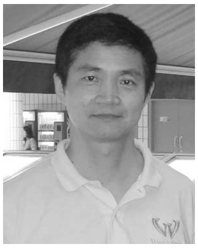
[26] X. Y. Zhou and G. Yin, "Markowitz mean-variance portfolio selection with regime switching: A continuous-time model," *SIAM J. Control Optim.*, vol. 42, pp. 1466–1482, 2003.

Qingshuo Song received the B.S. degree in automatic control and systems in 1996 and the M.A. degree in computer science in 1999 from Nankai University, Tianjin, China, and the Ph.D. degree in mathematics from Wayne State University, Detroit, MI, in 2006.

He is currently a nontenure track (NTT) Assistant Professor in the Department of Mathematics, University of Southern California, Los Angeles. His current research interests include stochastic control, mathematical finance, numerical analysis, and stochastic differential games.

G. George Yin (S'87–M'87–S'96–F'02) received the B.S. degree in mathematics from the University of Delaware, Newark, in 1983, and the M.S. degree in electrical engineering and the Ph.D. degree in applied mathematics from Brown University, Providence, RI, in 1987. Since 1996, he has been a Professor in the Department of Mathematics, Wayne State University, Detroit, MI.

Prof. Yin was the Editor of the SIAM Activity Group on Control and Systems Theory Newsletters. He has also been a member of the Mathematical Review Date Base Committee, the International Federation of Automatic Control (IFAC) Technical Committee on Modeling, Identification and Signal Processing, and various conference program committees. He is an Associate Editor of *Automatica* and *Society for Industrial and Applied Mathematics (SIAM) Journal on Control and Optimization*. From 1994 to 1998, he was an Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL. He is (or was) on the editorial board of eight other journals. He was the SIAM Representative to the 34th Conference on Decision and Control (CDC), the Co-chair of the 1996 American Mathematical Society (AMS)–SIAM Summer Seminar in Applied Mathematics, the Co-chair of the 2003 AMS–Institute of Mathematical Statistics (IMS)–SIAM Summer Research Conference: Mathematics of Finance, and Co-organizer of the 2005 Institute for Mathematics and its Applications (IMA) Workshop on Wireless Communications, and a Co-organizer of 2006 IMA Participating Institution (PI) Conference.



Zhimin Zhang received the B.S. degree in mathematics in 1982 and the M.S. degree in computational mathematics in 1985 from the University of Science and Technology of China, Hefei, China, and the Ph.D. degree in applied mathematics in 1991 from the University of Maryland, College Park.

From 1991 to 1999, he was an Assistant Professor and Associate Professor at Texas Tech University, Lubbock. He is currently a Professor in the Department of Mathematics, Wayne State University, Detroit, MI. His current research interests include numerical solutions for partial differential equations, computational mechanics, and scientific computing.

Prof. Zhang is an Associate Editor of the *Discrete and Continuous Dynamical Systems—Series B* and an Associated Editor of the *International Journal of Numerical Analysis and Modeling*. In 2006, he joined the Editorial Board of the *Journal of Computational Mathematics*. He was the organizer of the 2000 National Science Foundation (NSF)-Conference Board of the Mathematical Sciences (CBMS) Regional Conference in the Mathematical Sciences: "Superconvergence in Finite Element Methods" at Texas Tech University.