SIAM J. CONTROL OPTIM.
Vol. 41, No. 2, pp. 457–486

NUMERICAL APPROXIMATIONS FOR STOCHASTIC DIFFERENTIAL GAMES*

HAROLD J. KUSHNER†

Abstract. The Markov chain approximation method is a widely used, robust, relatively easy to use, and efficient family of methods for the bulk of stochastic control problems in continuous time for reflected-jump-diffusion-type models. It has been shown to converge under broad conditions, and there are good algorithms for solving the numerical problems if the dimension is not too high. Versions of these methods have been used in applications to various two-player differential and stochastic dynamic games for a long time, and proofs of convergence are available for some cases, mainly using PDE-type techniques. In this paper, purely probabilistic proofs of convergence are given for a broad class of such problems, where the controls for the two players are separated in the dynamics and cost function, and which cover a substantial class not dealt with in previous works. Discounted and stopping time cost functions are considered. Finite horizon problems and problems where the process is stopped on first hitting an a priori given boundary can be dealt with by adapting the methods of [H. J. Kushner and P. Dupuis, *Numerical Methods for Stochastic Control Problems, in Continuous Time,* 2nd ed., Springer-Verlag, Berlin, New York, 2001] as done in this paper for the treated problems. The essential conditions are the weak-sense existence and uniqueness of solutions, an "almost everywhere" continuity condition, and that a weak local consistency condition holds "almost everywhere" for the numerical approximations, just as for the control problem. There are extensions to problems with controlled variance and jumps.

Key words. stochastic differential games, numerical methods, Markov chain approximations

AMS subject classifications. 60F17, 65C30, 65C40, 91A15, 91A23, 93E25

PII. S0363012901389457

1. Introduction.

The Markov chain approximation method of [25, 26, 32] is an effective and widely used method for the numerical solution of virtually all of the standard forms of stochastic control problems with reflected-jump-diffusion models. It is robust and can be shown to converge under very broad conditions. In this paper, the basic ideas will be extended to two-player stochastic dynamic games with the same systems model, but where the controls for the two players are separated in the dynamics and cost functions, and for certain classes of stopping time problems. Such "separated" models occur, for example, in pursuit-evasion games, where each player controls its own dynamics, risk-sensitive and robust control [2, 3, 8, 18, 36], Lagrangian formulation of optimization under side constraints, and controlled large deviation problems [12]. When the robust control is for controlled queues in heavy traffic, with or without finite buffers, or for its fluid limits, then the state is confined to some convex polyhedron by boundary reflection [28]. See section 8 for a few illustrations. The minimizing and maximizing players will be called, respectively, players 1 and 2.

Early results concerning algorithms and convergence for stochastic games for finite-state Markov chain models are in [30, 31], and a survey is in [37]. The performance of all of these algorithms can be improved with the use of multigrid, Gauss–Seidel, and various accelerated versions. See [32] for additional references and more

*Received by the editors May 17, 2001; accepted for publication (in revised form) November 30, 2001; published electronically June 18, 2002. This work was partially supported by contract DAAD19-99-1-0223 from the Army Research Office and National Science Foundation grant ECS 0097447.

http://www.siam.org/journals/sicon/41-2/38945.html

†Applied Mathematics Department, Lefschetz Center for Dynamical Systems, Brown University, Providence, RI 02912 (hjk@dam.brown.edu).

detail concerning such accelerated algorithms.

Partial results for the convergence problem for approximations of various forms of continuous state and time dynamic games have appeared, but there does not seem to be a complete development for the fully stochastic problem for reflected-jump-diffusion models. The upper value for a deterministic game (an ODE model) was treated by the Markov chain approximation method in [34, 35]. Results for various deterministic problems are in [4, 5, 6, 7, 40, 41]. The actual numerical methods which are used in the computations tend to be of the Markov chain approximation type, although the proofs are sometimes based on subsequent PDE techniques.

In this paper, we will use purely probabilistic methods of proof. Such methods have the advantage of providing intuition concerning numerical approximations, they cover many of the problem formulations to date, and they converge under quite general conditions. The essential conditions are *weak-sense* existence and uniqueness of the solution to the controlled equations, "almost everywhere" continuity of the dynamical and cost rate terms, and a natural "local consistency" condition: The local consistency and continuity need hold only almost everywhere with respect to the measure of the basic model; hence discontinuities and severe singularities in the dynamics and cost function can be treated under appropriate conditions (see, in particular, Theorems 4.7 and 7.1 and the treatment of discontinuities and complex variational problems with singularities in [32]). Furthermore, the numerical approximations are represented as processes which are close to the original, which gives additional intuitive and practical meaning to the method. Indeed, the Markov chain approximation method seems to provide the intuition for many of the actual numerical methods which are used, no matter what the method of proof of convergence.

We will treat only a selection of problems. The basic controlled process $x(\cdot)$ is defined by (2.2) or, equivalently, (2.4). We concentrate on discounted and stopping time cost functions. Others, such as finite horizon problems and problems where the process is stopped on first hitting an a priori given boundary, can be dealt with by adapting the methods of [32] as done in this paper for the treated problems.

In many applications, the state of the actual physical problem is confined to a bounded set, and the reflection term in (2.2) ensures the correct boundary behavior. One example is the heavy traffic limit of controlled queueing networks with finite buffers [1, 28] or robust control of such systems as in [2, 3], where the set is a hyperrectangle. Then robust control or the optimization under side constraints would lead to a game problem with a hyperrectangular state space. Another example would be the control of large deviations for such problems, along the lines of [12]. If the system state is not a priori confined to a bounded set, then, for numerical purposes, it is commonly necessary to bound the state space artificially and then experiment with the bounds. Such problems which are bounded for numerical purposes often involve reflecting boundaries. For this reason, our basic model is confined to a state space G that is a convex polyhedron, and it is confined by a "reflection" on the boundary. In [32], the boundary of the state space was determined by a set of smooth curved surfaces. We restrict our attention to the simpler polyhedral case, since that is the one most widely used, and it avoids details which distract from the general development. However, the approximations of the more general boundaries that were used in [32] can be carried over without change to the problem of this paper. Similarly, for simplicity, we drop the jump term (which is treated in [32]) since including it for the game involves no new issues. See also [27] for a setup where the jumps themselves are controlled. Again, for simplicity, we do not allow the variance to be controlled.

However, if (see (2.2)) $w(\cdot) = (w_1(\cdot), w_2(\cdot))$, where the $w_i(\cdot), i = 1, 2$, are mutually independent, and we have the separated form $\sigma(x, u)dw = \sigma_1(x, u_1)dw_1 + \sigma_2(x, u_2)dw_2$, then the methods in [32, Chapter 13] or [26] can be adapted.

The methods to be used are based on the theory of weak convergence [10, 14] as they are applied in [32]. For any process with values in a complete and separable metric space S, let $D(S; 0, \infty)$ denote the space of S-valued paths on the time interval $[0, \infty)$ which are right continuous and have left-hand limits, and with the Skorohod topology used. The path space for the state process $x(\cdot)$ is $D(G; 0, \infty)$, where $G \subset \mathbb{R}^r$, r-dimensional Euclidean space. The tightness criterion to be used implicitly is Theorem 2.7b of [24], which is restated as [32, Theorem 9.2.1].

The development involves various concepts from stochastic control and game theory, weak-sense solutions, the Skorohod problem, and numerical methods for stochastic control, not all of which will be familiar to many readers. Because of this, to make the material as accessible as possible as well as to minimize detail, the development has been structured to take advantage of the results in [32] whenever possible. The analysis for the game problem is more difficult than that for the pure control problem, since we must work with strategies and not simply controls, the strategies of the two players might be dependent, and they need to be approximated in various ways for purposes of the analysis.

Sections 2 and 3 give the basic systems model and describe the numerical method. They also contain necessary background material. The dynamical model is the reflected SDE (2.2) or (2.4), also called the *Skorohod problem* [11, 32, 28]. (See also the beginning of section 4.) The conditions on the boundary of the state space are A2.1–A2.2. Condition A2.1 covers the great majority of cases of current interest, including those that arise from queueing and communications networks, as noted in section 2. The condition is trivial to verify for the special case where the state space is a hyperrectangle, with reflection directions being the interior normals. As is common in control theory when limits of a sequence of controls are involved, much of the analysis uses the notion of relaxed control, and the necessary definitions are given. The definitions of the upper and lower value of the game requires a precise definition of the class of allowed strategies. These are given in section 2. Later, we will define various subclasses of these sets which are needed in the approximation and limit proofs. The bulk of the paper works with weak-sense solutions. This allows the greatest generality, including the possibility of using Girsanov transformation methods for constructing solutions, hence the possibility of discontinuous dynamics. However, it comes at a price since the notation is more complicated than what would be required if Lipschitz conditions (hence strong-sense solutions) were used.

The numerical method, which is the Markov chain approximation procedure, is discussed in section 3. The actual ways of approximating the original problem to get the approximating chain and associated cost function are the same as in [32] for the pure control problem since it is the process for arbitrary controls that is approximated. The basic and natural *local consistency conditions* are stated. The approximation to the original process $x(\cdot)$ is a continuous time interpolation of the chain, and this interpolation $\psi^h(\cdot)$ is defined, and a useful representation is given. The upper and lower values for the game for the chain are defined.

The actual proof of convergence of the numerical method in Theorem 7.1 is not long. However, it depends on many approximations and estimates as well as on the fact that the original game has a value. These issues are dealt with in sections 4–6. Under a Lipschitz condition, Theorem 4.3 shows that the costs are well approximated (in a uniform sense) if the controls are, and Theorem 4.4 proves similar facts when there is only a weak-sense solution. Then it is shown that a fine discretization of any of the controls in space and time, and even slightly delaying the actions of any of the controls, changes the costs only slightly, again uniformly in the controls. Loosely speaking, the costs are continuous in the controls of either player and uniform in the control of the other. Theorem 4.7 shows, under appropriate conditions, how to get similar results when the dynamical and cost rate functions are discontinuous. These approximations are fundamental to the proof of the existence of the value of the game in section 5 since they imply that slight delays in any of the controls have little effect on the results, which in turn implies that "who goes first" is not too important.

Section 6 contains the final "auxiliary" result. In the proof of convergence of the numerical method, one needs to use ϵ-optimal strategies for the player who goes first. These strategies are for theoretical purposes only and *do not have any use in practice.* The construction of these strategies is much more complicated than what is required for the pure control problem, and it is done in Theorem 6.1.

The convergence of the numerical method is given in section 7. In the pure control problem of [32], the numerical approximations are controlled Markov chains, and one needs to show that the sequence of approximations to the optimal value function converges as the approximation parameter goes to zero. Here, the numerical approximations are games for Markov chains. They might or might not have a value, depending on the form of the approximation. However, one needs to show at least that the upper and lower values converge to the value of the original game as the approximation parameter goes to its limit. This is more difficult than the proof of convergence for the control problem, and one needs to keep careful track of the information available to the individual players.

Section 8 contains a brief discussion of some examples and extensions. The treatment of the ergodic cost case uses quite different methods and is in [29].

2. The system model.

Assumptions on the state space G. It is assumed that the system state $x(t)$ is confined to the set G by boundary reflections. Conditions A2.1 and A2.2 are common in treatments of SDEs with reflections and piecewise smooth boundaries [11, 32, 28].

A2.1. *G is a bounded convex polyhedron in r-dimensional Euclidean space $\mathbb{R}^r$ with an interior and a finite number of faces. Let d_i denote the direction of reflection to the interior of the ith face, assumed constant there. On any edge or corner, the reflection direction can be any nonnegative linear combination of the directions on the adjoining faces. Let $d(x)$ denote the set of reflection directions at $x \in \partial G$, the boundary of G. For an arbitrary corner or edge of ∂G, let $\bar{d}_i$ and $\bar{n}_i$ denote the direction of reflection and the interior normal, respectively, on the ith adjoining face. Then there are constants $a_i > 0$ (depending on the edge or corner) such that*

$$a_i \langle \bar{n}_i, \bar{d}_i \rangle > \sum_{j:j \neq i} a_j \left| \langle \bar{n}_i, \bar{d}_j \rangle \right| \quad \textit{for all } i. \tag{2.1}$$

A2.2. *There is a neighborhood $N(\partial G)$ and an extension of $d(\cdot)$ to $\overline{N(\partial G)}$ such that the following holds: For each $\epsilon > 0$, there is $\mu > 0$ which goes to zero as $\epsilon \to 0$ and such that if $x \in \overline{N(\partial G)} - \partial G$ and distance$(x, \partial G) \leq \mu$, then $d(x)$ is in the convex hull of $\{d(v); v \in \partial G, \text{distance}(x, v) \leq \epsilon\}$.*

In A2.3, the real variables c_i are the coefficients in the cost rate term $c'dy$, $c = \{c_i\}$, in (2.5), and U_i is the space of values for the control $u_i(t)$ of player i in (2.2).

A2.3. *$U_i, i = 1, 2,$ are compact subsets of some Euclidean space, and $c_i \geq 0$.*

A2.4. *The functions* $k_i(\cdot)$ *and* $b_i(\cdot)$ *are real-valued (resp.,* $\mathbb{R}^r$*-valued) and continuous on* $G \times U_i$*. Let* $\sigma(\cdot)$ *be a Lipschitz continuous matrix-valued function on* G*, with* r *rows and with the number of columns being the dimension of the Wiener process in* (2.2). *The* $b_i(\cdot, \alpha_i)$ *are Lipschitz continuous, uniformly in* α_i.

Later, the continuity and Lipschitz conditions in A2.4 will be replaced by A2.5 and either A2.6 or A2.7, and then we will be concerned with weak-sense solutions.

Comments on A2.1 and A2.2. Condition A2.2 is unrestrictive since one can always construct the extension. That A2.1 is quite natural can be seen from the following comments. First, suppose that the state space is being bounded for purely numerical reasons. Then the reflections are introduced merely to give a compact set G, which should be large enough so that the effects on the solution in the region of main interest are small. Then one often uses a hyperrectangle with normal reflection directions, in which case the right side of (2.1) is zero. Next, consider a heavy traffic queueing network model [22, 28, 39] where the state space is the nonnegative orthant, and the probability that an output of the ith processor goes to the queue for the jth processor is q_{ij}. Define the routing matrix $Q = \{q_{ij}; i, j\}$. If the spectral radius of Q is less than unity, then all customers will eventually leave the system, with probability one. The model is a special case of (2.2), and we can write $z(t) = [I - Q']y(t)$, where $y_i(\cdot)$ is nondecreasing, continuous, and can increase only at t, where $x_i(t) = 0$. In this case, A2.1 implies (see [11, 28]) the so-called completely S condition [22, 28, 38], which is essential to ensure important properties of the representation (2.2); for example, that $z(\cdot)$ has bounded variation with probability one. Also, A2.1 implies the Lipschitz condition and bound in Theorem 4.2.

The system model. Let $w(\cdot)$ be a standard vector-valued Wiener process with respect to a filtration $\{\mathcal{F}_t, t < \infty\}$, which might depend on the control. Let $u_i(\cdot), i = 1, 2$, be U_i-valued, measurable, and $\mathcal{F}_t$-adapted processes. Such processes are to be called *admissible controls.*[1] Keep in mind that the mere fact that $u_i(\cdot), i = 1, 2$, are admissible does not imply that they are acceptable controls for the game since the two players will have different information available depending on who "goes first." Furthermore, controls for the game are defined in terms of "strategies," as discussed at the end of this section. Nevertheless, for any controls with the correct information dependencies, there will be a filtration with respect to which $w(\cdot)$ is a standard vector-valued Wiener process, and to which the controls are adapted. The concept of admissibility will be used in getting useful approximations and bounds.

The dynamical model for the game process is the reflected SDE

$$x(t) = x(0) + \sum_{i=1}^{2} \int_0^t b_i(x(s), u_i(s))ds + \int_0^t \sigma(x(s))dw(s) + z(t), \tag{2.2}$$

where $u_i(\cdot)$ is the control for player i, $i = 1, 2$. The process $z(\cdot)$ is due to the boundary reflections and ensures that $x(t) \in G$. It has the representation

$$z(t) = \sum_i d_i y_i(t), \tag{2.3}$$

where $y(0) = 0$ and the $y_i(\cdot)$ are continuous, nondecreasing, and can increase only at t, where $x(t)$ is on the ith face of ∂G. The condition (2.1) implies that the set of reflection directions on any set of intersecting boundary faces are linearly independent.

[1] They will sometimes be referred to as admissible ordinary controls to distinguish them from relaxed controls.

This implies that the representation (2.3) is unique. See [28, Chapter 3] or [11, 21, 32] for a discussion of equations such as (2.2).

Relaxed controls $r_i(\cdot)$. Suppose that, for some filtration $\{\mathcal{F}_t, t < \infty\}$ and some standard vector-valued $\mathcal{F}_t$-Wiener process $w(\cdot)$, each $r_i(\cdot), i = 1, 2$, is a measure on the Borel sets of $U_i \times [0,\infty)$ such that $r_i(U_i \times [0,t]) = t$ and $r_i(A \times [0,t])$ is $\mathcal{F}_t$-measurable for each Borel set $A \subset U_i$. Then $r_i(\cdot)$ is said to be an *admissible relaxed control* for player i, with respect to $w(\cdot)$. If the Wiener process and filtration have been given or are obvious or unimportant, we simply say that $r_i(\cdot)$ is an admissible relaxed control for player i [15, 32]. For Borel sets $A \subset U_i$, we will write $r_i(A \times [0,t]) = r_i(A, t)$.

For almost all (ω, t) and each Borel set $A \subset U_i$, one can define the derivative

$$r_{i,t}(A) = \lim_{\delta \to 0} \frac{r_i(t, A) - r_i(t - \delta, A)}{\delta}.$$

Without loss of generality, we can suppose that the limit exists for all (ω, t). Then, for all (ω, t), $r_{i,t}(\cdot)$ is a probability measure on the Borel sets of U_i, and, for any bounded Borel set B in $U_i \times [0,\infty)$,

$$r_i(B) = \int_0^\infty \int_{U_i} I_{\{(\alpha_i,t)\in B\}} r_{i,t}(d\alpha_i)dt.$$

An ordinary control $u_i(\cdot)$ can be represented in terms of the relaxed control $r_i(\cdot)$, defined by its derivative $r_{i,t}(A) = I_A(u_i(t))$, where $I_A(u_i)$ is unity if $u_i \in A$ and is zero otherwise. The weak topology [32] will be used on the space of admissible relaxed controls. Relaxed controls are commonly used in control theory to prove existence theorems since any sequence of relaxed controls has a convergent subsequence.

Define the relaxed control $r(\cdot) = (r_1(\cdot) \times r_2(\cdot))$, with derivative $r_t(\cdot) = r_{1,t}(\cdot) \times r_{2,t}(\cdot)$. The $r(\cdot)$ is a measure on the Borel sets of $(U_1 \times U_2) \times [0,\infty)$, with marginals $r_i(\cdot), i = 1, 2$. Sometimes we will just write $r(\cdot) = (r_1(\cdot), r_2(\cdot))$ without ambiguity. The pair $(w(\cdot), r(\cdot))$ is called an *admissible pair* if each of the $r_i(\cdot)$ is admissible with respect to $w(\cdot)$.

In relaxed control terminology, (2.2) is written as

$$x(t) = x(0) + \sum_{i=1}^{2} \int_0^t \int_{U_i} b_i(x(s), \alpha_i) r_{i,s}(d\alpha_i)ds + \int_0^t \sigma(x(s))dw(s) + z(t). \tag{2.4}$$

The existence and uniqueness of solutions to (2.4) will be discussed in the next section. Until section 8, for $x(0) = x$ and $\beta > 0$, the cost function is

$$W(x, r_1, r_2) = E \int_0^\infty e^{-\beta t} \left[\sum_{i=1}^{2} \int_{U_i} k_i(x(s), \alpha_i) r_{i,t}(d\alpha_i)dt + c'dy(t) \right]. \tag{2.5}$$

Define $\alpha = (\alpha_1, \alpha_2)$, $u = (u_1, u_2)$, and $b(x, \alpha) = b_1(x, \alpha_1) + b_2(x, \alpha_2)$, and define $k(\cdot)$ analogously.

Weak-sense solution. Suppose that $(w(\cdot), r(\cdot))$ is admissible with respect to some filtration $\{\mathcal{F}_t, t < \infty\}$ on some probability space. If there is a probability space on which are defined a filtration $\{\tilde{\mathcal{F}}_t, t < \infty\}$ and an $\tilde{\mathcal{F}}_t$-adapted triple $(\tilde{x}(\cdot), \tilde{w}(\cdot), \tilde{r}(\cdot))$, where $(\tilde{w}(\cdot), \tilde{r}(\cdot))$ is admissible and has the same probability law as $(w(\cdot), r(\cdot))$, and the triple satisfies (2.4), then it is said that there is a *weak-sense* solution to (2.4) for $(w(\cdot), r(\cdot))$. (The associated reflection process $\tilde{z}(\cdot)$ is determined by $(\tilde{x}(\cdot), \tilde{w}(\cdot), \tilde{r}(\cdot))$.)

Unique weak-sense solution. Suppose that we are given two probability spaces (indexed by $i = 1, 2$) with filtrations $\{\mathcal{F}_t^i, t < \infty\}$ and on which are defined processes $(x^i(\cdot), w^i(\cdot), r^i(\cdot))$, where $w^i(\cdot)$ is a standard vector-valued $\mathcal{F}_t^i$-Wiener process, $(w^i(\cdot), r^i(\cdot))$ is an admissible pair, and $(x^i(\cdot), w^i(\cdot), r^i(\cdot))$ solves (2.4). If equality of the probability laws of $(w^i(\cdot), r^i(\cdot))$, $i = 1, 2$, implies equality of the probability laws of $(x^i(\cdot), w^i(\cdot), r^i(\cdot))$, $i = 1, 2$, then we say that there is a *unique weak-sense solution* to (2.4) for the admissible pair $(w^i(\cdot), r^i(\cdot))$.

When working with weak-sense solutions, condition A2.5 and either A2.6 or A2.7 will replace A2.4.

A2.5. *The functions $\sigma(\cdot), b_i(\cdot), k_i(\cdot), i = 1, 2$, are bounded and measurable. Equation* (2.4) *has a unique weak-sense solution for each admissible pair* $(w(\cdot), r(\cdot))$ *and each initial condition.*

A2.6. *The functions $\sigma(\cdot), b_i(\cdot)$, and $k_i(\cdot), i = 1, 2$, are continuous.*

In A2.7, let $(w(\cdot), r(\cdot))$ be an arbitrary admissible pair, and let $x(\cdot)$ be the corresponding solution. Condition A2.7 differs from A2.6 in that the dynamics can be discontinuous, provided that not much time is spent in a small neighborhood of the set of discontinuity. A "threshold" control example where A2.7 holds is where $\sigma(x)\sigma'(x)$ is uniformly positive definite in G, $b(x, \alpha) = \bar{b}(x, \alpha) + b_0(x)$, where $\bar{b}(\cdot) = \sum_i \bar{b}_i(\cdot)$ and $k_i(\cdot), \bar{b}_i(\cdot)$, and $\sigma(\cdot)$ are continuous, and $b_0(x)$ takes one of two values, depending on which side of a hyperplane x lies.

A2.7. *There is a Borel set $D_d \subset G$ such that $\sigma(\cdot), b_i(\cdot)$, and $k_i(\cdot), i = 1, 2$, are continuous when $x \notin D_d$, and, for each $\epsilon > 0$, there is $t_\epsilon > 0$, which goes to zero as $\epsilon \to 0$, and such that, for any real T,*

$$\lim_{\epsilon \to 0} \sup_{x(0)} \sup_{admis.\ r(\cdot)} \sup_{t_\epsilon \le t \le T} P\{x(t) \in N_\epsilon(D_d)\} = 0,$$

where $N_\epsilon(D_d)$ is an ϵ-neighborhood of D_d.

Comment on the Girsanov transformation method for defining solutions. When there is not a uniform Lipschitz condition (i.e., A2.3 does not hold), a common and useful approach to modeling uses the Girsanov transformation method [9, 23, 28, 32]. Here one starts with either a unique strong- or weak-sense solution and then introduces the control by a change of measure. Under appropriate conditions, the transformation is used to "shift" the drift term so that it includes the desired control. This procedure does not change the filtration or the probability space, but it does change the Wiener process. The new solution will also be weak-sense unique. See the references for more detail.

Classes of controls and strategies. Definitions. Let $\{\mathcal{F}_t, t < \infty\}$ be a filtration, and let $w(\cdot)$ be a standard vector-valued $\mathcal{F}_t$-Wiener process. Let $\mathcal{U}_i$ denote the set of controls $u_i(\cdot)$ for player i that are admissible with respect to $w(\cdot)$. For $\Delta > 0$, let $\mathcal{U}_i(\Delta) \subset \mathcal{U}_i$ denote the subset of admissible controls $u_i(\cdot)$ that are constant on the intervals $[k\Delta, k\Delta + \Delta)$, $k = 0, 1, \dots$, and where $u_i(k\Delta)$ is $\mathcal{F}_{k\Delta}$-measurable. Let B be a Borel subset of U_1. Let $\mathcal{L}_1(\Delta)$ denote the set of such piecewise constant controls for player 1 that are represented by functions $Q_{1k}(B; \cdot), k = 0, 1, \dots$, of the conditional probability type

$$\begin{aligned} P\left\{u_1(k\Delta) \in B \middle| w(s), u_2(s), s < k\Delta; u_1(l\Delta), l < k\right\} \\ = Q_{1k}\left(B; w(s), u(s), s < k\Delta\right), \end{aligned} \tag{2.6}$$

where $Q_{1k}(B; \cdot)$ is a measurable function for each Borel set B. Controls determined by (2.6) can be called strategies, owing to their explicit dependence on the past actions of both players.

If a rule for player 1 is given by the form (2.6), then, in the arguments of the cost functions, it will sometimes be written as $u_1(u_2)$ to emphasize its dependence on $u_2(\cdot)$. Although there is also dependence on $w(\cdot)$, that dependence is suppressed in the notation. Define $\mathcal{L}_2(\Delta)$ and the associated rules $u_2(u_1)$ for player 2 analogously. The same terminology will be used for relaxed controls. Thus $r_i(\cdot) \in \mathcal{U}_i$ means that $r_i(\cdot)$ is admissible, $r_i(\cdot) \in \mathcal{U}_i(\Delta)$ means that $r_i(\cdot)$ is admissible, the derivative $r_{i,t}(\cdot)$ is constant on the intervals $[k\Delta, k\Delta + \Delta)$, and $r_{i,t}(\cdot)$ is $\mathcal{F}_{k\Delta}$-measurable. Thus the difference between $\mathcal{L}_i(\Delta)$ and $\mathcal{U}_i(\Delta)$ is that, in the former case, the control is determined by a conditional probability law such as (2.6). However, the uniqueness condition A2.5 implies that it is only the probability law of $(w(\cdot), u_1(\cdot), u_2(\cdot))$ (or, more generally, of $(w(\cdot), r_1(\cdot), r_2(\cdot))$) that determines the law of the solution and hence the value of the cost. Thus we can always suppose that if the control for, say, player 1 is determined by a form such as (2.6), then (in relaxed control terminology) the law for $(w(\cdot), r_2(\cdot))$ is determined recursively by a conditional probability law

$$P\left\{\{w(s), r_2(s), k\Delta \le s \le k\Delta + \Delta\} \in \cdot \middle| w(s), r_2(s), u_1(s), s < k\Delta\right\}.$$

Theorems 4.5–4.7 imply that the values defined by (2.7) and (2.8) would not change if admissible relaxed controls were used in lieu of admissible ordinary controls.

Upper and lower values. For initial condition $x(0) = x$, define the upper and lower values for the game as

$$V^+(x) = \lim_{\Delta \to 0} \inf_{u_1 \in \mathcal{L}_1(\Delta)} \sup_{u_2 \in \mathcal{U}_2} W(x, u_1(u_2), u_2), \tag{2.7}$$

$$V^-(x) = \lim_{\Delta \to 0} \sup_{u_2 \in \mathcal{L}_2(\Delta)} \inf_{u_1 \in \mathcal{U}_1} W(x, u_1, u_2(u_1)). \tag{2.8}$$

Discussion of (2.7), (2.8). Let us interpret (2.7). For fixed $\Delta > 0$, consider the right side of (2.7). For each k, at time $k\Delta$, player 1 uses a rule of the form (2.6) to decide on the constant action that it will take on $[k\Delta, k\Delta + \Delta)$. That is, it "goes first." Player 2 can decide on its action at $t \in [k\Delta, k\Delta + \Delta)$ at the actual time that it is to be applied. (Its choice for the discrete instants $k\Delta$ is irrelevant.) Thus player 2 "goes last." Player 2 selects its strategy simply to be admissible. This operation yields admissible $u(\cdot) = (u_1(\cdot), u_2(\cdot))$. Under the Lipschitz condition A2.4, there is clearly a unique solution to (2.4). Alternatively, under the weak-sense existence and uniqueness assumption A2.5, there is a probability space on which are defined $(\tilde{w}(\cdot), \tilde{u}(\cdot))$ (with the same distribution as $(w(\cdot), u(\cdot))$) and on which is defined a solution to (2.4). The distribution of the set (solution, Wiener process, control) does not depend on the probability space. Thus, either way, the $\sup_{u_2 \in \mathcal{U}_2}$ is well defined for each rule for player 1. As $\Delta \to 0$, the inf sup is monotonically decreasing since player 1 can make decisions more often. Similar monotonicity was discussed in [20]. The analogous comments hold for (2.8). In section 4, it will be seen that the infs and sups could be taken over the relaxed controls without changing the results. Under our conditions, Theorem 5.1 says that there is a saddle point in that

$$V^+(x) = V^-(x) = V(x) \quad \text{for all } x \in G. \tag{2.9}$$

The use of limits of discrete strategies to define the upper and lower values goes back to [16, 17, 20], where discrete time games were used to approximate continuous time games. The Elliott–Kalton definition [13] does not require discretization and admits the widest class of strategies. However, various approaches based on discretized strategies are shown to yield the same values as those given by the Elliott–Kalton definition (see, for example, [19]). The references [4, 5, 6] all use various discrete time approximations in defining value, similar to (2.7) and (2.8). The numerical approximations converge to the value given by the definition (2.7)–(2.9).

3. The numerical procedure: The Markov chain approximation method.

The Markov chain approximation. Since some facts concerning the Markov chain approximation method of [25, 26, 32] will be needed when dealing with the convergence of the numerical approximation, let us recall the basic numerical procedure for the control problem where there is only one player. Loosely speaking, the method consists of two steps. The first step is the determination of a finite-state controlled Markov chain that has a continuous time interpolation that is an "approximation" of the process $x(\cdot)$. The second step solves the optimization problem for the chain and a cost function that approximates the one used for $x(\cdot)$. Let h denote the approximation parameter. Under a natural "local consistency" condition, the minimal cost function $V^h(x)$ for the controlled approximating chain converges to the minimal cost function for the original problem. The optimal control for the original problem is also approximated. The method is a robust and effective method for solving optimal control problems for reflected-jump-diffusions under very general conditions. The approximating chain and local consistency conditions are the same for the game problems of this paper. There are many methods for getting suitable approximating chains, and the references contain a comprehensive discussion. An advantage of the approach is that the approximations "stay close" to the physical model and can be adjusted to exploit local features. Our main aim is the proof of convergence for the game problem, so only the essential details of the numerical approximations will be given, and the reader is referred to the references for more information.

To construct the approximation, start by defining S_h, a discretization of $\mathbb{R}^r$. This can be done in many ways. For example, S_h might be a regular grid with the distance between points in any coordinate direction being h. The precise requirements, as spelled out below, are quite general. It is only the points in G and their immediate neighbors that will be of interest. The next step is to define the approximating controlled Markov chain ξ_n^h and its state space, which will be a subset of S_h. The state space for the chain is usually divided into two parts. The first part is $G_h = G \cap S_h$, on which the chain approximates the diffusion part of (2.4). If the chain tries to leave G_h, then it is returned immediately, consistently with the local reflection direction. Thus define ∂G_h^+ to be the set of points not in G_h to which the chain might move in one step from some point in G_h. The set ∂G_h^+ is an approximation to the reflecting boundary. This two-step procedure on the boundary simplifies both coding and analysis. In particular, it allows us to introduce a reflection process $z^h(\cdot)$ that is analogous to $z(\cdot)$. This "approximating" reflection process is needed to get the correct form for the limits of the approximating chain and for the components of the cost function that are due to the boundary reflection.

Local consistency on $\boldsymbol{G_h}$. First, we define local consistency at $x \in G_h$. Let $u_n^h = (u_{1,n}^h, u_{2,n}^h)$ denote the controls used at step n for the approximating chain ξ_n^h. Let $E_{x,n}^{h,\alpha}$ (resp., $\mathrm{covar}_{x,n}^{h,\alpha}$) denote the expectation (resp., the covariance), given all of the data to step n, when $\xi_n^h = x, u_n^h = \alpha$. Then the chain satisfies the following condition: There is a function $\Delta t^h(x,\alpha) > 0$ such that

$$
\begin{aligned}
E_{x,n}^{h,\alpha}\left[\xi_{n+1}^h - x\right] &= b(x,\alpha)\Delta t^h(x,\alpha) + o(\Delta t^h(x,\alpha)),\\
\operatorname{covar}_{x,n}^{h,\alpha}\left[\xi_{n+1}^h - x\right] &= a(x)\Delta t^h(x,\alpha) + o(\Delta t^h(x,\alpha)),\ a(x) = \sigma(x)\sigma'(x),\\
\lim_{h\to 0}\sup_{x,\alpha}\Delta t^h(x,\alpha) &= 0,\\
\|\xi_{n+1}^h - \xi_h^n\| &\le K_1 h
\end{aligned}
\tag{3.1}
$$

for some real K_1. With the straightforward methods in [32], $\Delta t^h(\cdot)$ is obtained automatically as a byproduct of getting the transition probabilities, and it will be used as an interpolation interval. Thus, in G, the conditional mean first two moments of $\xi_{n+1}^h - \xi_n^h$ are very close to those of the "differences" of the $x(\cdot)$ of (2.4). The interpolation interval $\Delta t^h(x,\alpha)$ can always be selected so that it does not depend on the control α (or even on the state x), and this is often the choice since it simplifies both the coding and numerical computations.

Remark concerning discontinuous dynamical and cost terms. The consistency condition (3.1) need not hold at all points. For example, consider a case where A2.7 holds: Let $k(\cdot),\sigma(\cdot)$ be continuous, and let $b(\cdot)$ have the form $b(x,\alpha) = b_0(x) + \bar{b}(x,\alpha)$, where $\bar{b}(\cdot)$ is continuous but $b_0(\cdot)$ is discontinuous at $D_d \subset G$. If A2.7 holds for D_d, then we do not need local consistency there. A2.7 would hold if the "noise" $\sigma(x)dw$ "drives" the process away from the set D_d, no matter what the control. See [32, discussion in section 5.5 and Theorem 10.5.3, and also the discussion concerning discontinuous dynamics in section 10.2] for examples and more detail.

Local consistency on the reflecting boundary ∂G_h^+. From points in ∂G_h^+, the transitions of the chain are such that they move to G_h, with the conditional mean direction being a reflection direction at x. More precisely,

$$
\lim_{h\to 0}\sup_{x\in\partial G_h^+}\text{distance}(x, G_h) = 0,
\tag{3.2}
$$

and there are $\theta_1 > 0$ and $\theta_2(h) \to 0$ as $h \to 0$ such that, for all $x \in \partial G_h^+$,

$$
\begin{aligned}
E_{x,n}^{h,\alpha}\left[\xi_{n+1}^h - x\right] &\in \left\{a\gamma : \gamma \in d(x), \theta_2(h) \ge a \ge \theta_1 h\right\},\\
\Delta t^h(x,\alpha) &= 0 \ \text{ for } x \in \partial G_h^+.
\end{aligned}
\tag{3.3}
$$

The last line of (3.3) says that the reflection from states on ∂G_h^+ is instantaneous. Reference [32] has an extensive discussion of straightforward methods of obtaining useful approximations, which can also be used for the game problem.

A cost function. Define $\Delta t_n^h = \Delta t^h(\xi_n^h, u_n^h)$ and $t_n^h = \sum_{l=0}^{n-1}\Delta t_l^h$. When $\xi_n^h \in \partial G_h^+$, we can write (modulo an asymptotically negligible term) $\xi_{n+1}^h - \xi_n^h = \sum_i d_i \delta y_{i,n}^h$, where $\delta y_{i,n}^h \ge 0$ and represents the increments in the direction d_i. The $\delta y_{i,n}^h = 0$ for $\xi_n^h \notin \partial G_h^+$. See also the representation of $z^h(\cdot)$ above (3.9). One choice of discounted cost function for the approximating chain and initial condition $x = x(0)$ is

$$
W^h(x,u^h) = E\sum_{n=0}^{\infty} e^{-\beta t_n^h}\left[k(\xi_n^h, u_n^h)\Delta t_n^h I_{\{\xi_n^h \in G_h\}} + c'\delta y_n^h\right].
\tag{3.4}
$$

Admissible controls and the values. Let $p^h(x,y|u)$ denote the transition probability of the chain for $u = (u_1,u_2)$, $u_1 \in U_1, u_2 \in U_2$. We will define the strategies for the game analogously to what was done in (2.6). If player i goes first, its strategy is defined by a conditional probability law of the type

$$P\left\{u_{i,n}^n \in \cdot \middle| \xi_l^h, l \leq n;\ u_l^h, l < n\right\}.$$

The class of such rules is called $\mathcal{U}_i^h(1)$. If player i goes last, then its strategy is defined by a conditional probability law of the type

$$P\left\{u_{i,n}^n \in \cdot \middle| \xi_l^h, l \leq n, u_l^h, l < n;\ u_{j,n}^h, j \neq i\right\}.$$

The class of such strategies is called $\mathcal{U}_i^h(2)$. Let $\{\delta\tilde{w}_n^h, n < \infty\}$ be mutually independent random variables and such that $\delta\tilde{w}_n^h$ is independent of the "past" $\{\xi_l^h, l \leq n, u_l^h, l < n\}$. For technical reasons, in section 7, the conditioning data might be augmented by $\{\delta\tilde{w}_l^h, l \leq n\}$, but the Markov property

$$P\left\{\xi_{n+1}^h = x \middle| \xi_l^h, u_l^h, l \leq n\right\} = p^h\left(\xi_n^h, x \middle| u_n^h\right)$$

will always hold.

The same notation $\mathcal{U}_i^h(k)$ is used for the admissible relaxed controls. Define the upper values, respectively, as

$$V^{+,h}(x) = \inf_{u_1 \in \mathcal{U}_1^h(1)} \sup_{u_2 \in \mathcal{U}_2^h(2)} W^h(x, u_1, u_2), \tag{3.5}$$

$$V^{-,h}(x) = \sup_{u_2 \in \mathcal{U}_2^h(1)} \inf_{u_1 \in \mathcal{U}_1^h(2)} W^h(x, u_1, u_2). \tag{3.6}$$

In interpreting the cost function and the interpolations to be defined below, keep in mind that $\Delta t^h(x,\alpha) = 0$ for $x \in \partial G_h^+$. For $x \in G_h$, the dynamic programming equation for the upper value is ($\alpha = (\alpha_1, \alpha_2)$)

$$V^{+,h}(x) = \min_{\alpha_1 \in U_1} \{ \max_{\alpha_2 \in U_2} E_x^\alpha [e^{-\beta \Delta t^h(x,\alpha)} V^{+,h}(\xi_1^h) + k(x,\alpha)\Delta t^h(x,\alpha)]\}, \tag{3.7}$$

and, for $x \in \partial G_h^+$, it is

$$V^{+,h}(x) = E_x\left[V^{+,h}(\xi_1^h) + c'\delta y_1^h\right]. \tag{3.8}$$

Here E_x^α denotes the expectation, given initial state x, with control pair α used, and E_x is the expectation, given initial state x (the reflection direction is not controlled). The equations are analogous for the lower value. Owing to the contraction implied by the discounting, there is a unique solution to (3.7) [35]. If desired, the transition probabilities could be constructed so that $\Delta t^h(\cdot)$ does not depend on α, and we have the separated form[2]

$$p^h(x, y|\alpha) = \bar{p}_1(x, y|\alpha_1) + \bar{p}_2(x, y|\alpha_2).$$

Such a form is useful for establishing the existence of a value for the game for the chain [31, 30], but it is not needed for the convergence of the numerical method.

Continuous time interpolation. The chain ξ_n^h is defined in discrete time, but $x(\cdot)$ is defined in continuous time. Only the chain is needed for the numerical

[2]For example, for the latter use, the splitting method of [32, subsection 5.3.2].

computations. However, for the proofs of convergence, the chain must be interpolated into a continuous time process which approximates $x(\cdot)$. The interpolation intervals are suggested by the $\Delta t^h(\cdot)$ in (3.1). We will use a Markovian interpolation, called $\psi^h(\cdot)$. Let $\{\Delta\tau_n^h, n < \infty\}$ be conditionally mutually independent and "exponential" random variables in that

$$P_{x,n}^{h,\alpha}\left\{\Delta\tau_n^h \geq t\right\} = e^{-t/\Delta t^h(x,\alpha)}.$$

Note that $\Delta\tau_n^h = 0$ if ξ_n^h is on the reflecting boundary ∂G_h^+. Define $\tau_0^h = 0$, and, for $n > 0$, set $\tau_n^h = \sum_{i=0}^{n-1} \Delta\tau_i^h$. The τ_n^h will be the jump times of $\psi^h(\cdot)$. Now define $\psi^h(\cdot)$ and the interpolated reflection processes by

$$\psi^h(t) = x(0) + \sum_{\tau_{i+1}^h \leq t} [\xi_{i+1}^h - \xi_i^h],$$

$$Z^h(t) = \sum_{\tau_{i+1}^h \leq t} [\xi_{i+1}^h - \xi_i^h] I_{\{\xi_i^h \in \partial G_h^+\}},$$

$$z^h(t) = \sum_{\tau_{i+1}^h \leq t} E_i^h [\xi_{i+1}^h - \xi_i^h] I_{\{\xi_i^h \in \partial G_h^+\}}.$$

Define the continuous time interpolations $u_i^h(\cdot)$ of the controls analogously. Let $r_i^h(\cdot)$ denote the relaxed control representation of $u_i^h(\cdot)$. The process $\psi^h(\cdot)$ is a continuous time Markov chain. When the state is x and control pair is α, the jump rate out of $x \in G_h$ is $1/\Delta t^h(x,\alpha)$. So the conditional mean interpolation interval is $\Delta t^h(x,\alpha)$; i.e., $E_{x,n}^{h,\alpha}[\tau_{n+1}^h - \tau_n^h] = \Delta t^h(x,\alpha)$.

Define $\tilde{z}^h(\cdot)$ by $Z^h(t) = z^h(t) + \tilde{z}^h(t)$. Note that this representation splits the effects of the reflection into two parts. The first is composed of the "conditional mean" parts $E_i^h[\xi_{i+1}^h - \xi_i^h] I_{\{\xi_i^h \in \partial G_h^+\}}$, and the second is composed of the perturbations about these conditional means [32, section 5.7.9]. The process $z^h(\cdot)$ is a reflection term of the classical type. Both components can change only at t, where $\psi^h(t)$ can leave G_h. Suppose that at some time t, $Z^h(t) - Z^h(t-) \neq 0$, with $\psi^h(t-) = x \in G_h$. Then by (3.3), $z^h(t) - z^h(t-)$ points in a direction in $d(N_h(x))$, where $N_h(x)$ is a neighborhood with radius that goes to zero as $h \to 0$. The process $\tilde{z}^h(\cdot)$ is the "error" due to the centering of the increments of the reflection term about their conditional means and has bounded (uniformly in x, h) second moments, and it converges to zero, as will be seen in Theorem 3.1. By A2.1, A2.2, and the local consistency condition (3.3), we can write (modulo an asymptotically negligible term)

$$z^h(t) = \sum_i d_i y_i^h(t),$$

where $y_i^h(0) = 0$ and $y_i^h(\cdot)$ is nondecreasing and can increase only when $\psi^h(t)$ is arbitrarily close (as $h \to 0$) to the ith face of ∂G.

The interpolated cost criterion. The cost criterion (3.4) can be written (modulo an asymptotically negligible error), where we use relaxed control terminology, $x(0) = x$, and $r_i^h(\cdot)$ is the relaxed control representation of $u_i^h(\cdot)$, as

$$W^h(x, r^h) = E \int_0^\infty e^{-\beta t} \left[\sum_{i=1}^{2} \int_{U_i} k_i(\psi^h(s), \alpha_i) r_{i,t}^h(d\alpha_i) dt + c' dy^h(t) \right]. \tag{3.9}$$

In the numerical computations, the controls are ordinary and not relaxed, but it will be convenient to use the relaxed control terminology when taking limits. The proof of Theorem 7.1 implies that there is $\rho^h \to 0$ as $h \to 0$ such that

$$V^{+,h}(x) \leq V^{-,h}(x) + \rho^h. \tag{3.10}$$

This implies that either the upper or lower numerical game gives an approximation to the original game.

A representation for $\psi^h(\cdot)$. The process $\psi^h(\cdot)$ has a representation which makes it appear close to (2.4) and which is useful in the convergence proofs. Let $\xi_0^h = x$. If $a(\cdot)$ is not uniformly positive definite, then augment the probability space by adding a standard vector-valued Wiener process $\tilde{w}(\cdot)$, where, for each n, $\delta\tilde{w}_{n+1}^h = \tilde{w}(\tau_n^h + \cdot) - \tilde{w}(\tau_n^h)$ is independent of the "past" $\{\psi^h(s), u^h(s), \tilde{w}(s), s \leq \tau_n^h\}$. Then, by [32, sections 5.7.3 and 10.4.1], we can write

$$\begin{aligned} \psi^h(t) = x \ &+ \int_0^t b(\psi^h(s), u^h(s))ds \\ &+ \int_0^t \sigma(\psi^h(s))dw^h(s) + Z^h(s) + \epsilon^h(s), \end{aligned} \tag{3.11}$$

where $\psi^h(t) \in G$. The process $\epsilon^h(\cdot)$ is due to the $o(\cdot)$ terms in (3.1) and is asymptotically unimportant in that, for any T, $\lim_h \sup_{x,r^h} \sup_{s\leq T} E|\epsilon^h(s)|^2 = 0$. The process $w^h(\cdot)$ is a martingale with respect to the filtration induced by $(\psi^h(\cdot), u^h(\cdot), w^h(\cdot))$ and converges weakly to a standard (vector-valued) Wiener process. The $w^h(t)$ is obtained from $\{\psi^h(s), \tilde{w}(s), s \leq t\}$. All of the processes in (3.11) are constant on the intervals $[\tau_n^h, \tau_{n+1}^h)$.

Let $|z^h|(T)$ denote the variation of the process $z^h(\cdot)$ on the time interval $[0, T]$. Then we have the following theorem from [32].

THEOREM 3.1 (Theorem 11.1.3 and (5.7.5) [32]). *Assume* A2.1, A2.2, *and the local consistency conditions, and let $b(\cdot)$ and $\sigma(\cdot)$ be bounded and measurable. Then, for any $T < \infty$, there are $K_2 < \infty$ and δ_h, where $\delta_h \to 0$ as $h \to 0$, and which do not depend on the controls or initial condition, such that*

$$E\left|z^h\right|^2(T) \leq K_2, \tag{3.12}$$

$$E\sup_{s\leq T}\left|\tilde{z}^h(s)\right|^2 = \delta_h E\left|z^h\right|(T). \tag{3.13}$$

The inequalities hold for $y^h(\cdot)$ replacing $z^h(\cdot)$.

4. Auxiliary results: Bounds and approximations. This section is concerned with various estimates and approximations of the solution to (2.4) which are uniform in the control. The proofs of convergence of any numerical approximations involve approximations of the underlying process, especially when control is involved, and the results of this section will be used in section 6 to obtain nearly optimal strategies of a particular type that will play a fundamental role in the convergence proofs of the numerical algorithms. Furthermore, the approximations will be used in section 5 to show that the game has a value. This is critical in showing that the numerical approximations actually converge to the desired value. The approximations imply, among other things, that slight delays in the controls of any of the players affect the costs only slightly. Delaying the control of the second player is equivalent to that player "going first" since its actual applied control at any time will depend on "old" information. This idea will be used in the next section to prove that the game has a value. The first part of the following theorem is [11, Theorem 2.2]. The inequality (4.3) is [32, Theorem 11.1.1].

DEFINITION 4.1 (the Skorohod problem). *Assume* A2.1 *and* A2.2, *and let the components of the $\mathbb{R}^r$-valued function $\psi(\cdot)$ be right continuous and have left-hand limits. Consider the equation $\bar{x}(t) = \psi(t) + \bar{z}(t)$, $x(t) \in G$. Then $\bar{x}(\cdot)$ is said to solve the Skorohod problem* [11, 32] *if the following holds. The components of $\bar{z}(\cdot)$ are right continuous with $\bar{z}(0) = 0$, and $\bar{z}(\cdot)$ is constant on the time intervals where $\bar{x}(t)$ is in the interior of G. The variation $|\bar{z}|(t)$ of $\bar{z}(\cdot)$ on each $[0, t]$ is finite. There is measurable $\gamma(\cdot)$ with values $\gamma(t) \in d(\bar{x}(t))$, the set of reflection directions at $\bar{x}(t)$, such that $\bar{z}(t) = \int_0^t \gamma(s) d|\bar{z}|(s)$. Thus $\bar{z}(\cdot)$ can change only when $\bar{x}(t)$ is on the boundary of G, and then its "increment" is in a reflection direction at $\bar{x}(t)$.*

THEOREM 4.2. *Assume* A2.1 *and* A2.2. *Let $\psi(\cdot) \in D(\mathbb{R}^r; 0, \infty)$, and consider the Skorohod problem $\bar{x}(t) = \psi(t) + \bar{z}(t)$, $x(t) \in G$. Then there is a unique solution $(\bar{x}(\cdot), \bar{z}(\cdot))$ in $D(\mathbb{R}^{2r}; 0, \infty)$. There is $K < \infty$ depending only on the $\{d_i\}$ such that*

$$|\bar{x}(t)| + |\bar{z}(t)| \leq K \sup_{s \leq t} |\psi(s)|, \tag{4.1}$$

and, for any $\psi^i(\cdot) \in D(\mathbb{R}^r; 0, \infty), i = 1, 2$, and corresponding solutions $(\bar{x}^i(\cdot), \bar{z}^i(\cdot))$,

$$|\bar{x}_1(t) - \bar{x}_2(t)| + |\bar{z}_1(t) - \bar{z}_2(t)| \leq K \sup_{s \leq t} |\psi_1(s) - \psi_2(s)|. \tag{4.2}$$

Consider (2.4), *where $b(\cdot)$ and $\sigma(\cdot)$ are bounded and measurable, and use the representation* (2.3) *for the reflection process $z(\cdot)$. Then, for any $T < \infty$, there is a constant K_1 which does not depend on the initial condition or controls and such that*

$$\sup_{x \in G} E \left| y(1) \right|^2 \leq K_1. \tag{4.3}$$

Approximations under the Lipschitz condition A2.4. Suppose that the Lipschitz and continuity condition A2.4 holds. Then the bound (4.1) and Lipschitz condition (4.2) ensure a unique strong-sense solution to the SDE (2.2) or (2.4) for any admissible controls. The proofs of the convergence of the numerical methods and of the existence of the value depend on our ability to approximate the controls. This is simplest under the Lipschitz condition A2.4, and we start with that case. Then the same approximations will be shown to hold if A2.5 and either A2.6 or A2.7 replace A2.4.

For each admissible relaxed control $r(\cdot)$, let $r^\epsilon(\cdot)$ be admissible relaxed controls with respect to the same filtration and Wiener process $w(\cdot)$ and that satisfy

$$\lim_{\epsilon \to 0} \sup_{r_i \in \mathcal{U}_i} E \sup_{t \leq T} \left| \int_0^t \int_{U_i} \phi_i(\alpha_i) \left[r_{i,s}(d\alpha_i) - r^\epsilon_{i,s}(d\alpha_i) \right] ds \right| = 0, \quad i = 1, 2, \tag{4.4}$$

for each bounded and continuous real-valued nonrandom function $\phi_i(\cdot)$ and each $T < \infty$. For future use, note that if (4.4) holds, then it also holds for functions $\phi_i(\cdot)$ of (t, α_i) that are continuous except when t takes some value in a finite set $\{t_i\}$. Let $x(\cdot)$ and $x^\epsilon(\cdot)$ denote the solutions to (2.4) corresponding to $r(\cdot)$ and $r^\epsilon(\cdot)$, respectively, with the same Wiener process used. In particular,

$$x^\epsilon(t) = x(0) + \int_0^t \int_{U_1 \times U_2} b(x^\epsilon(s), \alpha) r^\epsilon_s(d\alpha) ds + \int_0^t \sigma(x^\epsilon(s)) dw(s) + z^\epsilon(t). \tag{4.5}$$

Define

$$\rho^\epsilon(t) = \int_0^t \int_{U_1 \times U_2} b(x(s), \alpha) \left[r_s(d\alpha) - r_s^\epsilon(d\alpha) \right] ds.$$

The processes $x(\cdot), x^\epsilon(\cdot)$, and $\rho^\epsilon(\cdot)$ depend on $r(\cdot)$, but this dependence is suppressed in the notation. The next theorem shows that the set $\{x(\cdot)\}$ over all admissible controls is equicontinuous in probability in the sense that (4.6) holds, and that the costs corresponding to $r(\cdot)$ and $r^\epsilon(\cdot)$ are arbitrarily close for small ϵ, uniformly in $r(\cdot)$.

THEOREM 4.3. *Assume* A2.1 *and* A2.2*, and let $b(\cdot), \sigma(\cdot)$ be bounded and measurable. Then, for each real $\lambda > 0$,*

$$\lim_{\Delta \to 0} \sup_{x(0)} \sup_t \sup_{r_1 \in \mathcal{U}_1} \sup_{r_2 \in \mathcal{U}_2} P \left\{ \sup_{s \leq \Delta} |x(t+s) - x(t)| \geq \lambda \right\} = 0. \tag{4.6}$$

Now add the assumptions A2.3 *and* A2.4*, and let $(r(\cdot), r^\epsilon(\cdot))$ satisfy* (4.4) *for each bounded and continuous $\phi_i(\cdot), i = 1, 2$, and $T < \infty$. Define $\Delta^\epsilon(t) = \sup_{s \leq t} |x(s) - x^\epsilon(s)|^2$. Then, for each t,*

$$\lim_{\epsilon \to 0} \sup_{x(0)} \sup_{r_1 \in \mathcal{U}_1} \sup_{r_2 \in \mathcal{U}_2} E \left| \sup_{s \leq t} \rho^\epsilon(s) \right|^2 = 0, \tag{4.7}$$

$$\lim_{\epsilon \to 0} \sup_{x(0)} \sup_{r_1 \in \mathcal{U}_1} \sup_{r_2 \in \mathcal{U}_2} \left[E\Delta^\epsilon(t) + E \sup_{s \leq t} |z(s) - z^\epsilon(s)|^2 \right] = 0, \tag{4.8}$$

$$\lim_{\epsilon \to 0} \sup_x \sup_{r_1 \in \mathcal{U}_1} \sup_{r_2 \in \mathcal{U}_2} |W(x, r) - W(x, r^\epsilon)| = 0. \tag{4.9}$$

Proof. Assume the conditions in the first sentence of the theorem. Define $\psi(\cdot)$ by

$$\psi(t) = \int_0^t \int_{U_1 \times U_2} b(x(s), \alpha) r_s(d\alpha) ds + \int_0^t \sigma(x(s)) dw(s).$$

Then

$$x(t+\delta) - x(t) = [\psi(t+\delta) - \psi(t)] + [z(t+\delta) - z(t)].$$

By Theorem 4.2, there is $K < \infty$ which does not depend on the control or initial condition and such that

$$\sup_{s \leq \delta} |x(t+s) - x(t)| + \sup_{s \leq \delta} |z(t+s) - z(t)| \leq K \sup_{s \leq \delta} \left[\psi(t+s) - \psi(t) \right].$$

Now using standard estimates for SDEs to evaluate the fourth moments of the right side of the last inequality yields, for some $K_1 < \infty$,

$$\sup_{x(0), t} \sup_{r_1 \in \mathcal{U}_1} \sup_{r_2 \in \mathcal{U}_2} E \sup_{s \leq \delta} |x(t+s) - x(t)|^4 \leq K_1 \delta^2, \tag{4.10}$$

which implies Kolmogorov's criterion for equicontinuity in probability, which is (4.6) [33, Proposition III.5.3]. Write

$$\begin{aligned} x(t) - x^\epsilon(t) \ = \ & \int_0^t \int_{U_1 \times U_2} \left[b(x(s), \alpha) - b(x^\epsilon(s), \alpha) \right] r_s^\epsilon(d\alpha) ds + \rho^\epsilon(t) \\ & + \int_0^t \left[\sigma(x(s)) - \sigma(x^\epsilon(s)) \right] dw(s) + z(t) - z^\epsilon(t). \end{aligned}$$

Now assume the Lipschitz condition A2.4. Then the Lipshitz condition (4.2) together with standard estimates for SDEs, imply that there is a constant K not depending on $(r(\cdot), r^\epsilon(\cdot))$ or on the initial condition $x(0)$ and such that

(4.11)
$$E\Delta^\epsilon(t) \le K\left[E\sup_{s\le t}|\rho^\epsilon(s)|^2 + (t+1)\int_0^t E\Delta^\epsilon(s)ds + E\sup_{s\le t}|z(s)-z^\epsilon(s)|^2\right],$$
$$E\sup_{s\le t}|z(s)-z^\epsilon(s)|^2 \le K\left[E\sup_{s\le t}|\rho^\epsilon(s)|^2 + (t+1)\int_0^t E\Delta^\epsilon(s)ds\right].$$

Suppose that, in the definition of $\rho^\epsilon(\cdot)$, the function $b(x(t),\alpha)$ was replaced by a bounded nonrandom function $\phi(t,\alpha)$, which is continuous except when t takes values in some finite set $\{t_i\}$. Then (4.7) and (4.8) would hold by (4.4) and the use of Gronwall's inequality on the first line of (4.11), after the second line is substituted in to eliminate $z(\cdot)-z^\epsilon(\cdot)$. The equicontinuity in probability (4.6) and the boundedness and continuity of $b(\cdot)$ imply that $b(x(t),\alpha)$ can be approximated arbitrarily well by replacing $x(t)$ by $x(k\mu)$ for $t\in[k\mu, k\mu+\mu), k=0,1,\dots,$ where μ can be chosen independently of $r(\cdot)$. Doing this approximation and using (4.4) imply (4.7) and (4.8).

Now we turn our attention to (4.9). By (4.7), (4.8), and the discounting, the parts of $W(x,r^\epsilon)$ that involve $k(\cdot)$ converge to the corresponding parts of $W(x,r)$. As noted below (2.3), the linear independence of the reflection directions on any set of intersecting boundary faces which is implied by (2.1) implies that $z(\cdot)$ uniquely determines $y(\cdot)$ with probability one. Thus $y^\epsilon(\cdot)$ converges to $y(\cdot)$ with probability one. This convergence, the uniform integrability of the set $\{|y^\epsilon(t+1)-y^\epsilon(t)|; t<\infty, \text{ all } r(\cdot), \epsilon>0\}$ (which is implied by (4.3) and the compactness of G), and the discounting imply that the component of $W(x,r^\epsilon)$ involving $y^\epsilon(\cdot)$ converges to the component of $W(x(0),r)$ involving $y(\cdot)$. □

Weak-sense solutions. The next theorem uses only weak-sense solutions and does not require the Lipschitz condition A2.4. Except for the uniformity assertion, it is a slight variation of [32, Theorem 10.1.2] or, equivalently, of [26, Theorem 3.5.2]. The method of proof, using specially selected probability spaces, is very useful in general when dealing with sequences of solutions that are defined in the weak sense.

THEOREM 4.4. *Assume* A2.1–A2.3, A2.5, *and* A2.6. *Let $r(\cdot)$ and $r^\epsilon(\cdot)$, $\epsilon>0$, be admissible with respect to some Wiener process $w^r(\cdot)$ and satisfy* (4.4). *For each $\epsilon>0$, there is a probability space with an admissible pair $(\tilde w^{r,\epsilon}(\cdot), \tilde r^\epsilon(\cdot))$ which has the same probability law as $(w^r(\cdot), r^\epsilon(\cdot))$ and on which is defined a solution $(\tilde x^{r,\epsilon}(\cdot), \tilde y^{r,\epsilon}(\cdot))$ to* (2.4). *Let $x^r(\cdot)$ denote the solution to* (2.4), *corresponding to $(w^r(\cdot), r(\cdot))$, and let $z^r(\cdot)=\sum_i d_i y_i^r(\cdot)$ denote the associated reflection process. Let $F(\cdot)$ be a bounded and continuous real-valued function on the path space of the canonical set $(x(\cdot), y(\cdot), r(\cdot))$. Then the approximation of the solutions by using $r^\epsilon(\cdot)$ is uniform in that*

$$\lim_{\epsilon\to 0}\sup_{x(0)}\sup_{r_1\in\mathcal{U}_1}\sup_{r_2\in\mathcal{U}_2}\left|EF(\tilde x^{r,\epsilon}(\cdot), \tilde y^{r,\epsilon}(\cdot), \tilde r^\epsilon(\cdot)) - EF(x^r(\cdot), y^r(\cdot), r(\cdot))\right| = 0. \tag{4.12}$$

Now let $F(\cdot)$ be only continuous with probability one with respect to the measure of any solution set $(x(\cdot), y(\cdot), r(\cdot))$. Then, if $(x^n(\cdot), y^n(\cdot), r^n(\cdot))$ converges weakly to $(x(\cdot), y(\cdot), r(\cdot))$, $F(x^n(\cdot), y^n(\cdot), r^n(\cdot))$ converges weakly to $F(x(\cdot), y(\cdot), r(\cdot))$. Also, (4.12) *continues to hold.*

Proof. Let $F(\cdot)$ be bounded and continuous. Let $(w^r(\cdot), r(\cdot))$ be an admissible pair on some probability space, with associated solution process $x^r(\cdot)$ and reflection process $z^r(\cdot) = \sum_i d_i y_i^r(\cdot)$. Since $(w^r(\cdot), r^\epsilon(\cdot))$ is an admissible pair, by the existence part of A2.5, there is a probability space on which are defined processes $(\tilde{x}^{r,\epsilon}(\cdot), \tilde{y}^{r,\epsilon}(\cdot), \tilde{w}^{r,\epsilon}(\cdot), \tilde{r}^\epsilon(\cdot))$, where the last two members are an admissible pair with the same law as $(w^r(\cdot), r^\epsilon(\cdot))$, and $\tilde{x}^{r,\epsilon}(\cdot)$ is the associated weak-sense solution to (2.4), with reflection process $\tilde{z}^{r,\epsilon}(\cdot) = \sum_i d_i \tilde{y}^{r,\epsilon}(\cdot)$. Any weak-sense limit (as $\epsilon \to 0$) $(\tilde{x}^r(\cdot), \tilde{y}^r(\cdot), \tilde{w}^r(\cdot), \tilde{r}(\cdot))$ of the quadruple must solve (2.4), and $\tilde{w}^r(\cdot)$ is a standard vector-valued Wiener process with respect to the filtration generated by $(\tilde{x}^r(\cdot), \tilde{w}^r(\cdot), \tilde{r}(\cdot))$. For a proof of such a characterization of a related limit, see [32, Theorem 11.1.2].

By (4.4) and the weak convergence, $(\tilde{w}^r(\cdot), \tilde{r}(\cdot))$ has the probability law of $(w^r(\cdot), r(\cdot))$. Thus by the uniqueness part of A2.5, $(\tilde{x}^r(\cdot), \tilde{y}^r(\cdot), \tilde{w}^r(\cdot), \tilde{r}(\cdot))$ has the probability law of $(x^r(\cdot), y^r(\cdot), w^r(\cdot), r(\cdot))$. This yields convergence (as $\epsilon \to 0$) in (4.12) for each pair $(w^r(\cdot), r(\cdot))$ and initial condition $x(0)$.

Suppose that the uniformity (in $r(\cdot)$ and $x(0)$) of the convergence in (4.12) does not hold. Then there are $x(0), x_n \to x(0)$, all in G, $\rho > 0$, $\epsilon_n \to 0$, bounded and continuous $F(\cdot)$, and for each n there is a probability space on which are defined an admissible pair $(w^n(\cdot), r^n(\cdot))$, an associated solution $(x^n(\cdot), y^n(\cdot))$, and approximations $r_i^{n,\epsilon_n}(\cdot)$ to $r_i^n(\cdot), i = 1, 2$, satisfying

$$(4.13) \qquad \lim_{n\to\infty} E \sup_{t\le T} \left| \int_0^t \int_{U_i} \phi_i(\alpha_i) \left[r_{i,s}^n(d\alpha_i) - r_{i,s}^{n,\epsilon_n}(d\alpha_i) \right] ds \right| = 0, \quad i = 1, 2,$$

for each bounded and continuous real-valued nonrandom function $\phi_i(\cdot)$ and each $T < \infty$, and with the following additional properties. For each n, there is a probability space on which are defined an admissible pair $(\tilde{w}^{n,\epsilon}(\cdot), \tilde{r}^{n,\epsilon_n}(\cdot))$, which has the law of $(w^n(\cdot), r^{n,\epsilon_n}(\cdot))$, and associated solution $(\tilde{x}^{n,\epsilon_n}(\cdot), \tilde{y}^{n,\epsilon_n}(\cdot))$ with initial conditions $\tilde{x}^{n,\epsilon_n}(0) = x_n$ such that

$$(4.14) \qquad \liminf_n \left| EF(x^n(\cdot), y^n(\cdot), r^n(\cdot)) - EF(\tilde{x}^{n,\epsilon_n}(\cdot), \tilde{y}^{n,\epsilon_n}(\cdot), \tilde{r}^{n,\epsilon_n}(\cdot)) \right| \ge \rho.$$

Equation (4.13) is implied by (4.4).

Now take a weakly convergent subsequence of $\{\tilde{x}^{n,\epsilon_n}(\cdot), \tilde{y}^{n,\epsilon_n}(\cdot), \tilde{w}^{n,\epsilon_n}(\cdot), \tilde{r}^{n,\epsilon_n}(\cdot)\}$, with limit $(\tilde{x}(\cdot), \tilde{y}(\cdot), \tilde{w}(\cdot), \tilde{r}(\cdot))$. Take a weakly convergent subsequence of $\{x^n(\cdot), y^n(\cdot), w^n(\cdot), r^n(\cdot), r^{n,\epsilon_n}(\cdot)\}$ with limit $(x(\cdot), y(\cdot), w(\cdot), r(\cdot), \hat{r}(\cdot))$. By (4.13), $r(\cdot) = \hat{r}(\cdot)$. Also, $(x(\cdot), y(\cdot), w(\cdot), r(\cdot))$ and $(\tilde{x}(\cdot), \tilde{y}(\cdot), \tilde{w}(\cdot), \tilde{r}(\cdot))$ both solve (2.4) with initial condition $x(0)$. Since $(w(\cdot), r(\cdot))$ has the same law as $(w(\cdot), \hat{r}(\cdot))$, hence as $(\tilde{w}(\cdot), \tilde{r}(\cdot))$, the quadruples in the last sentence must be identical in law, which (together with the weak convergence) contradicts (4.14). Thus (4.12) holds.

Now let $F(\cdot)$ be merely bounded and measurable. The first assertion of the last paragraph of the theorem follows from [14, Theorem 3.1(f), Chapter 3]. With this in hand, the uniformity of the convergence in (4.12) is treated as for the case of continuous $F(\cdot)$. □

Finite-valued and piecewise constant approximations $r^\epsilon(\cdot)$ in (4.4): Definitions. Now some approximations of subsequent interest will be defined. They are just piecewise constant and finite-valued ordinary controls. Consider the following discretization of the U_i. Given $\mu > 0$, partition U_i into a finite number of disjoint subsets $C_i^l, l \le p_i$, each with diameter no greater than $\mu/2$. Choose a point $\alpha_i^l \in C_i^l$. Henceforth let p_i be some given function of μ.

Now, given admissible $(r_1(\cdot), r_2(\cdot))$, define the approximating admissible relaxed control $r_i^\mu(\cdot)$ on the control value space $\{\alpha_i^l, l \le p_i\}$ by its derivative as $r_{i,t}^\mu(\alpha_i^l) =$

$r_{i,t}(C_i^l)$. Denote the set of such controls over all $\{C_i^l, \alpha_i^l, l \le p_i\}$ by $\mathcal{U}_i(\mu)$. Let $\mathcal{U}_i(\mu,\delta)$ denote the subset of $\mathcal{U}_i(\mu)$ that are ordinary controls and constant on the intervals $[l\delta, l\delta+\delta), l = 0, 1, \dots$. Another subclass $\mathcal{U}_i(\mu,\delta,\Delta)$ will be defined above Theorem 4.6.

Finite-valued controls. The proof that (4.4) holds in the next two theorems is straightforward and the details are left to the reader. Under the Lipschitz conditions in A2.4, (4.9) follows from Theorem 4.3, and this implies (4.17). Under A2.5 and A2.6, use Theorem 4.4 to get (4.17).

THEOREM 4.5. *Assume* A2.1–A2.3, A2.5, A2.6*, and the above approximation of* $r_i(\cdot)$ *by* $r_i^\mu(\cdot) \in \mathcal{U}_i(\mu), i = 1, 2$*. Then* (4.4) *and* (4.9) *hold for* μ *replacing* ϵ*, no matter what the* $\{C_i^l, \alpha_i^l\}$*. The same result holds if we approximate only one of the* $r_i(\cdot)$*.*

Finite-valued, piecewise-constant, and "delayed" approximations. The proof that the game has a value in section 6 depends on showing that the cost changes little if the controls of any player are "delayed" since that implies that the order in which the players act is not important. Let $r_i^\mu(\cdot) \in \mathcal{U}_i(\mu)$, where the control-space values are $\{\alpha_i^l,\ l \le p_i\}$. Let $\Delta > 0$. Define the "backward" differences $\Delta_{i,k}^l = r_i^\mu(\alpha_i^l, k\Delta) - r_i^\mu(\alpha_i^l, k\Delta - \Delta)$, $l \le p_i$, $k = 1, \dots$. Define the piecewise constant ordinary controls $u_i^{\mu,\Delta}(\cdot) \in \mathcal{U}_i(\mu,\Delta)$ on the interval $[k\Delta, k\Delta+\Delta)$ by

$$u_i^{\mu,\Delta}(t) = \alpha_i^l \ \text{ for } \ t \in \left[k\Delta + \sum_{\nu=1}^{l-1} \Delta_{i,k}^\nu,\ k\Delta + \sum_{\nu=1}^{l} \Delta_{i,k}^\nu\right). \tag{4.15}$$

Note that, on $[k\Delta, k\Delta+\Delta)$, $u^{\mu,\Delta}(\cdot)$ takes the value α_i^l on a time interval of length $\Delta_{i,k}^l$. Note also that the $u_i^{\mu,\Delta}(\cdot)$ are "delayed" in that the values of $r_i(\cdot)$ on $[k\Delta-\Delta, k\Delta)$ determine the values of $u_i^{\mu,\Delta}(\cdot)$ on $[k\Delta, k\Delta+\Delta)$. Thus $u_i^{\mu,\Delta}(\cdot)$ is $\mathcal{F}_{k\Delta-}$-measurable. This delay will play an important role in the next two sections. Let $r_i^{\mu,\Delta}(\cdot)$ denote the relaxed control representation of $u_i^{\mu,\Delta}(\cdot)$.

The intervals $\Delta_{i,k}^l$ in (4.15) are just real numbers. For use in section 6, it is important to have them be some multiple of some small $\delta > 0$, where Δ/δ is an integer. Consider one method of doing this. Divide $[k\Delta, k\Delta+\Delta)$ into Δ/δ subintervals of length δ each. To each value α_i^l first assign (the integer part) $[\Delta_{i,k}^l/\delta]$ subintervals of length δ. Then assign the remaining unassigned subintervals to the values α_i^l at random with probabilities proportional to the residual (unassigned) lengths $\Delta_{i,k}^l - [\Delta_{i,k}^l/\delta]\delta$, $i \le p_i$. Call the resulting control $u_i^{\mu,\delta,\Delta}(\cdot)$, with relaxed control representation $r_i^{\mu,\delta,\Delta}(\cdot)$. Let $\mathcal{U}_i(\mu,\delta,\Delta)$ denote the set of such controls. If $u_i^{\mu,\delta,\Delta}(\cdot)$ is obtained from $r_i(\cdot)$ in this way, then we will henceforth write it as $u_i^{\mu,\delta,\Delta}(\cdot|r_i)$ to emphasize that fact. Similarly, if $u_i^{\mu,\Delta}(\cdot)$ is obtained from $r_i(\cdot)$, then it will be written as $u_i^{\mu,\Delta}(\cdot|r_i)$. Let $r_{i,t}^{\mu,\Delta}(\cdot|r_i)$ denote the time derivative of $r_i^{\mu,\Delta}(\cdot|r_i)$. As stated in the next theorem, for fixed μ and small δ, $u_i^{\mu,\delta,\Delta}(\cdot|r_i)$ well approximates $u_i^{\mu,\Delta}(\cdot|r_i)$ uniformly in $r_i(\cdot)$ and $\{\alpha_i^l\}$ in that (4.4) holds in the sense that, for each $\mu > 0$ $\Delta > 0$ and bounded and continuous $\phi_i(\cdot)$,

$$\lim_{\delta\to 0} \sup_{r_i \in \mathcal{U}_i} E \sup_{t \le T} \left| \int_0^t \int_{U_i} \phi_i(\alpha_i)[r_{i,s}^{\mu,\Delta}(d\alpha_i|r_i) - r_{i,s}^{\mu,\delta,\Delta}(d\alpha_i|r_i)]ds \right| = 0,\ i = 1, 2. \tag{4.16}$$

THEOREM 4.6. *Assume* A2.1–A2.3, A2.5*, and* A2.6*. For* $r_i(\cdot) \in \mathcal{U}_i$*, approximate as above the theorem to get* $r_i^{\mu,\Delta}(\cdot|r_i) \in \mathcal{U}_i(\mu,\Delta)$ *and* $r_i^{\mu,\delta,\Delta}(\cdot|r_i) \in \mathcal{U}_i(\mu,\delta,\Delta)$*. Then* (4.4) *holds for* $r_i^{\mu,\Delta}(\cdot|r_i)$ *and* (μ,Δ) *replacing* $r_i^\epsilon(\cdot)$ *and* ϵ*, respectively. Also,* (4.16) *holds and*

$$\lim_{\Delta\to 0}\lim_{\delta\to 0}\sup_x \sup_{r_1\in\mathcal{U}_1}\sup_{r_2\in\mathcal{U}_2} |W(x,r_1,r_2)-W(x,r_1,u_2^{\mu,\delta,\Delta}(\cdot|r_2))| = 0. \tag{4.17}$$

For each $\epsilon > 0$*, there are* $\mu_\epsilon > 0$ *and* $\delta_\epsilon > 0$ *such that, for* $\mu \le \mu_\epsilon$ *and* $\delta \le \delta_\epsilon$ *and* $r_i(\cdot) \in \mathcal{U}_i, i = 1, 2$*, there are* $u_i^{\mu,\delta}(\cdot) \in \mathcal{U}_i(\mu,\delta)$ *such that* (4.4) *holds for* $u_i^{\mu,\delta}(\cdot)$ *and* (μ,δ) *replacing* $r_i^\epsilon(\cdot)$ *and* ϵ*, respectively, and*

$$\sup_x \sup_{r_1\in\mathcal{U}_1}\sup_{r_2\in\mathcal{U}_2} |W(x,r_1,r_2)-W(x,r_1,u_2^{\mu,\delta})| \le \epsilon. \tag{4.18}$$

The expressions (4.17) *and* (4.18) *hold with the indices* 1 *and* 2 *interchanged.*

THEOREM 4.7. *If* A2.7 *replaces* A2.6 *in Theorems* 4.4–4.6*, then their conclusions continue to hold.*

Proof. Only a few details will be given. The last part of Theorem 4.4 will be used, and we need only identify the $F(\cdot)$ for the present case. Theorem 4.2 and (4.6) required only measurability and boundedness of $b(\cdot)$ and $\sigma(\cdot)$. Also, the tightness of any solution sequence $\{x^\epsilon(\cdot), y^\epsilon(\cdot), w^\epsilon(\cdot), r^\epsilon(\cdot)\}$ requires only the measurability and boundedness of $b(\cdot)$ and $\sigma(\cdot)$.

For each $t > 0$, define the bounded real-valued function $F(\phi(\cdot), m(\cdot))$ on the product path space of $x(\cdot)$ and $r(\cdot)$ by

$$F(\phi(\cdot), m(\cdot)) = \int_0^t \int_{U_1\times U_2} b(\phi(s),\alpha) m_s(d\alpha)ds.$$

Under A2.7, $F(\phi(\cdot), m(\cdot))$ is continuous with probability one with respect to the measure induced by any pair $(x(\cdot), r(\cdot))$ solving (2.4). Let $(x^\epsilon(\cdot), y^\epsilon(\cdot), w^\epsilon(\cdot), r^\epsilon(\cdot))$ satisfy

$$x^\epsilon(t) = x(0) + \int_0^t \int_{U_1\times U_2} b(x^\epsilon(s),\alpha) r^\epsilon(d\alpha)ds + \int_0^t \sigma(x^\epsilon(s))dw^\epsilon(s) + z^\epsilon(t) \tag{4.19}$$

and converge weakly to $(x(\cdot), y(\cdot), w(\cdot), r(\cdot))$ as $\epsilon \to 0$.

For the sake of simplicity and without loss of generality, suppose that the Skorohod representation is used so that all processes are defined on the same probability space and the weak convergence is equivalent to convergence with probability one [14, Theorem 1.8, Chapter 3]. First, suppose that $\sigma(\cdot)$ is continuous but $b(\cdot)$ is not. By the asserted almost everywhere continuity of $F(\cdot)$ and the weak convergence and Skorohod representation, the integral in (4.19) involving $b(\cdot)$ converges to

$$\int_0^t \int_{U_1\times U_2} b(x(s),\alpha) r_s(d\alpha)ds$$

with probability one. Discontinuous $k(\cdot)$ is treated in the same way. Also, the stochastic integral and reflection term converge to those for the limit. The proof for the convergence of the stochastic integral uses a finite sum approximation $\sum_{l\gamma\le t}\sigma(x^\epsilon(l\gamma))\cdot [w^\epsilon(l\gamma+\gamma) - w^\epsilon(l\gamma)]$ and a standard estimate of the errors in this approximation. The proof for the reflection direction is similar to that in [32, Theorem 11.1.2]. The uniformity in $r(\cdot)$ of the approximations, as asserted in (4.9), (4.12), (4.17), and (4.18) in Theorems 4.4–4.6, is shown by a contradiction argument as in Theorem 4.4.

Now, suppose that $\sigma(\cdot)$ is discontinuous but that A2.7 holds. For $\rho > 0$, define the real-valued function $f^\rho(\cdot)$ by

$$f^\rho(x) = \begin{cases} 1 & \text{for } \operatorname{dist}(x, D_d) \ge \rho, \\ \operatorname{dist}(x, D_d)/\rho & \text{otherwise.} \end{cases}$$

The function $\sigma(\cdot)f^\rho(\cdot)$ is continuous, and, for each $T < \infty$, A2.7 implies that

$$\lim_{\rho\to 0} \sup_{x(0),r} \sup_{\epsilon} E \sup_{t\le T} \left| \int_0^t \sigma(x^\epsilon(s)) \left[1 - f^\rho(x^\epsilon(s))\right] dw^\epsilon(s) \right|^2 = 0. \tag{4.20}$$

Now approximate $\sigma(\cdot)$ by $\sigma(\cdot)f^\rho(\cdot)$, and use (4.20) to get the convergence of the stochastic integrals

$$\int_0^t \sigma(x^\epsilon(s))dw^\epsilon(s) \to \int_0^t \sigma(x(s))dw(s). \qquad \square$$

5. Existence of the value of the game.

Motivating the proof that the value exists, i.e., that (2.9) holds. Let player 1 go first, and have a control that is constant on intervals of length Δ_1. Theorems 4.5–4.7 imply that, if the control for player 2 is delayed by more than Δ_1 and discretized in time, then the costs change little for small Δ_1. This delay will mean that player 1 will know player 2's actions before it selects its own. This, in turn, is equivalent to player 2 going first, which (together with the fact that the costs change little) essentially implies that the upper and lower values are as close as we wish. The proof formalizes this idea.

THEOREM 5.1. *Assume* A2.1–A2.3, A2.5, *and either* A2.6 *or* A2.7. *Then the game has a value in that* (2.9) *holds.*

Proof. Let $\Delta_1 > 0$. Let $\Delta, \delta, \mu, \epsilon$ be positive with Δ/δ being an integer. By Theorems 4.5–4.7, for small enough μ, δ, Δ, and large Δ/δ,

$$|W(x, u_1(r_2), r_2) - W(x, u_1(r_2), u_2^{\mu,\delta,\Delta}(\cdot|r_2))| \le \epsilon \tag{5.1}$$

for all $u_1(\cdot) \in \mathcal{L}_1(\Delta_1)$ and $r_2(\cdot) \in \mathcal{U}_2$. Also, for all $\Delta_1 > 0$,

$$\begin{aligned}&\left| \inf_{u_1\in\mathcal{L}_1(\Delta_1)} \sup_{r_2\in\mathcal{U}_2} W(x, u_1(r_2), r_2) - \inf_{u_1\in\mathcal{L}_1(\Delta_1)} \sup_{r_2\in\mathcal{U}_2} W(x, u_1(r_2), u_2^{\mu,\delta,\Delta}(\cdot|r_2)) \right| \\ &\qquad \le \epsilon.\end{aligned} \tag{5.2}$$

The results analogous to (5.1)–(5.2) hold if player 1 goes last. It follows that, in computing the upper or lower values, we can use either relaxed or ordinary controls for the player that goes last.

By the definition (2.7), for each $\Delta_1 > 0$,

$$V^+(x) \le \inf_{u_1\in\mathcal{L}_1(\Delta_1)} \sup_{r_2\in\mathcal{U}_2} W(x, u_1(r_2), r_2). \tag{5.3}$$

Let $\epsilon > 0$. By (5.2), there is $\Delta_\epsilon > 0$ such that, for $\Delta \le \Delta_\epsilon$, there are $\mu > 0$ and $\delta > 0$ such that, for all $\Delta_1 > 0$,

$$\begin{aligned}&\inf_{u_1\in\mathcal{L}_1(\Delta_1)} \sup_{r_2\in\mathcal{U}_2} W(x, u_1(r_2), r_2) \\ &\quad \le \inf_{u_1\in\mathcal{L}_1(\Delta_1)} \sup_{r_2\in\mathcal{U}_2} W(x, u_1(r_2), u_2^{\mu,\delta,\Delta}(\cdot|r_2)) + \epsilon \\ &\quad \le \inf_{u_1\in\mathcal{L}_1(\Delta_1)} \sup_{r_2\in\mathcal{U}_2} W(x, u_1(u_2^{\mu,\delta,\Delta}(\cdot|r_2)), u_2^{\mu,\delta,\Delta}(\cdot|r_2)) + \epsilon.\end{aligned} \tag{5.4}$$

Now let $\Delta_1 < \Delta$ with Δ/Δ_1 being an integer. Recall that the process $u_2^{\mu,\delta,\Delta}(\cdot|r_2)$ is constant on the intervals $[l\delta, l\delta+\delta), l = 0, 1, \dots$, and that $u_2^{\mu,\delta,\Delta}(t|r_2)$ is $\mathcal{F}_{q\Delta}$-measurable for $t \in [q\Delta, q\Delta+\Delta)$. Thus, for integers k and q such that $k\Delta_1 \in [q\Delta, q\Delta+\Delta)$, the inf sup and the use of $u_2^{\mu,\delta,\Delta}(\cdot|r_2)$ on the right of (5.4) can be interpreted to mean that, at each such time $k\Delta_1$, player 1 knows all of player 2's actions on the entire interval $[k\Delta_1, q\Delta+\Delta)$ as well as the data on the "past" up to $k\Delta_1$. Thus one computes the value of the main term on the right side of (5.4) as if player 2 goes first: For $k\Delta_1 \in [q\Delta, q\Delta+\Delta)$, player 1 uses a rule which can be represented in the form

$$P\{u_1(k\Delta_1) \in \cdot \big| u_1(l\Delta_1), l < k;\ u_2^{\mu,\delta,\Delta}(l\delta|r_2), l\delta < q\Delta+\Delta;\ w(s), s < k\Delta_1\}. \tag{5.5}$$

Since it is only the joint probability law that matters, it can be supposed that the value of $u_2(t) = u_2^{\mu,\delta,\Delta}(t|r_2)$ which is actually applied on $[q\Delta, q\Delta+\Delta)$ is determined by a conditional probability law which can be represented in the form

$$P\left\{(u_2(q\Delta+l\delta), l\delta < \Delta) \in \cdot \big| u_2(l\delta), l\delta < q\Delta;\ u_1(s),\ w(s), s < q\Delta\right\}, \tag{5.6}$$

where the $u_2(l\delta)$ take values in a μ-discretization of U_2. Let $\mathcal{L}_2(\mu, \delta, \Delta)$ denote the set of such rules for player 2. The main term on the right side of (5.4) involves arbitrary strategies for player 2 but which are discretized in space and time and delayed. By this fact and the use of the form (5.5), the $u_2^{\mu,\delta,\Delta}(\cdot|r_2)$ can be replaced by a control $u_2(\cdot)$ in $\mathcal{L}_2(\mu,\delta,\Delta)$, and we can write

$$\begin{aligned}
&\inf_{u_1\in\mathcal{L}_1(\Delta_1)} \sup_{r_2\in\mathcal{U}_2} W(x, u_1(u_2^{\mu,\delta,\Delta}(\cdot|r_2)), u_2^{\mu,\delta,\Delta}(\cdot|r_2)) \\
&\qquad = \inf_{u_1\in\mathcal{L}_1(\Delta_1)} \sup_{u_2\in\mathcal{L}_2(\mu,\delta,\Delta)} W(x, u_1(u_2), u_2) \\
&\qquad = \inf_{u_1\in\mathcal{U}_1(\Delta_1)} \sup_{u_2\in\mathcal{L}_2(\mu,\delta,\Delta)} W(x, u_1(u_2), u_2).
\end{aligned}$$

Now, since player 2 can be considered to "go first," we can write

$$\begin{aligned}
&\inf_{u_1\in\mathcal{U}_1(\Delta_1)} \sup_{u_2\in\mathcal{L}_2(\mu,\delta,\Delta)} W(x, u_1(u_2), u_2) \\
&\qquad = \sup_{u_2\in\mathcal{L}_2(\mu,\delta,\Delta)} \inf_{u_1\in\mathcal{U}_1(\Delta_1)} W(x, u_1, u_2(u_1)).
\end{aligned} \tag{5.7}$$

By (5.3), (5.4), and (5.6),

$$V^+(x) \le \sup_{u_2\in\mathcal{L}_2(\mu,\delta,\Delta)} \inf_{u_1\in\mathcal{U}_1(\Delta_1)} W(x, u_1, u_2(u_1)) + \epsilon. \tag{5.8}$$

For small Δ_1 and large Δ/Δ_1,

$$\begin{aligned}
&\left| \sup_{u_2\in\mathcal{L}_2(\mu,\delta,\Delta)} \inf_{u_1\in\mathcal{U}_1(\Delta_1)} W(x, u_1, u_2(u_1)) - \sup_{u_2\in\mathcal{L}_2(\mu,\delta,\Delta)} \inf_{u_1\in\mathcal{U}_1} W(x, u_1, u_2(u_1)) \right| \\
&\qquad \le \epsilon.
\end{aligned}$$

It now follows from this, (5.8), and the definition of $V^-(x)$ that, for small Δ and μ and large Δ/δ and Δ/Δ_1,

$$\begin{aligned}
V^+(x) &\le \sup_{u_2\in\mathcal{L}_2(\mu,\delta,\Delta)} \inf_{u_1\in\mathcal{U}_1(\Delta_1)} W(x, u_1, u_2(u_1)) + \epsilon \\
&\le \sup_{u_2\in\mathcal{L}_2(\mu,\delta,\Delta)} \inf_{u_1\in\mathcal{U}_1} W(x, u_1, u_2(u_1)) + \epsilon \\
&\le \sup_{u_2\in\mathcal{U}_2(\delta)} \inf_{u_1\in\mathcal{U}_1} W(x, u_1, u_2(u_1)) + \epsilon \le V^-(x) + 2\epsilon.
\end{aligned} \tag{5.9}$$

Since ϵ is arbitrary and $V^+(x) \ge V^-(x)$, the theorem is proved. $\square$

6. An auxiliary result: Nearly optimal policies. The proof of convergence of the numerical method in the next section will require the use of particular ϵ-optimal minimizing (resp., maximizing) strategies for player 1 when it goes first (resp., for player 2 when it goes first). Such strategies will be constructed in this section. They are for mathematical purposes only and do not have any practical value otherwise.

In order to motivate the construction, we will first recall the method of proof used for the pure control problem (where there is only a minimizing player) in [32, Chapters 10 and 11]. Let $r^h(\cdot)$ denote the continuous time interpolation of the relaxed control representation of the optimal control for the approximating chain ξ_n^h. Thus the optimal cost $V^h(x)$ equals $W^h(x, r^h)$. The corresponding set $\{\psi^h(\cdot), z^h(\cdot), w^h(\cdot), r^h(\cdot)\}$ was shown to be tight. The limit $(x(\cdot), z(\cdot), w(\cdot), r(\cdot))$ of any weakly convergent subsequence was shown to satisfy the (one-player form of) (2.4). Hence it cannot be better than an optimal solution for (2.4). This implies that $\liminf_h V^h(x) \geq V(x)$, the minimal value of the cost for (2.4).

To finish the convergence proof in [32], we had to show that $\limsup_h V^h(x) \leq V(x)$. This was done in the following way. Given arbitrary $\epsilon > 0$, a special ϵ-optimal control for (2.4) was constructed. This special control was such that it could be adapted for use on the approximating chain, and for small h the interpolated chain well approximated the limit process under that control. In more detail, let $r^\epsilon(\cdot)$ denote the relaxed control form of this special ϵ-optimal control for (2.4), with Wiener process $w^\epsilon(\cdot)$ and associated solution and reflection process $(x^\epsilon(\cdot), z^\epsilon(\cdot))$. Let $r^{\epsilon,h}(\cdot)$ denote the relaxed control form of the adaptation of this special control for use on the chain ξ_n^h, interpolated to continuous time, and let $(\psi^{\epsilon,h}(\cdot), z^{\epsilon,h}(\cdot), w^{\epsilon,h}(\cdot))$ denote the continuous time interpolation of the corresponding solution, reflection process and "pre-Wiener" process in the representation (3.11). Since $r^{\epsilon,h}(\cdot)$ is no better than the optimal control for the chain, $V^h(x) \leq W^h(x, r^{\epsilon,h})$. By the method of construction of $r^{\epsilon,h}(\cdot)$, the set $(\psi^{\epsilon,h}(\cdot), z^{\epsilon,h}(\cdot), w^{\epsilon,h}(\cdot), r^{\epsilon,h}(\cdot))$ converged weakly to the set $(x^\epsilon(\cdot), z^\epsilon(\cdot), w^\epsilon(\cdot), r^\epsilon(\cdot))$, with ϵ-optimal cost $W(x, r^\epsilon)$. Since ϵ is arbitrary, we have $\limsup_h V^h(x) \leq V(x)$, which completes the proof that $V^h(x) \to V(x)$. See the references for full details.

Such an ϵ-optimal control for (2.4) (whether minimizing or maximizing) for the player that goes first plays a similar role for the game problem of this paper. The construction follows the general lines of what was done in [32, Theorem 10.3.1], but there are some very important differences since we must work with strategies, where the two controls depend on each other, which is not the case for the pure (i.e., one-player) control problem. The construction is done as it is since we know little about nearly optimal policies in general. For example, we do know whether there are smooth ϵ-optimal feedback controls for either player, in general.

THEOREM 6.1. *Assume* A2.1–A2.3, A2.5, *and either* A2.6 *or* A2.7. *Let player* 1 *go first. Then, for each* $\epsilon > 0$, *there is an* ϵ*-optimal minimizing control law for player* 1 *with the following properties. For positive* Δ, δ, *and* ρ, *let* δ/ρ *and* Δ/δ *be integers. The control is constant on the intervals* $[k\Delta, k\Delta + \Delta)$, $k = 0, 1, \dots$, *finite-valued, the value at* $k\Delta$ *is* $\mathcal{F}_{k\Delta}$*-measurable, and, for small* $\lambda > 0$, *it is defined by the conditional probability law* (*which defines the function* $q_{i,k}(\cdot)$)

$$
\begin{aligned}
&P\left\{u_1(k\Delta) = \gamma \,\middle|\, u_1(l\Delta), l < k;\, w(s), r_2(s), s < k\Delta\right\} \\
&= P\{u_1(k\Delta) = \gamma \big| w(l\lambda), l\lambda < k\Delta;\, u_1(l\Delta), l < k;\, u_2^{\mu,\rho,\delta}(l\rho|r_2), l\rho < k\Delta\} \qquad (6.1)\\
&= q_{1,k}(\gamma;\, w(l\lambda), l\lambda < k\Delta;\, u_1(l\Delta), l < k;\, u_2^{\mu,\rho,\delta}(l\rho|r_2), l\rho < k\Delta).
\end{aligned}
$$

The function $q_{1k}(\cdot)$ *is continuous in the* w*-arguments for each value of the others. Since the rule* (6.1) *depends on* $r_2(\cdot)$ *only via* $u_2^{\mu,\rho,\delta}(\cdot|r_2)$ (*which is defined above*

Theorem 4.6), *we write the rule as* $\bar{u}_1^\epsilon(u_2^{\mu,\rho,\delta}(\cdot|r_2))$. *In particular, for small* λ, μ, Δ *and large* δ/ρ *and* Δ/δ, *it satisfies the inequality*

$$\sup_{r_2\in\mathcal{U}_2} W(x, \bar{u}_1^\epsilon(u_2^{\mu,\rho,\delta}(\cdot|r_2)), r_2) \leq V(x) + \epsilon. \tag{6.2}$$

Also, if $r_2^n(\cdot)$ *is a sequence which converges weakly to some* $r_2(\cdot)$, *then*

$$\limsup_n W(x, \bar{u}_1^\epsilon(u_2^{\mu,\rho,\delta}(\cdot|r_2^n)), r_2^n) \leq V(x) + \epsilon. \tag{6.3}$$

For each $r_2(\cdot)$ *and* $l = 0, 1, \ldots$, *let* $\tilde{u}_2^{\mu,\rho,\delta}(l\rho|r_2)$ *be a control that differs from* $u_2^{\mu,\rho,\delta}(l\rho|r_2)$ *by at most* μ *in absolute value. Then* (6.2) *and* (6.3) *hold for the perturbation* $\tilde{u}_2^{\mu,\rho,\delta}(\cdot|r_2)$ *replacing* $u_2^{\mu,\rho,\delta}(\cdot|r_2)$.

Similarly, if player 2 *goes first, then there is an* ϵ*-optimal control rule of the same type: In particular, and with the analogous terminology,*

$$\inf_{r_1\in\mathcal{U}_1} W(x, r_1, \bar{u}_2^\epsilon(u_1^{\mu,\rho,\delta}(\cdot|r_1))) \geq V(x) - \epsilon, \tag{6.4}$$

and (6.4) *continues to hold with the perturbation* $\tilde{u}_1^{\mu,\rho,\delta}(\cdot|r_1)$ *replacing* $u_1^{\mu,\rho,\delta}(\cdot|r_1)$.

Proof. Recall the approximation of the U_i given above Theorem 4.5: Given $\mu_1 > 0$, U_i was partitioned into a finite number of disjoint subsets $C_i^l, l \leq p_i$, each with diameter no greater than $\mu_1/2$. A point α_i^l in each C_i^l was chosen. These are the values of γ in (6.1). Let player 1 go first. Given $\epsilon > 0$, there are $\Delta > 0, \mu_1 > 0$, and an $\epsilon/8$-optimal rule for player 1 which can be represented in the "conditional probability" form

$$P\left\{u_1(k\Delta) = \gamma \middle| u_1(l\Delta), l < k;\, w(s), r_2(s), s \leq k\Delta\right\}. \tag{6.5}$$

Call the rule $u_1^\epsilon(r_2)$. Then, by the $\epsilon/8$-optimality, for all $r_2(\cdot)$, we have

$$W(x, u_1^\epsilon(r_2), r_2) \leq V^+(x) + \epsilon/8. \tag{6.6}$$

The rule (6.5) needs to be approximated so that it depends only on selected samples of the data.

Whatever $r_2(\cdot)$, $u_2^{\mu,\rho,\delta}(\cdot|r_2)$ is also an admissible control. Hence, by (6.6), for all $r_2(\cdot) \in \mathcal{U}_2$,

$$W(x, u_1^\epsilon(u_2^{\mu,\rho,\delta}(\cdot|r_2)), u_2^{\mu,\rho,\delta}(\cdot|r_2)) \leq V^+(x) + \epsilon/8. \tag{6.7}$$

Indeed, (6.7) holds for all $u_2^{\mu,\rho,\delta}(\cdot|r_2)$, irrespective of $r_2(\cdot)$. Let μ, ρ, and δ be positive numbers with Δ/δ and δ/ρ being integers. For small μ and δ and large δ/ρ, Theorems 4.5–4.7 imply that

$$W(x, u_1^\epsilon(u_2^{\mu,\rho,\delta}(\cdot|r_2)), r_2) \leq W(x, u_1^\epsilon(u_2^{\mu,\rho,\delta}(\cdot|r_2)), u_2^{\mu,\rho,\delta}(\cdot|r_2)) + \epsilon/8 \tag{6.8}$$

for all $r_2(\cdot) \in \mathcal{U}_2$. The control law $u_1^\epsilon(u_2^{\mu,\rho,\delta}(\cdot|r_2))$ can be represented in the form

$$\begin{aligned} &P\{u_1(k\Delta) = \gamma \big| u_1(l\Delta), l < k;\, r_2(s), w(s), s < k\Delta\} \\ &= P\{u_1(k\Delta) = \gamma \big| w(s), s < k\Delta\ \ u_1(l\Delta); l < k;\, u_2^{\mu,\rho,\delta}(l\rho|r_2), l\rho < k\Delta\}. \end{aligned} \tag{6.9}$$

Let $r_2^n(\cdot)$ converge to $r_2(\cdot)$ as $n \to \infty$. Then the discrete approximations $u_2^{\mu,\rho,\delta}(\cdot|r_2^n)$ do not necessarily converge to $u_2^{\mu,\rho,\delta}(\cdot|r_2)$ as $n \to \infty$. They will converge if the limit "mass" on the boundaries of the sets $\{C_2^m, m \le p_2\}$ into which U_2 is subdivided is zero—more particularly, if

$$r_2(l\delta + \delta, \partial C_2^m) - r_2(l\delta, \partial C_2^m) = 0$$

for all m, l [14, Chapter 3, Theorem 3.1(f)]. Such convergence is hard to ensure for arbitrary $r_2(\cdot)$. The problem is due to the fact that the sets C_2^n are not all closed so that part of the boundary of some set will actually be in a neighboring set. Owing to our use of a μ-discretization of the U_i, the worst that can happen is that $u_2^{\mu,\rho,\delta}(l\rho|r_2^n)$ will differ from $u_2^{\mu,\rho,\delta}(l\rho|r_2)$ by at most μ for each l in the limit as $n \to \infty$. For each $r_2(\cdot)$, let $\tilde{u}_2^{\mu,\rho,\delta}(\cdot|r_2)$ be any admissible control satisfying

$$\sup_l |u_2^{\mu,\rho,\delta}(l\rho|r_2) - \tilde{u}_2^{\mu,\rho,\delta}(l\rho|r_2)| \le \mu.$$

Then, as $n \to \infty$ and $r_2^n(\cdot) \to r_2(\cdot)$, we will have $u_1^\epsilon(u_2^{\mu,\rho,\delta}(\cdot|r_2^n)) \to u_1^\epsilon(\tilde{u}_2^{\mu,\rho,\delta}(\cdot|r_2))$ for some perturbation $\tilde{u}_2^{\mu,\rho,\delta}(\cdot|r_2)$ that differs from $\tilde{u}_2^{\mu,\rho,\delta}(\cdot|r_2)$ by a most μ at each time point. For small μ and large δ/ρ, it will be seen that inequality (6.12) holds, and that is all that will be needed.

For small μ and δ and large δ/ρ, (6.8) yields

$$W(x, u_1^\epsilon(\tilde{u}_2^{\mu,\rho,\delta}(\cdot|r_2)), r_2) \le W(x, u_1^\epsilon(\tilde{u}_2^{\mu,\rho,\delta}(\cdot|r_2)), \tilde{u}_2^{\mu,\rho,\delta}(\cdot|r_2)) + \epsilon/8 \tag{6.10}$$

for all $r_2(\cdot) \in \mathcal{U}_2$. Inequalities (6.10) and (6.6) (with all $r_2(\cdot)$ replaced by $\tilde{u}_2^{\mu,\rho,\delta}(\cdot|r_2)$) imply that, for all $r_2(\cdot)$ and all such perturbations $\tilde{u}_2^{\mu,\rho,\delta}(\cdot|r_2)$,

$$W(x, u_1^\epsilon(\tilde{u}_2^{\mu,\rho,\delta}(\cdot|r_2)), r_2) \le V^+(x) + 2\epsilon/8. \tag{6.11}$$

Hence, for small μ and δ and large δ/ρ, the rule (6.9), but with any perturbation $\tilde{u}_2^{\mu,\rho,\delta}(\cdot|r_2)$ used in lieu of $u_2^{\mu,\rho,\delta}(\cdot|r_2)$, still yields a $2\epsilon/8$-optimal rule for player 1 if it goes first. Furthermore, if $r_2^n(\cdot)$ converges to $r_2(\cdot)$, then

$$\limsup_n W(x, u_1^\epsilon(u_2^{\mu,\rho,\delta}(\cdot|r_2^n)), r_2^n) = W(x, u_1^\epsilon(\tilde{u}_2^{\mu,\rho,\delta}(\cdot|r_2)), r_2) \le V^+(x) + 2\epsilon/8 \tag{6.12}$$

for some perturbation $\tilde{u}_2^{\mu,\rho,\delta}(\cdot|r_2)$.

The next step is to approximate the right side of (6.9) so that it depends only on samples of the $w(\cdot)$. Other than the $w(\cdot)$-variables, owing to the discretizations of time and control value, the conditioning data in (6.9) takes only a finite number of values. By the martingale convergence theorem, as $\lambda \to 0$, the function defined by

$$\begin{aligned} &q_{1,k}^\lambda(\gamma;\, w(l\lambda), l\lambda < k\Delta;\, u_1(l\Delta), l < k;\, u_2^{\mu,\rho,\delta}(l\rho|r_2), l\rho < k\Delta) \\ &\quad \equiv\ P\{u_1(k\Delta) = \gamma \big| w(l\lambda), l\lambda < k\Delta;\, u_1(l\Delta), l < k;\, u_2^{\mu,\rho,\delta}(l\rho|r_2), l\rho < k\Delta\} \end{aligned} \tag{6.13}$$

converges to

$$P\{u_1(k\Delta) = \gamma \big| w(s), s < k\Delta;\, u_1(l\Delta), l < k;\, u_2^{\mu,\rho,\delta}(l\rho|r_2), l\rho < k\Delta\}$$

for almost all $w(\cdot)$, for each value of the other conditioning variables. Thus, for small enough λ, the rule (6.9) can be approximated by (6.13). If the new rule is called $\hat{u}_1^{\epsilon,\lambda}(u_2^{\mu,\rho,\delta}(\cdot|r_2))$, then, for small λ,

$$W(x, \hat{u}_1^{\epsilon,\lambda}(u_2^{\mu,\rho,\delta}(\cdot|r_2)), r_2) \le V^+(x) + 6\epsilon/8,$$

and for small λ, μ, ρ, and δ and large Δ/δ and δ/ρ, the same inequality holds if the perturbation $\tilde{u}_2^{\mu,\rho,\delta}(\cdot|r_2)$ replaces $u_2^{\mu,\rho,\delta}(\cdot|r_2)$. We can suppose, without loss of generality, that Δ/λ is an integer.

There is one more approximation since we will require that the function $q_{1,k}(\cdot)$ in (6.1) be continuous in the $w(l\lambda)$-variables for each value of the others. Fix k, and let m denote the dimension of $w(1)$. Let $n = [k\Delta/\lambda] - 1$. Let $\bar{\alpha}$ denote the canonical value of the entire set $\{u_1(l\Delta), l < k\}$, and let $\bar{\beta}$ denote the canonical value of the entire set $\{u_2(l\rho), l\rho < k\Delta\}$. Let $w_\nu, v_\nu, \nu \le n$, be vectors in $\mathbb{R}^m$. For $\kappa > 0$, define the smoothed function

$$
\begin{aligned}
&q_{1,k}^{\lambda,\kappa}\left(\gamma\middle|\bar{\alpha},\bar{\beta};\, w_\nu, \nu \le n\right) \\
&= \frac{1}{[2\pi\kappa]^{nm/2}} \int \cdots \int e^{-|w_\nu - v_\nu|^2/[2\kappa]} q_{1,k}^{\lambda}\left(\gamma\middle|\bar{\alpha},\bar{\beta};\, v_\nu, \nu \le n\right) dv_1 \cdots dv_n.
\end{aligned}
\tag{6.14}
$$

The smoothed function defined by (6.14) is continuous in $\{w_\nu, \nu \le n\}$ for each value of $\bar{\alpha}, \bar{\beta}, \gamma$. As $\kappa \to 0$, for each $\bar{\alpha}, \bar{\beta}, \gamma$, it converges to $q_{1,k}^{\lambda}\left(\gamma\middle|\bar{\alpha},\bar{\beta};\, w_\nu, \nu \le n\right)$ for almost all (Lebesgue measure) $\{w_\nu, \nu \le n\}$. Since the measure of $\{w(l\lambda), l\lambda < k\Delta\}$ is absolutely continuous with respect to Lebesgue measure, the convergence is for almost all $w(\cdot)$. Finally, defining the function $q_{1,k}(\cdot)$ in (6.1) by $q_{1,k}^{\lambda,\kappa}(\cdot)$ for small enough λ and κ and calling the resulting control law $\bar{u}_1^\epsilon(u_2^{\mu,\rho,\delta}(\cdot|r_2))$, we have (6.1) and (6.3), and $q_{1,k}(\cdot)$ is continuous in the w-variables for each value of the others.

If players 1 and 2 are interchanged in all of the above arguments, then we get an ϵ-optimal (maximizing) rule analogous to the form (6.1) for player 2, and (6.4) holds. $\square$

7. Convergence of the numerical solutions. The next theorem establishes the convergence of the numerical procedure. It supposes the local consistency condition (3.1)–(3.3) everywhere, but recall the remarks concerning discontinuous dynamical and cost terms below (3.1). We do not show the convergence of the controls. In numerical examples, the sequence of optimal feedback controls for the chain does converge as well, and, in all examples of which we are aware, it is of a form that can be shown to be optimal. This would be the case if the optimal feedback controls $\bar{u}_i^h(\cdot)$ for the chain converged to feedback controls $\bar{u}_i(\cdot)$, where the convergence is uniform and the limits are continuous outside of an arbitrarily small neighborhood of a set D_d satisfying A2.7, and the process (2.4) under the $\bar{u}_i(\cdot)$ is unique in the weak sense. Then $W(x, \bar{u}_1, \bar{u}_2) = V(x)$.

Theorem 7.1. *Assume the local consistency conditions* (3.1)–(3.3)*,* A2.1–A2.3, A2.5, *and either* A2.6 *or* A2.7. *Then* $V^{\pm,h}(x) \to V(x)$ *as* $h \to 0$.

Proof. Let player 1 go first. Given $\epsilon > 0$, let us adapt the ϵ-optimal (minimizing) rule $\bar{u}_1^\epsilon(u_2^{\mu,\rho,\delta}(\cdot|r_2))$ for (2.4) that is defined by (6.1) for use on the chain. With player 1 using this rule for the Markov chain model, for each integer k player 1 uses a constant control value on the interpolated time interval $[k\Delta, k\Delta + \Delta)$. The continuous time interpolation of the relaxed control representation of the control processes which are used for the two players will be written as $r^h(\cdot) = (r_1^h(\cdot), r_2^h(\cdot))$. Thus, for some small positive μ, ρ, δ, and λ, the adaptation of the rule (6.1) for player 1 for the chain can be represented by the form, where $w^h(\cdot)$ is the "pre-Wiener" process in (3.11),

$$
\begin{aligned}
&P\left\{u_1^h(k\Delta) = \gamma\middle|u_1^h(l\Delta), l < k;, r_2^h(s), w^h(s), s \le t\right\} \\
&= P\{u_1^h(k\Delta) = \gamma\big|u_1^h(l\Delta), l < k;\, u_2^{\mu,\rho,\delta}(l\rho|r_2^h), l\rho < k\Delta;\, w^h(l\lambda), l\lambda < k\Delta\} \\
&= q_{1,k}(\gamma\big|u_1^h(l\Delta), l < k;\, u_2^{\mu,\rho,\delta}(l\rho|r_2^h), l\rho < k\Delta;\, w^h(l\lambda), l\lambda < k\Delta).
\end{aligned}
\tag{7.1}
$$

Given the rule (7.1) for player 1, player 2 selects a maximizing control at each state transition. Let $u_2^h(\cdot)$ denote player 2's optimal choice. Then $r_2^h(\cdot)$ is its relaxed control representation, and $r_1^h(\cdot)$ is the relaxed control representation of the realization of player 1's actions which are determined by (7.1).

Choose a weakly convergent subsequence of $\{\psi^h(\cdot), r^h(\cdot), w^h(\cdot), y^h(\cdot)\}$ (abusing terminology, for simplicity this subsequence is also indexed by h). This converges weakly to a solution $(x(\cdot), r(\cdot), w(\cdot), y(\cdot))$ of (2.4) and $W^h(x, r^h) \to W(x, r)$. The proofs of these facts are the same as for the pure control problem in [32, Theorems 11.1.2 and 11.1.5]. Let us use the Skorohod representation [14, Theorem 1.8, Chapter 3] so that all processes are defined on the same probability space, and weak convergence becomes convergence with probability one. If

$$r_2(l\delta+\delta, \partial C_2^m) - r_2(l\delta, \partial C_2^m) = 0$$

for all m, l, then, using the continuity of $q_{1k}(\cdot)$ in the w-variables,

$$\begin{aligned} &q_{1,k}(\gamma\big|u_1^h(l\Delta), l<k;\ u_2^{\mu,\rho,\delta}(l\rho|r_2^h), l\rho<k\Delta;\ w^h(l\lambda), l\lambda<k\Delta) \\ &\quad \to q_{1,k}(\gamma\big|u_1(l\Delta), l<k;\ u_2^{\mu,\rho,\delta}(l\rho|r_2), l\rho<k\Delta;\ w(l\lambda), l\lambda<k\Delta) \end{aligned} \tag{7.2}$$

with probability one,

$$W^h(x, \bar u_1^\epsilon(u_2^{\mu,\rho,\delta}(\cdot|r_2^h)), r_2^h) \to W(x, \bar u_1^\epsilon(u_2^{\mu,\rho,\delta}(\cdot|r_2)), r_2), \tag{7.3}$$

and (6.2) holds. In any case, as noted in the paragraph below (6.9),

$$W^h(x, \bar u_1^\epsilon(u_2^{\mu,\rho,\delta}(\cdot|r_2^h)), r_2^h) \to W(x, \bar u_1^\epsilon(\tilde u^{\mu,\rho,\delta}(\cdot|r_2)), r_2), \tag{7.4}$$

where $\tilde u^{\mu,\rho,\delta}(\cdot|r_2)$ is a perturbation of the type defined in Theorem 6.1. We have, where $\tilde r_i(\cdot)$ is the relaxed control representations of the canonical $\tilde u_i(\cdot)$,

$$\begin{aligned} V^{+,h}(x) &= \inf_{\tilde u_1 \in \mathcal{U}_1^h(1)} \sup_{\tilde u_2 \in \mathcal{U}_2^h(2)} W^h(x, \tilde u_1, \tilde u_2) \le \sup_{\tilde u_2 \in \mathcal{U}^h(2)} W^h(x, \bar u_1^\epsilon(u_2^{\mu,\rho,\delta}(\cdot|\tilde r_2)), \tilde r_2) \\ &= W^h(x, \bar u_1^\epsilon(u_2^{\mu,\rho,\delta}(\cdot|r_2^h)), r_2^h) \to W(x, \bar u_1^\epsilon(\tilde u_2^{\mu,\rho,\delta}(\cdot|r_2)), r_2). \end{aligned} \tag{7.5}$$

This and the inequality (6.3) imply that, for any $\epsilon > 0$,

$$\limsup_h V^{+,h}(x) \le V(x) + \epsilon. \tag{7.6}$$

Now repeat the procedure, but with player 2 going first. Use the analogue of the ϵ-optimal rule (6.1) for player 2. Then, given that rule for player 2, let player 1 optimize (minimize). Writing $r^h(\cdot)$ for the actual control process, we have the analogue of (7.5), namely,

$$\begin{aligned} V^{-,h}(x) &= \sup_{\tilde u_2 \in \mathcal{U}_2^h(1)} \inf_{\tilde u_1 \in \mathcal{U}_1^h(2)} W^h(x, \tilde u_1, \tilde u_2) \ge \inf_{\tilde u_1 \in \mathcal{U}_1^h(2)} W^h(x, \tilde r_1, \bar u_2^\epsilon(u_1^{\mu,\rho,\delta}(\cdot|\tilde r_1))) \\ &= W^h(x, r_1^h, \bar u_2^\epsilon(u_1^{\mu,\rho,\delta}(\cdot|r_1^h))) \to W(x, r_1, r_2) = W(x, r_1, \bar u_2^\epsilon(\tilde u_1^{\mu,\rho,\delta}(\cdot|r_1))), \end{aligned}$$

where $\tilde u_1^{\mu,\rho,\delta}(\cdot|r_1)$ is a perturbation of $u_1^{\mu,\rho,\delta}(\cdot|r_1)$. Using this and (6.4) yields

$$\liminf_h V^{-,h}(x) \ge V(x) - \epsilon. \tag{7.7}$$

Finally, (7.5) and (7.7) yield

$$\limsup_h V^{+,h}(x) - \liminf_h V^{-,h}(x) \le 2\epsilon,$$

and the proof is concluded since $\epsilon > 0$ is arbitrary. □

8. Comments and extensions.

8.1. Examples with separated dynamics. Only a few comments will be made since the interest is in reminding the reader of the connection between risk-sensitive, robust, constrained, and large deviations control and differential games.

Risk-sensitive control. Let $\epsilon > 0$, and consider the problem of minimizing

$$\Lambda^\epsilon(u_1) = \lim_{T\to\infty} \frac{1}{T} \log E_x \exp\left[\frac{1}{\epsilon}\int_0^T L(x(s), u_1(s))ds\right]$$

for bounded and continuous $L(\cdot)$ with dynamics given by

$$dx = b(x, u_1)dt + \left[\frac{\epsilon}{2\gamma}\right]^{1/2} \sigma(x)dw + dz,$$

where $u_1(t) \in U_1$, a compact set. This is part of the subject of risk-sensitive control [18]. Under appropriate conditions, the solution reduces to that of a differential game with separated dynamics. Let $\bar{\Lambda}^\epsilon = \inf_{u_1} \Lambda^\epsilon(u_1)$, and define $a(x) = \sigma(x)\sigma'(x)$, assumed positive definite for each x. For x in the interior of G, the Isaacs equation is [18]

$$\begin{aligned}\bar{\Lambda}^\epsilon = &\frac{\epsilon}{4\gamma^2}\sum_{i,j} V_{x_i x_j}(x) a_{ij}(x)\\ &+ \max_{u_2}[V_x'(x)u_2 - \gamma^2 |u_2|^2] + \min_{u_1}\left[V_x'(x)b(x,u_1) + L(x,u_1)\right].\end{aligned}$$

This corresponds to a two-person game with cost rate $k(x,u) = L(x,u_1) - \gamma^2|u_2|^2$. Only u_1 appears in the dynamical equation.

In applications, the set U_1 is often unbounded. Effective approaches to dealing with unbounded sets for the control problem are in the chapters concerning the variational problems in [32], and they can be adapted to the game problem under appropriate conditions.

Constrained optimization via the Lagrangian method. Consider the model (2.4), but with only one control $u_1(\cdot)$. Let $q_i(\cdot), i \le p$, be bounded, continuous, and continuously differentiable real-valued functions on G, and consider the minimization of

$$E\int_0^\infty e^{-\beta t} k_1(x(t), u_1(t))dt$$

for a bounded and continuous function subject to the constraints $Eq_i(x(t)) \le 0$ for almost all t, $i = 1, \ldots, \mu$. The problem can be formulated as a game, via the introduction of Lagrange multipliers $u_{2,i}(t) \ge 0, i = 1, \ldots, p$. Define

$$W(x, u_1, u_2) = E\int_0^\infty e^{-\beta t}\left[k_1(x(t), u_1(t)) + u_2'(t)q(x(t))\right]dt.$$

Then the solution is obtained from the game with upper (and lower as well, since the game has a value) value

$$\lim_{\Delta\to 0}\ \inf_{u\in\mathcal{U}_1(\Delta)}\ \sup_{u_2\in\mathcal{U}_2} W(x, u_1(u_2), u_2).$$

Here the set U_2 is $[0, \infty)$. However, for numerical purposes, one bounds the interval and then experiments with the bound until the desired solution is obtained.

Controlled large deviations problems. Consider the problem in controlled large deviations where one wishes to minimize (over choice of a control) the large deviations estimate of the probability of an event, say the probability that a set will be exited over some time interval. The mathematical formulation of such problems for diffusion-type models often reduces to that of a game, where the dynamics and cost function are separated, analogously to the forms of $b(\cdot)$ and $k(\cdot)$ in (2.4) and (2.5). See for example, the development in [12].

8.2. Stopping time problems and pursuit-evasion games.

Stopping cost not depending on who stops first. Suppose that player i now has a choice of an $\mathcal{F}_t$-stopping time τ_i as well as of the controls. Define $\tau = \min\{\tau_1, \tau_2\}$. For a continuous function $g(\cdot)$, replace (2.5) by

$$\begin{aligned} W(x, r, \tau) \;=\; & E \int_0^\tau e^{-\beta t} \left[\int_{U_i} \sum_{i=1}^{2} k_i(x(t), \alpha_i) r_{i,t}(d\alpha_i) dt + c' dy(t) \right] \\ & + E e^{-\beta\tau} g(x(\tau)). \end{aligned} \tag{8.1}$$

Thus, in this model, the stopping cost $g(x(\tau))$ does not depend on who selects the stopping time.

The control spaces such as $\mathcal{U}_i$, $\mathcal{U}_i(\Delta), \mathcal{L}_i(\Delta)$, and $\mathcal{U}_i(\mu, \delta, \Delta)$, etc. need to be extended so that they include the stopping times. Let $\overline{\mathcal{U}}_i$ be the set of pairs $(u_i(\cdot), \tau)$, where $u_i(\cdot) \in \mathcal{U}_i$ and τ is an $\mathcal{F}_t$-stopping time. Let $\overline{\mathcal{U}}_i(\Delta)$ denote the subset where $u_i(\cdot) \in \mathcal{U}_i(\Delta)$ and τ takes values $k\Delta, k = 0, 1 \ldots$, where the set $\{\omega : \tau = k\Delta\}$ is $\mathcal{F}_{k\Delta}$-measurable. Similarly, $\overline{\mathcal{U}}_i(\mu, \delta, \Delta)$ denotes the subset of $\overline{\mathcal{U}}_i(\Delta)$, where $u_i(\cdot) \in \mathcal{U}_i(\mu, \delta, \Delta)$. Let $\overline{\mathcal{L}}_1(\Delta)$ denote the set of controls in $\overline{\mathcal{U}}_1(\Delta)$ for player 1 which can be represented in the form

$$\begin{aligned} & P\left\{\tau_1 > k\Delta \middle| w(s), u_2(s), s < t; u_1(l\Delta), l < k, \tau_1 \geq k\Delta\right\}, \\ & P\left\{u_1(k\Delta) \in \cdot \middle| w(s), u_2(s), s < t; u_1(l\Delta), l < k;\ \tau_1 > k\Delta\right\}. \end{aligned} \tag{8.2}$$

Define $\mathcal{L}_2(\Delta)$ analogously for player 2.

The definitions of the upper and lower values in (2.6) are replaced by, respectively,

$$\begin{aligned} V^+(x) \;&=\; \lim_{\Delta \to 0} \inf_{u_1, \tau_1 \in \overline{\mathcal{L}}_1(\Delta)} \sup_{(u_2, \tau_2) \in \overline{\mathcal{U}}_2} W(x, u_1, u_2, \tau), \\ V^-(x) \;&=\; \lim_{\Delta \to 0} \sup_{(u_2, \tau_2) \in \overline{\mathcal{L}}_2(\Delta)} \inf_{(u_1, \tau_1) \in \overline{\mathcal{U}}_1} W(x, u_1, u_2, \tau). \end{aligned} \tag{8.3}$$

The first line of (8.3) is to be understood as follows. Suppose that the game has not stopped by time $k\Delta$. Then, at $k\Delta$, player 1 goes first and decides whether to stop based on data to time $k\Delta-$. If it stops, the game is over. If not, it selects the control value $u_1(k\Delta)$ (which it will use until $(k\Delta + \Delta)-$ or until player 2 stops, whichever comes first) based on data to time $k\Delta-$. If the game is not stopped at $k\Delta$ by player 1, then player 2 has the opportunity to stop at any time on $[k\Delta, k\Delta + \Delta)$, with the decision to stop at any time being based on all data to that time. Until it stops (if it does), it chooses admissible control values $u_2(\cdot)$. The procedure is then repeated at time $k\Delta + \Delta$, and so forth. With these changes and minor (mostly notational) modifications, the previous theorems continue to hold. In particular, Theorem 7.1 holds.

Stopping cost depends on who stops first. Now let the cost be

$$W(x,r,\tau_1,\tau_2) = E\int_0^\tau e^{-\beta t}\left[\int_{U_i}\sum_{i=1}^{2} k_i(x(t),\alpha_i)r_{i,t}(d\alpha_i)dt + c'dy(t)\right] + Ee^{-\beta\tau_1}g_1(x(\tau_1))I_{\{\tau_1<\tau_2\}} + Ee^{-\beta\tau_2}g_2(x(\tau_2))I_{\{\tau_2\le\tau_1\}}, \tag{8.4}$$

where $\tau = \min\{\tau_1,\tau_2\}$ and the $g_i(\cdot)$ are bounded and continuous. The proof in Theorem 5.1 that the game has a value does not carry over to the present case, since the stopping cost depends on who stops first. However, if the game has a value, then Theorem 7.1 holds.

Consider the approximating Markov chain. Let player 1 go first, and let I_1 denote the indicator of the event that player 1 stops at the current step. Then the Bellman equation for the (for example) upper value is

$$\begin{aligned} &V^{+,h}(x) = \min_{I_1,\alpha_1}\{g_1(x)I_1, \\ &(1-I_1)\max[\max_{\alpha_2}(E_x^\alpha e^{-\beta\Delta t^h(x,\alpha)}V^{+,h}(\xi_1^h) + k(x,\alpha)\Delta t^h(x,\alpha)), g_2(x)]\}. \end{aligned} \tag{8.5}$$

Acknowledgments. The author greatly appreciates the many helpful suggestions of Paul Dupuis and Wendell Fleming.

REFERENCES

[1] E. Altman and H. J. Kushner, *Admission control for combined guaranteed performance and best effort communications systems under heavy traffic*, SIAM J. Control Optim., 37 (1999), pp. 1780–1807.

[2] J. A. Ball, M. Day, and P. Kachroo, *Robust feedback control of a single server queueing system*, Math. Control Signals Systems, 12 (1999), pp. 307–345.

[3] J. A. Ball, M. Day, P. Kachroo, and T. Yu, *Robust L_2-gain for nonlinear systems with projection dynamics and input constraints: An example from traffic control*, Automatica J. IFAC, 35 (1999), pp. 429–444.

[4] M. Bardi, S. Bottacin, and M. Falcone, *Convergence of discrete schemes for discontinuous value functions of pursuit-evasion games*, in New Trends in Dynamic Games and Applications, G.I. Oldser, ed., Birkhäuser Boston, Boston, 1995, pp. 273–304.

[5] M. Bardi, M. Falcone, and P. Soravia, *Fully discrete schemes for the value function of pursuit-evasion games*, in Advances in Dynamic Games and Applications, T. Basar and A. Haurie, eds., Birkhäuser Boston, Boston, 1994, pp. 89–105.

[6] M. Bardi, M. Falcone, and P. Soravia, *Numerical methods for pursuit-evasion games via viscosity solutions*, in Stochastic and Differential Games: Theory and Numerical Methods, M. Bardi, T. E. S. Raghavan, and T. Parthasarathy, eds., Birkhäuser Boston, Boston, 1999, pp. 105–175.

[7] M. Bardi and P. Soravia, *Approximation of differential games of pursuit-evasion by discrete time games*, in Differential Games: Developments in Modelling and Computation, R. P. Hämäläinen and H. K. Ehtamo, eds., Springer-Verlag, Berlin, New York, 1991, pp. 131–143.

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

[9] V. E. Benes, *Existence of optimal stochastic control laws*, SIAM J. Control, 9 (1971), pp. 446–472.

[10] P. Billingsley, *Convergence of Probability Measures*, 2nd ed., Wiley, New York, 1999.

[11] P. Dupuis and H. Ishii, *On Lipschitz continuity of the solution mapping to the Skorokhod problem, with applications*, Stochastics Stochastic Rep., 35 (1991), pp. 31–62.

[12] P. Dupuis and W. M. McEneaney, *Risk-sensitive and robust escape criteria*, SIAM J. Control Optim., 35 (1997), pp. 2021–2049.

[13] R. J. Elliott and N. J. Kalton, *Existence of Value in Differential Games*, Mem. Amer. Math. Soc., 126, AMS, Providence, RI, 1972.

[14] S. N. Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence*, Wiley, New York, 1986.

[15] W. F. Fleming, *Generalized solutions in optimal stochastic control*, in Differential Games and Control Theory III, E. Roxin, P. T. Liu, and R. Sternberg, eds., Marcel Dekker, New York, 1977, pp. 147–165.
[16] W. H. Fleming, *The convergence problem for differential games*, J. Math. Anal. Appl., 3 (1961), pp. 102–116.
[17] W. H. Fleming, *The convergence problem for differential games* II, in Advances in Game Theory, Ann. of Math. Stud. 52, Princeton University Press, Princeton, NJ, 1964, pp. 195–210.
[18] W. H. Fleming and W. M. McEneaney, *Risk-sensitive control on an infinite time horizon*, SIAM J. Control Optim., 33 (1995), pp. 1881–1915.
[19] W. H. Fleming and P. E. Souganidis, *On the existence of value functions for two-player zero-sum differential games*, Indiana Univ. Math. J., 38 (1989), pp. 293–314.
[20] A. Friedman, *Differential Games*, Wiley, New York, 1971.
[21] J. M. Harrison and M. I. Reiman, *Reflected Brownian motion on an orthant*, Ann. Probab., 9 (1981), pp. 302–308.
[22] J. M. Harrison and R. J. Williams, *Brownian models of open queueing networks with homogeneous customer populations*, Stochastics Stochastics Rep., 22 (1987), pp. 77–115.
[23] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer-Verlag, New York, 1988.
[24] T. G. Kurtz, *Approximation of Population Processes*, CBMS-NSF Regional Conf. Ser. in Appl. Math. 36, SIAM, Philadelphia, 1981.
[25] H. J. Kushner, *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations*, Academic Press, New York, 1977.
[26] H. J. Kushner, *Numerical methods for stochastic control problems in continuous time*, SIAM J. Control Optim., 28 (1990), pp. 999–1048.
[27] H. J. Kushner, *Jump-diffusions with controlled jumps: Existence and numerical methods*, J. Math. Anal. Appl., 249 (2000), pp. 179–198.
[28] H. J. Kushner, *Heavy Traffic Analysis of Controlled Queueing and Communication Networks*, Springer-Verlag, Berlin, New York, 2001.
[29] H. J. Kushner, *Numerical Approximations for Stochastic Differential Games: The Ergodic Case*, Report, Applied Math., Brown University, Providence, RI, 2001.
[30] H. J. Kushner and S. G. Chamberlain, *Finite state stochastic games: Existence theorems and computational procedures*, IEEE Trans. Automat. Control, 14 (1969), pp. 248–255.
[31] 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., 26 (1969), pp. 560–575.
[32] H. J. Kushner and P. Dupuis, *Numerical Methods for Stochastic Control Problems in Continuous Time*, 2nd ed., Springer-Verlag, Berlin, New York, 2001.
[33] J. Neveu, *Mathematical Foundations of the Calculus of Probability*, Holden-Day, San Francisco, 1965.
[34] O. Pourtallier and M. Tidball, *Approximation of the Value Function for a Class of Differential Games with Target*, Research report 2942, INRIA, Le Chesnay, France, 1996.
[35] O. Pourtallier and B. Tolwinski, *Discretization of Isaac's Equation*, Report, INRIA, Le Chesney, France, 1992.
[36] P. Soravia, *H_∞ control of nonlinear systems: Differential games and viscosity solutions*, SIAM J. Control Optim., 34 (1996), pp. 1071–1097.
[37] T. E. S. Raghavan and J. A. Filar, *Algorithms for stochastic games: A survey*, Z. Oper. Res., 35 (1991), pp. 437–472.
[38] M. I. Reiman and R. J. Williams, *A boundary property of semimartingale reflecting Brownian motions*, Probab. Theory Related Fields, 77 (1988), pp. 87–97.
[39] M. R. Reiman, *Open queueing networks in heavy traffic*, Math. Oper. Res., 9 (1984), pp. 441–458.
[40] M. Tidball, *Undiscounted zero-sum differential games with stopping times*, in New Trends in Dynamic Games and Applications, G. J. Oldser, ed., Birkhäuser Boston, Boston, 1995, pp. 305–322.
[41] M. Tidball and R. L. V. González, *Zero-sum differential games with stopping times: Some results and about its numerical resolution*, in Advances in Dynamic Games and Applications, T. Basar and A. Haurie, eds., Birkhäuser Boston, Boston, 1994, pp. 106–124.