2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Admissible Controls, Modeling, and Optimization for a New Class of Nonlinear Stochastic Delay Systems

Harold J. Kushner
Applied Math, Brown University, Providence RI 02912 USA, hjk@dam.brown.edu

Abstract. The paper deals with several fundamental issues that have not been previously addressed in the modeling and optimization of nonlinear stochastic delay systems. For an example, consider the special case of a system with a delayed control term of the form $f(u_1(t+\theta_1), u_2(t+\theta_2))$, where the delays $\theta_i < 0$ are different and $f(\cdot)$ is not the sum of functions of each of the controls separately. The class of adapted relaxed controls is no longer adequate as the class of admissible controls, at least in the sense that the infimum of the costs over this class is not the infimum over the class of ordinary controls, and the limit of convergent sequences might be meaningless. We deal with such issues of admissibility and optimization for a large class of systems that includes the above example. The appropriate extensions and the proofs are not obvious. The issues are crucial for the convergence of numerical approximations to optimal control problems, as well as for the optimization problem to be well-defined.

I. Introduction and Motivation.

We are concerned with a fundamental issue that has not been previously addressed in the modeling, optimization, and numerical optimization of nonlinear stochastic delay systems and which arises with many new classes of models. The proofs of the existence of optimal controls generally use weak convergence methods. One chooses a minimizing sequence of admissible controls, shows that the set of (paths, controls) is tight, and that the limit of any weakly convergent subsequence is an optimal process. This has required the use of relaxed controls since a minimizing sequence of ordinary controls will not have a meaningful convergent subsequence in general, and (least for a compact control-value space, and the use of the weak topology), any sequence of relaxed controls is compact; i.e., closed under weak convergence. With this method, the limit control would be in relaxed form. Under quite weak conditions any relaxed control can be approximated by a piecewise-constant ordinary control in the sense that the (approximating controls, associated paths) converge weakly to the (original relaxed control and associated path), and that the associated costs converge. This approximation property is essential for the method to work, since it assures that the use of relaxed controls does not enlarge the range of values of the cost function. In particular, the infimum of the cost function in this class equals the infimum over the class of ordinary controls (e.g, [10, Theorem 10.1.2]).

Similar issues arise in the proofs of convergence of numerical approximations to optimal control problems, say by the Markov chain approximation method [10]. In the convergence proofs, the approximating chain is interpolated into a continuous-time process, and the continuous-time interpolation of the controls is represented in relaxed control form. It is then shown that the set of approximations (parameterized by the approximation parameter) is tight and that the limit of any weakly convergent subsequence is optimal. The requires the existence of an optimal control for the original process, as well as arbitrarily good piecewise-constant and finite-valued approximations to it.

For the classical models with no delays, the above results can be shown under very broad conditions. The same is true for the traditional types of nonlinear stochastic delay systems, such as (2.1) below and those in [9], which develops numerical methods for getting optimal controls and value functions or to evaluate performance for a large class of such systems, with possibly reflecting boundaries, by extensions of the Markov chain approximation methods. The methods are the same as those for the non-delay models, using relaxed controls. Then the set of adapted relaxed controls can be taken to be the set of admissible controls.

There are many important classes of models with delays for which the adaptation of these methods has not been done and is not obvious. For motivation consider the problem $dx(t) = f_1(x_t, u_1(t+\theta_1), u_2(t+\theta_2))dt + f_2(x_t, u_0(t+\theta_3), u_0(t+\theta_4))dt + dw(t)$, where x_t is the path memory segment at t, $w(\cdot)$ is a Wiener process, the functions $f_i(\cdot)$ are NOT sums of functions of their individual arguments,[1] and the delays $\theta_i < 0$ take different values. The set of relaxed controls representations is not adequate since it does not capture the JOINT values of the controls at the delay times, and the question of piecewise-constant approximations is open. The difficulty with getting the approximations is that all of the values $\{u_i(t)\}$ are computed at time t, but applied to the dynamics only after the individual delays. So it is the representations (what ever that will be) of the set of controls $\{u_i(t)\}$ that must be approximate by admissible controls, such that, when delayed, we get a good approximation to the joint values that are needed for the dynamics as in the above equation.

One of our motivations came from a problem of admission

Supported by NSF DMS-0804822, ARO contract FA9550-09-1-0378, AFOSR contract W911NF-09-1-0155

[1] E.g., $f_i(\cdot)$ is the product of the controls at the different delays.

control in a network. Jobs arrive at a portal and send a message to a downstream buffer/router requesting admission. The acceptance arrives at the portal after the round trip transportation delay, and if accepted the source can start sending data. If rejected, the source has the possibility of reapplying after a (perhaps random) delay. Then a new control at the router will determine the acceptance. This leads to a model where controls appear in a product form with different delays, and which modulates the arrival process. This paper is devoted to such matters for a new class of systems.

II. The Model

A currently common form. The paths of all of the considered process will be confined to a convex bounded polyhedron G by boundary reflections. To put our model into perspective, first recall the model that was used for the development of numerical algorithms for the approximation of optimal controls and values in [9, Chapter 9], and covers most forms used in the study of delay systems:

$$\begin{aligned} dx(t) &= c(x(t),u(t))dt + \sigma(x(t))dw(t) + dz(t) \\ &+dt\int_{-\bar\theta}^{0} b(x(t+\theta),u(t+\theta),\theta)d\mu_a(\theta) \\ &+dt\int_{\theta=-\bar\theta}^{0} p(x(t+\theta))d_\theta y(t+\theta). \end{aligned} \tag{2.1}$$

$y_i(\cdot)$ is the component of $z(\cdot)$ due to reflection from the ith face of G and $u(t)\in U$, a compact set. We use reflected diffusions with delays since they model many problems of interest and there is little data concening them. If there is no reflecting boundary, then simply drop all of the $z(\cdot)$ and $y(\cdot)$ processes. Reflecting boundaries with delays of the reflection terms occur frequently in the modeling of communications systems. For example some of the state space constraints might be due to finite buffers. Loosely speaking, in simple models of the internet TCP control, information (non acknowledgements) concerning buffer overflows is sent to the source after a transportation delay, and affects the source rate. Numerical data for such problems is in [4], [5], and attests to the possibilities of the numerical methods in [9]. In such problems it is selected components of the reflection process that are fed back after a delay.

The model (2.1) was extended in [6], which was concerned with numerical methods, to allow more general forms of the delayed reflection term component and there was a controlled and delayed Poisson measure driving term as well. For both of these forms the class of admissible controls is just the class of adapted relaxed controls, and the proof of this fact follows the usual lines described in the Introduction. The situation is much more complicated for the models in the sequel.

In models such as (2.1) the vector-valued control $u(\cdot)$ appears in the dynamics in a form that is like a function of the control at one delay plus a function at another delay, etc., what we call an "additive" form. It was easy to show that the set of admissible controls is just the set of relaxed controls and to get arbitrarily good piecewise-constant ordinary control approximations to any relaxed control, in that the corresponding paths and costs are arbitrarily close. It was this property that allowed the use of the relaxed controls as admissible controls, since its use did not enlarge the range of the cost function values and the set of paths and relaxed controls is closed under weak convergence.

As noted in the Introduction, we need to define the class of admissible controls not only to get a well-defined model, but also to get existence (of an optimal control) proofs and limit and approximation theorems. The existence of such piecewise-constant approximations ensures that we have not enlarged the range of values of the cost function, and is necessary for the convergence proofs for numerical approximations to optimal controls and values.

Our model. The model considered in this paper is the following. Let $0 \ge \theta_k \cdots > \theta_1 = -\bar\theta$, where $\bar\theta$ is the maximum delay, and consider (for $t\ge 0$, x_t denotes the part of $x(\cdot)$ on $[t-\bar\theta,t]$)

$$\begin{aligned} dx(t) &= dt\int_{-\bar\theta}^{0} b_a(x_t,u_a(t+\theta),\theta)\mu_a(d\theta) + \sigma(x_t)dw(t) + \\ & b(x_t,u_1(t+\theta_1),u_2(t+\theta_2),\ldots,u_k(t+\theta_k),t)dt + dz(t), \end{aligned} \tag{2.2}$$

where $u_i\in U_i$, $u_a(\cdot)\in U_a$, all compact, and $U=\prod_{i=1}^k U_i$. The values $u_a(t), u_i(t), i=1,\ldots,$ are determined at time t, but they appear in the dynamics with the appropriate delays. The memory segments are $u_i(t+\theta), \theta\in[\theta_i,0], i\le k$, and $(u_a(t+\theta),x(t+\theta)), \theta\in[-\bar\theta,0]$. A delayed reflection term could be included, but since it does not play a major role in the concerns of this paper, for expositional simplicity it will not be included.

We will concentrate on the discounted cost function, for $\beta>0$ and bounded and continuous $k(\cdot)$:

$$W(\bar x,r) = E_{\bar x}^r \int_0^\infty \int_U e^{-\beta t}\left[k(x(t),\alpha)r(d\alpha\, dt) + q'dy(t)\right], \tag{2.3}$$

where the expectation is for the process under initial path data $\bar x$ and the use of relaxed control $r(\cdot)$.

III. Assumptions

Weak convergence. Let S be a complete and separable metric space with metric $\rho(\cdot)$. Define $D(S;[-\bar\theta,\infty))$ to be the space of S-valued functions on $[-\bar\theta,\infty)$ that are right-continuous with left-hand limits. The symbol $\Rightarrow$ denotes weak convergence. The following criterion for tightness will be used.

Theorem 3.1. [7, Theorem 2.7b.] *Let $X^n(\cdot)$ be processes with paths in $D(S;[-\bar\theta,\infty))$. For each $\delta>0$ and rational $t<\infty$, let there be a compact set $S_{\delta,t}\subset S$ such that $\sup_n P(X^n(t)\notin S_{\delta,t})\le\delta$. Let $\mathcal{F}_t^n$ be the σ-algebra determined by $\{X^n(s),\ s\le t\}$ and let $\mathcal{T}_n(T)$ be the set of $\mathcal{F}_t^n$-stopping times that are no bigger than T. Suppose that*

$$\lim_{\delta_1\to 0}\limsup_n \sup_{\tau\in\mathcal{T}_n(T)} E\min\{1,\rho(X^n(\tau+\delta_1),X^n(\tau))\} = 0$$

for each $T < \infty$. Then $\{X^n(\cdot),\ n < \infty\}$ is tight in $D(S; [0,\infty))$.

The weak topology is used on the space of relaxed controls. Thus $r^n(\cdot) \Rightarrow r(\cdot)$ if and only if $\int\int \phi(\alpha,s) r^n(d\alpha\, ds) \to \int\int \phi(\alpha,s) r(d\alpha\, ds)$ for all continuous real-valued functions $\phi(\cdot)$ with compact support. The space of relaxed controls with this topology is compact. Otherwise S is a vector-valued Euclidean space with the Skorokhod topology [3]. For A a Borel set in U, and $a \le b$, we will write $r(A,[a,b]) = r(A \times [a,b])$, and $r(A,b)$ if $a = 0$.

Assumptions (A3.1) and (A3.2) below on the constraint set G are slightly weaker than those in [10, Section 5.7]. They are quite standard in the treatment of diffusions with constraining boundaries, and include the so-called "completely S" conditions that are used in modeling stochastic networks. References [1], [2] and [8, Section 3.5] discuss their role in applications in queueing and communications systems. These assumptions are made only because they ensure the bounds (3.1) and (3.2), which in turn ensure that the reflection component of the cost functions are well-defined, and that the various sequences of reflection processes are tight.

A3.0. *$b(\cdot)$ and $\sigma(\cdot)$ are bounded and continuous. The continuity in x_t is in the sup norm topology.*

After the admissible controls are defined, we will impose a weak-sense uniqueness condition.

A3.1. *G is compact and is the intersection of a finite number of closed half spaces in Euclidean r-space $I\!R^r$ and is the closure of its interior. Let ∂G_i, $i = 1, \ldots,$ denote the faces of G, and n_i the interior normal to ∂G_i. In the interior of ∂G_i, the reflection direction is the unit vector d_i, and $\langle d_i, n_i\rangle > 0$ for each i. The possible reflection directions at points on the intersections of the ∂G_i are in the convex hull of the directions on the adjoining faces. No more than r constraints are active at any boundary point.*

A3.2. *For an arbitrary corner or edge, let d_i denote the directions of reflection on the adjoining faces only. Then there are constants $a_i > 0$ such that for all such i,*

$$a_i \langle n_i, d_i\rangle > \sum_{j\ne i} a_j \left|\langle n_i, d_j\rangle\right|.$$

With this notation, $y_i(\cdot)$ can increase only when $x(t) \in \partial G_i$ and $z(t) = \sum_i d_i y_i(t)$.

Bounds on the reflection term. The following growth estimate is needed to ensure that the cost function is well defined, and for the proofs of tightness in convergence. Let $x_i(t) = \psi_i(t) + z_i(t)$, define $x = x_1 - x_2, z = z_1 - z_2$, etc., and let $|z|(a,b)$ denote the variation of $z(\cdot)$ on $[a,b]$. Then (A3.1) and (A3.2) imply that there is $C < \infty$ such that [1][2]

$$|z|(0,t)| + |y(0,t)| \le C \sup_{s\le t} |\psi(s)|. \tag{3.1}$$

[2]See [10, Theorem 11.1.1] and [8, Theorem 3.5.1]. Our (A2.2) is the (A3.5.4) of the last reference and as noted in [8, Section 3.5], it implies the condition (A3.5.3) there, which in turn (see [1]) ensures (3.1).

For (2.3) or (3.1) and τ a bounded stopping time,

$$\lim_{\delta\to 0} \sup_{\bar{x},r,\tau} E^r_{\bar{x}} |z|^2(\tau, \tau+\delta) = 0, \tag{3.2}$$

which implies that, for $T > 0$, $\sup_{\bar{x},r,t} E^r_{\bar{x}} |z|^2(t, t+T) < \infty$.

IV. Admissible Controls

Let $\mathcal{R}$ denote the class of adapted relaxed controls $r(\cdot)$, the relaxed control extensions of the set of ordinary controls $u(\cdot) = (u_a(\cdot), u_1(\cdot), u_2(\cdot), \ldots, u_k(\cdot))$ and defined on $[-\bar{\theta}, \infty)$. Let $\overline{\mathcal{R}}$ denote the class of relaxed controls $\bar{r}(\cdot)$, defined on $[0,\infty)$, where $\bar{r}(\cdot)$ is the class of relaxed control extensions of the set of delayed ordinary controls $(\bar{u}_1(\cdot), \ldots, \bar{u}_k(\cdot)) = (u_1(\cdot+\theta_1), u_2(\cdot+\theta_2), \ldots, u_k(\cdot+\theta_k))$. By the ith marginal $r_i(\cdot)$ of $r(\cdot)$ we mean the specialization to the ith component of the control, and analogously define the ith marginal $\bar{r}_i(\cdot)$. Let $r'_i(d\alpha_i, t)$ and $\bar{r}'_i(d\alpha_i, t)$ denote the left-hand derivatives of the ith marginals $r_i(\cdot)$ and $\bar{r}_i(\cdot)$, resp. Then, by the definitions, the marginals of $\bar{r}(\cdot)$ are delayed in that $\bar{r}_i(d\alpha_i\, dt) = \bar{r}'_i(d\alpha_i, t)dt = r'_i(d\alpha_i, t+\theta_i)dt$.

The control $u_a(\cdot)$ plays a minor role since it appears "additively" and the piecewise-constant approximations are the usual ones. For simplicity of notation and development, the $b_a(\cdot)$ term will be dropped, and we work with $k = 2$. The development and results are the same for the general case. The definitions mean (meas=Lebesgue measure)

$$r(A_1 \times A_2, [0,t]) = \text{meas}\left\{s : u_i(s) \in A_i, i \le 2, \text{ on } [0,t]\right\},$$

$$\bar{r}(A_1 \times A_2, [0,t]) = \text{meas}\left\{s : u_i(s+\theta_i) \in A_i, i \le 2, \text{ on } [0,t]\right\}.$$

On dropping $b_a(\cdot)$ and using $k = 2$, (2.2) can be written as

$$\begin{aligned} x(t) = x(0) &+ \int_0^t \sigma(x_s) dw(s) + z(t) \\ &+ \int_0^t \int_U b(x_s, \alpha_1, \alpha_2) \bar{r}(d\alpha_1\, d\alpha_2\, ds). \end{aligned} \tag{3.1}$$

The following assumption will be used.

A4.1. *There is a unique weak-sense solution to* (2.2) *and* (3.1) *for each* $(r(\cdot), \bar{r}(\cdot), w(\cdot))$.

To show that $(\mathcal{R}, \overline{\mathcal{R}})$, the set of pairs $(r(\cdot), \bar{r}(\cdot))$, can be used as the class of admissible controls, we need to show that it (together with the set of corresponding solutions to (3.1)) is closed under weak convergence, and that any control in it can be approximated by a piecewise-constant control, with arbitrarily small changes in the cost and solution paths. Next we show that $\bar{r}(\cdot)$ cannot be obtained from $r(\cdot)$. Then we prove the closure property and show that any control in $(\mathcal{R}, \overline{\mathcal{R}})$ can be approximated by a piecewise-constant ordinary control.

To show that we cannot get $\bar{r}(\cdot)$ from $r(\cdot)$ in general, consider the example where $u_1^n(\cdot) = u_2^n(\cdot)$ and $u_1^n(t) = 1, t \in [l/n, l/n + 1/n)$ for l odd, $u_1^n(t) = 0$ for l even. Let $\theta_2 = 0$ and $|\theta_1|$ an odd multiple of $1/n$ for the sequence n that is used. Then $r^n(\cdot) \to r(\cdot), \bar{r}^n(\cdot) \to \bar{r}(\cdot)$. The limits satisfy $r'((a,a),t) = 1/2$ for $a = 1$ or 0, and $\bar{r}'((a,a),t) = 0$ for $a = 1$ or 0. Thus $r(\cdot)$ does not yield the joint distributions of the controls, when each is evaluated at a different time, which is what is needed in (3.1). Hence $\bar{r}(\cdot)$ must be included in the definition of admissible control.

V. Limits of Convergent Sequences

The weak-sense uniqueness assumption (A4.1) means that the probability law of $(r(\cdot), \bar{r}(\cdot), w(\cdot))$ uniquely implies that of $(x(\cdot), z(\cdot), r(\cdot), \bar{r}(\cdot), w(\cdot))$. Thus we should speak of admissible sets $(r(\cdot), \bar{r}(\cdot), w(\cdot))$ instead of admissible controls $r(\cdot)$.

Let C denote the set of bounded, continuous, and real-valued functions of their arguments. The initial path data on $[-\bar{\theta}, 0]$ is denoted by $\bar{x}$.

5.1. Theorem 5.1. *Let $w^n(\cdot)$ be a Wiener process, $(r^n(\cdot), \bar{r}^n(\cdot)) \in (\mathcal{R}, \overline{\mathcal{R}})$, adapted to $w^n(\cdot)$. Let $(r^n(\cdot), \bar{r}^n(\cdot), w^n(\cdot)) \Rightarrow (r(\cdot), \bar{r}(\cdot), w(\cdot))$. Then $w(\cdot)$ is a Wiener process, $(r(\cdot), \bar{r}(\cdot)) \in (\mathcal{R}, \overline{\mathcal{R}})$, and is adapted to $w(\cdot)$. If $x^n(\cdot), z^n(\cdot)$ correspond to $(r^n(\cdot), \bar{r}^n(\cdot), w^n(\cdot))$, then*

$$(x^n(\cdot), z^n(\cdot), r^n(\cdot), \bar{r}^n(\cdot), w^n(\cdot)) \Rightarrow (x(\cdot), z(\cdot), r(\cdot), \bar{r}(\cdot), w(\cdot)),$$

where $x(\cdot), z(\cdot)$ corresponds to $(r(\cdot), \bar{r}(\cdot), w(\cdot))$. The other limit processes are nonanticipative with respect to $w(\cdot)$. The costs converge to that for the limit.

Proof. First we show the consistency of the weak-sense limit of the controls. For any $\phi(\cdot) \in C$ and $i = 1, 2$, where $d_s(s+\theta_i)$ is the s-differential at time $(s+\theta_i)$,

$$\int_0^t \int_U \phi(\alpha_i, s)\bar{r}^n(d\alpha_1\, d\alpha_2\, ds) = \int_0^t \int_{U_i} \phi(\alpha_i, s) r_i^n(d\alpha_i\, d_s(s+\theta_i)).$$

In the limit

$$\int_0^t \int_{U_i} \phi(\alpha_i, s)\bar{r}_i(d\alpha_i\, ds) = \int_0^t \int_{U_i} \phi(\alpha_i, s) r_i(d\alpha_i\, d_s(s+\theta_i)),$$

which implies that the ith-marginal $\bar{r}_i(\cdot)$ is the delayed $r_i(\cdot)$.

It follows from (3.1), (3,2), (A3.0), and Theorem 3.1, that $\{z^n(\cdot)\}$ is tight and asymptotic continuous. Then the tightness and asymptotically continuity of $\{x^n(\cdot)\}$ follows. Suppose that $(x^n(\cdot), z^n(\cdot), r^n(\cdot), \bar{r}^n(\cdot), w(\cdot)) \Rightarrow (x(\cdot), z(\cdot), r(\cdot), \bar{r}(\cdot), w(\cdot))$. Then $w(\cdot)$ must be a Wiener process.

Nonanticipativity of the limit $(x(\cdot), z(\cdot), r_i(\cdot), \bar{r}_i(\cdot), i \le k)$ with respect to $w(\cdot)$ can be shown by the following martingale method. Let $h(\cdot), \phi_m(\cdot), \bar{\phi}_m(\cdot), m \le q, f(\cdot)$, all be in C, with $f(\cdot)$ having its partial derivatives up to second order bounded and continuous. For arbitrary nonnegative p, q, t, τ, and $t_l \le t, l \le p$, by Itô's Lemma we have[3]

$$Eh\left(x^n(t_l), z^n(t_l), w^n(t_l), \langle r_i^n, \phi_m\rangle(t_l), \langle \bar{r}_i^n, \bar{\phi}_m\rangle(t_l), \right. \\ \left. l \le p, m \le q, i = 1, 2\right)[w^n(t+\tau) - w^n(t)] = 0. \tag{5.1}$$

(5.1) holds in the limit, where the n is dropped. Due to the arbitrariness of $t, \tau, t_l, \phi_m(\cdot), \bar{\phi}_m(\cdot), f(\cdot), h(\cdot), p, q, m, w(\cdot)$ is a martingale with respect to the filtration generated by $(x(\cdot), z(\cdot), r(\cdot), \bar{r}(\cdot), w(\cdot))$. Hence the other limit processes are nonanticipative.

To show that that $x(\cdot)$ satisfies (3.1) take weak-sense limits of the terms in

$$x^n(t) = x(0) + \int_0^t \int_U b(x_s^n, \alpha)\bar{r}^n(d\alpha\, ds) + \int_0^t \sigma(x_s^n) dw^n(s) + z^n(t), \tag{5.2}$$

which yields (3.1). The proof that $z(\cdot)$ is the reflection process for $x(\cdot)$ is the same as that in [10, Section 11.1.2]. Owing to the weak-sense uniqueness of the solution to (3.1), any sequence $(r^n(\cdot), \bar{r}^n(\cdot), w^n(\cdot))$ satisfying the hypotheses yields the same limit process, in the sense of measure.

The existence of an optimal admissible control is shown by taking a minimizing sequence $(r^n(\cdot), \bar{r}^n(\cdot), w^n(\cdot))$, and showing that there is a subsequence that converges to an admissible set, and that the associated costs converge to that under the associated limit process and control. ■

5.2. Approximating the controls. Introduction. To minimize the notation and algebra we henceforth suppose that all θ_i are integral multiples of the δ that are used. We can do without this assumption, at the expense of more detail. Owing to the continuity of $b(\cdot)$ and the weak-sense uniqueness in (A4.1), the process and costs can be well approximated by using fine enough finite subsets of the U_i, so we can (w.l.o.g.) suppose that $U_i = \{a_{i,j}, j\}$, a finite set for all i. Theorem 5.1 says that $(\mathcal{R}, \overline{\mathcal{R}})$ is closed in that it contains the weak-sense limits of convergent subsequences, and any weakly convergent subsequence of $(x^n(\cdot), z^n(\cdot), r^n(\cdot), \bar{r}^n(\cdot), w^n(\cdot))$ yields an adapted pair $(r(\cdot), \bar{r}(\cdot)) \in (\mathcal{R}, \overline{\mathcal{R}})$), and the limits satisfy (3.1). To show that $(\mathcal{R}, \overline{\mathcal{R}})$ can be taken to be the set of admissible controls we need to show any $r(\cdot)$ with $(r(\cdot), \bar{r}(\cdot)) \in (\mathcal{R}, \overline{\mathcal{R}})$) can be approximated by a piecewise-constant ordinary control that yields an approximation to $\bar{r}(\cdot)$. We use the definitions $\bar{r}_i(a, [b, c]) = \bar{r}_i(\{a\} \times [b, c])$ and $\bar{r}((a_1, a_2), [l\delta, l\delta + \delta)) = \bar{r}(\{a_1, a_2\} \times [l\delta, l\delta + \delta))$, with analogous definitions for $r_i(\cdot)$.

Two cases will be considered. The first where the components $u_i(\cdot)$ or $r_i(\cdot)$ can be chosen independently of one another, in that there are no a priori constraints connecting them, such as $u_1(\cdot) = u_2(\cdot)$. The second case is where the controls are dependent, in particular, $u_1(\cdot) = u_2(\cdot)$. The approximation in the latter case is more subtle, since only one control is involved, but plays a double role in that the correlation of the values at the delays must be consistent with the original pair of delayed controls.

Let $\delta > 0$ be the approximation parameter, and let $u^\delta(\cdot) = (u_i^\delta(\cdot), i \le k)$ denote an approximation that is constant on $[l\delta, l\delta + \delta), l = 0, 1, \ldots$. If there were no delays or if the system were of the form of (2.1) then the approximation is standard and simple: On $[l\delta, l\delta + \delta)$, we define $u^\delta(t) = (u_i^\delta(t), i \le k) = (a_{m_i} \in U_i, i \le k)$ on a subinterval of length $r((a_{m_i}, i \le k), [l\delta - \delta, l\delta))$. The relaxed control representation of $u^\delta(\cdot)$ converges weakly to $r(\cdot)$. The

[3] $\langle r, \phi\rangle(t) = \int_0^t \int_U \phi(\alpha, s) r(d\alpha\, ds)$.

process paths also converge weakly and we have the desired approximation.

The development of the piecewise-constant approximation when there are delays that are unequal is far more subtle, since while $u_i^\delta(t)$ is computed at time t, it is applied at $t+|\theta_i|$, and we must approximate the joint distributions of the set of delayed processes that is given by $\bar{r}(\cdot)$, the relaxed control representation of $(u(\cdot+\theta_i), i \le k)$. In all cases, we will approximate $\bar{r}(\cdot)$ on the interval $[l\delta-\delta, l\delta)$. Then it will be applied to the dynamics on $[l\delta, l\delta+\delta)$. This procedure will assure nonanticipativity.

5.3. Piecewise-constant approximations. Independent controls. The first step is to approximate $\bar{r}(\cdot)$ (or, in the case of ordinary controls, $\bar{u}(\cdot) = \{u_i(\cdot+\theta_i)\}$) since that is what appears in the dynamics. $\bar{u}^\delta(\cdot)$ must be a piecewise-costant admissible control, with all components $u_i^\delta(t)$ computed at time t. Define $\bar{u}^\delta(t) = \{u_i^\delta(t+\theta_i\}$. The construction will be in two parts. First we construct $\bar{u}^\delta(\cdot)$, an approximation to $\bar{r}(\cdot)$, and then we construct admissible $u^\delta(\cdot)$, which will yield $\bar{u}^\delta(\cdot)$ when the components are delayed.

Constructing $\bar{u}^\delta(\cdot)$. Approximate $\bar{r}(\cdot)$ on the interval $[l\delta-\delta, l\delta)$. and apply it to the dynamics on $[l\delta, l\delta+\delta)$. Let $\{\mathcal{F}_t\}$ be the basic filtration. Starting with the largest delay, $|\theta_1|$, divide $[l\delta, l\delta+\delta)$ into successive subintervals of lengths $\bar{r}_1(a_{1,j}, [l\delta-\delta, l\delta))$, $j = 1, \ldots$. As will be seen, it is important that we have subdivided for the largest delay first- to assure a nonanticipative approximation.

The next step is to take each of the subintervals of $[l\delta, l\delta+\delta)$ of length $\bar{r}_1(a_{1,j}, [l\delta-\delta, l\delta))$ that were just constructed, and further subdivide it into successive subintervals whose lengths are the joint values

$$\bar{r}((a_{1,j}, a_{2,n}), [l\delta-\delta, l\delta)), \quad a_{2,n} \in U_2,\ n = 1, \ldots.$$

Define $\bar{u}_i^\delta(\cdot), i = 1, 2$, by setting $\bar{u}^\delta(t) = (a_{1,j}, a_{2.n})$ on the (j, n)-th subinterval.[4] See Figure 5.1 for an illustration for $U_1 = U_2 = \{1, 0\}$, $a_{i,1} = 1, a_{i,2} = 0, i = 1, 2$.

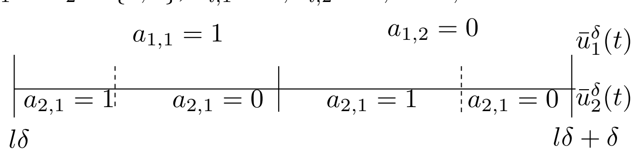

Fig. 5.1. Control-value regions for $\bar{u}_i^\delta(\cdot), i = 1, 2$.

The lengths of the subintervals that we constructed are arbitrary real numbers. If we want them to be integral multiples of some small $\rho > 0$, then further divide each of them into smaller subintervals, where δ/ρ =integer, so that the resulting controls are constant on each of them. On the "remainder" subintervals, define the controls in any way at all. If $\rho/\delta \to 0$, then this further approximation has no asymptotic consequences.

Let $\bar{r}^\delta(\cdot)$ denote the relaxed control representation of $\{\bar{u}_i^\delta(\cdot), i = 1, 2\}$. On $[l\delta, l\delta+\delta)$, the components $\bar{u}_i^\delta(\cdot)$ are $\mathcal{F}_{(l\delta+\theta_i)-}-$adapted, since $|\theta_1| > |\theta_2|$. (This would not have been the case if we subdivided first for the control with the shorter delay.) By the construction, $\bar{r}^\delta(\cdot) \Rightarrow \bar{r}(\cdot)$, as $\delta \to 0$.

Constructing $u^\delta(\cdot)$. We have just approximated $\bar{r}(\cdot)$. by the piecewise-constant $\bar{u}^\delta(\cdot)$. Now we must use $\bar{u}^\delta(\cdot)$ to construct an admissible piecewise-constant control $u^\delta(\cdot)$ which yields that $\bar{u}^\delta(\cdot)$. This is important since it is $u^\delta(t)$ that that we would actually compute at time t.

Again, we start with the component of $\bar{u}^\delta(\cdot)$ with the longest delay, namely $\bar{u}_1^\delta(\cdot)$, and define $u_1^\delta(t) = \bar{u}_1^\delta(t+|\theta_1|)$. On the interval $[l\delta, l\delta+\delta)$, $u_1^\delta(\cdot)$ is $\mathcal{F}_{l\delta}$-adapted, since its values on the subintervals of $[l\delta, l\delta+\delta)$ are determined by

$$\begin{aligned}&\bar{r}_1(a_{1,j}, [l\delta+|\theta_1|-\delta, l\delta+|\theta_1|)) \\ &\qquad = r_1(a_{1,j}, [l\delta-\delta, l\delta)),\ a_{1,j} \in U_1, j = 1, \ldots.\end{aligned}$$

Next define the component $u_2^\delta(t) = \bar{u}_2^\delta(t+|\theta_2|)$ with the next longest delay. This component is also $\mathcal{F}_{l\delta}$-adapted on $[l\delta, l\delta+\delta)$, since its values on the subintervals of $[l\delta, l\delta+\delta)$ that we have constructed are determined by

$$\bar{r}((a_{1,j}, a_{2,n}), [l\delta-\delta+|\theta_2|, l\delta+|\theta_2|)),\ a_{1,j} \in U_1, a_{2,n} \in U_2,$$

which is $\mathcal{F}_{l\delta}-$adapted since $\theta_1 < \theta_2$. Subdividing using the longest delay first, then the next longest delay, with the successive subdivisions conditioned on those for the earlier ones, assures that $u^\delta(\cdot)$ is $\mathcal{F}_{l\delta}$-adapted on $[l\delta, l\delta+\delta)$.

The procedure gives us a piecewise-constant admissible control that yields $\bar{u}^\delta(\cdot)$ when the components are appropriately delayed. Let $r^\delta(\cdot)$ denote the relaxed control representation of $u^\delta(\cdot)$.

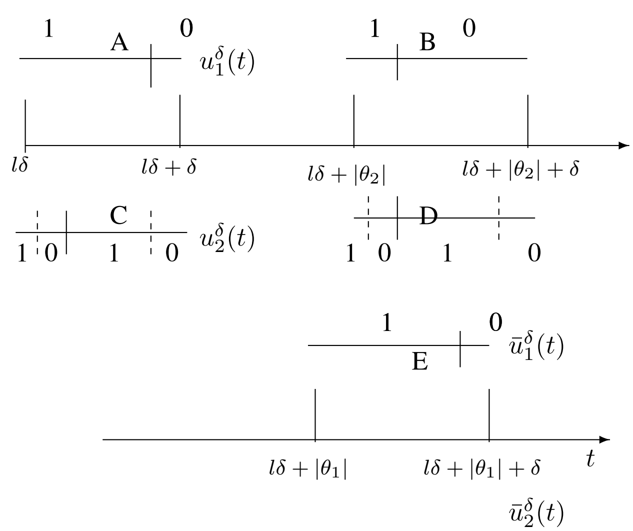

Fig 5.2. The piecewise-constant approximation $u^\delta(\cdot)$.

The construction is illustrated in Figure 5.2 for the example of Figure 5.1. The figure is split, the lower part should be to the right of the upper part. The lines B and E indicate the division of the intervals $[l\delta+|\theta_2|, l\delta+|\theta_2|+\delta)$ and $[l\delta+|\theta_1|, l\delta+|\theta_1|+\delta)$, resp., into the subintervals for the first component $\bar{u}_1^\delta(\cdot)$ on those intervals. The line E, shifted left to the interval $[l\delta, l\delta+\delta)$ gives the values of $u_1^\delta(\cdot)$ on $[l\delta, l\delta+\delta)$. This division is just the line A. The values of

[4] If there are more than two controls, then continue the subdivisions for u_3, etc.

$\bar{u}_2^\delta(\cdot)$ on $[l\delta+|\theta_2|, l\delta+|\theta_2|+\delta))$ are indicated on the line D. We now shift this to the interval $[l\delta, l\delta+\delta)$ as indicated by the line C, to get the value of $u_2^\delta(\cdot)$ on $[l\delta, l\delta+\delta)$.

Theorem 5.1 together with the next theorem shows that $(\mathcal{R}, \overline{\mathcal{R}})$ satisfies all of the requirements for being the class of admissible controls. Since we are working with weak convergence and weak-sense uniqueness, the terminology in the next Theorem is a little loose. For each δ, we should construct a new probability space with a copy of $(r^\delta(\cdot), \bar{r}^\delta(\cdot)), w(\cdot))$ and on which there are the associated solution processes to (3.1). But, we try to keep the notation simple. The meaning should be clear.

Theorem 5.2. *Let $r^\delta(\cdot)$ be the relaxed control representation of $u^\delta(\cdot)$. $\bar{r}^\delta(\cdot)$ can be constructed from the $u_i^\delta(\cdot), i=1,2$, simply by delaying the component $u_i^\delta(\cdot)$ by $|\theta_i|$. Then $(r^\delta(\cdot), \bar{r}^\delta(\cdot), w(\cdot)) \Rightarrow (r(\cdot), \bar{r}(\cdot), w(\cdot))$. If $x^\delta(\cdot)$ and $z^\delta(\cdot)$ are the path and reflection process corresponding to $(r^\delta(\cdot), \bar{r}^\delta(\cdot), w(\cdot))$, then $(x^\delta(\cdot), z^\delta(\cdot), r^\delta(\cdot), \bar{r}^\delta(\cdot), w(\cdot)) \Rightarrow (x(\cdot), z(\cdot), r(\cdot), \bar{r}(\cdot), w(\cdot))$, where $x(\cdot), z(\cdot)$ correspond to $(r(\cdot), \bar{r}(\cdot), w(\cdot))$. The costs converge as $\delta \to 0$. The set $(\mathcal{R}, \overline{\mathcal{R}})$ can be considered to be the class of admissible controls.*

Proof. Since we started with $(r(\cdot), \bar{r}(\cdot), w(\cdot))$, it is easy to show the weak convergence $(r^\delta(\cdot), \bar{r}^\delta(\cdot), w(\cdot)) \Rightarrow (r(\cdot), \bar{r}(\cdot), w(\cdot))$ from the construction. The proof of the convergence of $(x^\delta(\cdot), z^\delta(\cdot), r^\delta(\cdot), \bar{r}^\delta(\cdot), w(\cdot), \text{costs})$ is like that in Theorem 5.1. This convergence and Theorem 5.1, which showed the closure of $(\mathcal{R}, \overline{\mathcal{R}})$, shows that $(\mathcal{R}, \overline{\mathcal{R}})$ is the closure of the piecewise-constant controls, and that the use of $(\mathcal{R}, \overline{\mathcal{R}})$ as the class of admissible controls does not affect the range of the costs. ■

5.4. The piecewise-constant approximation for non-independent controls. The above construction depended on there being no constraints tying the values of the components together. If that is not the case, then the method fails. We now consider the basic case where $u_0(\cdot) = u_1(\cdot) = u_2(\cdot)$, so that at time t, we compute $u_0(t)$, and then apply it to the dynamics at times $t+|\theta_1|$ and $t+|\theta_2|$. Let U_0 be the compact control-value space for $u_0(\cdot)$. For simplicity of exposition, we continue to use $k=2$, but the conclusion holds for any value of k. Let $r_0(\cdot)$ denote the relaxed control extension of $u_0(\cdot)$, and $\bar{r}_0(\cdot)$ the relaxed control extension of $(u_0(\cdot+\theta_1), u_0(\cdot+\theta_2))$. The marginals $(\bar{r}_{0,1}(\cdot), \bar{r}_{0,2}(\cdot))$ of $\bar{r}_0(\cdot)$ are consistent with $r_0(\cdot)$ in that for $A_i \subset U_0$ and times $0 \le t_1 < t_2$, $\bar{r}_{0,i}(A_i, [t_1, t_2]) = r_0(A_i, [t_1+\theta_i, t_2+\theta_i])$. Owing to the generality of the form $b(x_t, u_0(t+\theta_1), u_0(t+\theta_2))$, we must choose the approximating control such that the joint distributions or values in $\bar{r}(\cdot)$ are approximated.

The construction. Since it is the joint values that are to be approximated, the construction will be via a conditional probability formula that approximates the joint distributions, and then a martingale approach to the law of large numbers is employed to verify the construction.

Owing to the way that law of large numbers will be used, we work with a finer interval than δ. For $\delta > 0$ and $\rho > 0$, where $\delta/\rho = q_0$, a large integer, let $u^{\delta,\rho}(\cdot)$ denote the approximating control. It will be constant on intervals $[k\rho, k\rho+\rho), k = 0, 1, \ldots$. Let a_i denote the elements of the set U_0, which, w.l.o.g., we continue to suppose is finite.

$r_0(\cdot)$ is defined on $[\theta_1, \infty)$. In $[\theta_1, \theta_1+\delta)$, use any value $u^{\delta,\rho}(t) \in U_0$. Since $\delta \to 0$, the value is unimportant. In the interval $l\delta + q\rho + \theta_1 \le \theta_2, l > 0$, construct the control via the conditional probability fromula

$$\begin{aligned} P\Big\{u_0^{\delta,\rho}(l\delta+q\rho+\theta_1) = a\Big|&\\ u_0^{\delta,\rho}(s), s < l\delta+q\rho+\theta_1; \bar{r}_0(\cdot), w(\cdot)\Big\}& \\ = r_0(a, [\theta_1+l\delta-\delta, \theta_1+l\delta))/\delta.& \end{aligned} \tag{5.3}$$

So far we have assigned values on the time interval $[\theta_1, \theta_2)$. For $t \ge \theta_2$, we are obliged to take the correlations into account. I.e., for $l\delta + q\rho \ge 0$, the value of $u_0^{\delta,\rho}(l\delta+q\rho+\theta_2)$ should depend on $u_0^{\delta,\rho}(l\delta+q\rho+\theta_1)$, which has already been assigned. This is done via the conditional probability formula

$$\begin{aligned} P\Big\{u_0^{\delta,\rho}(l\delta+q\rho+\theta_2) = a_2\Big| u_0^{\delta,\rho}(s+\theta_2), s < l\delta+q\rho;&\\ u_0^{\delta,\rho}(l\delta+q\rho+\theta_1) \in A_1; \bar{r}_0(\cdot), w(\cdot)\Big\}&\\ = P\Big\{u_0^{\delta,\rho}(l\delta+q\rho+\theta_2) = a_2\Big|&\\ u_0^{\delta,\rho}(l\delta+q\rho+\theta_1) \in A_1; \bar{r}_0(\cdot), w(\cdot)\Big\}&\\ = \frac{\bar{r}_0((A_1, a_2), [l\delta-\delta, l\delta))}{\bar{r}_0((A_1, U_0), [l\delta-\delta, l\delta))}&\\ = \frac{\bar{r}_0((A_1, a_2), [l\delta-\delta, l\delta))}{r_0(A_1, [l\delta-\delta+\theta_1, l\delta+\theta_1))}.& \end{aligned} \tag{5.4}$$

By the construction in (5.4), the distribution of the control, conditioned the entire past, depends only on the control value $|\theta_1 - \theta_2|$ units of time ago, and on $\bar{r}(\cdot), w(\cdot)$, and not on any other past controls. Since the right-hand side of (5.4) is $\mathcal{F}_{l\delta+\theta_2}$-adapted, so is $u_0^{\delta,\rho}(t+\theta_2)$ for $t \in [l\delta, l\delta+\delta)$. This completes the construction.

To get the expression for the conditional probability of $u_0^{\delta,\rho}(l\delta+q\rho+\theta_1)$ given the past, just add $\theta_1 - \theta_2 < 0$ to $l\delta$ in (5.4). This moves us back in time by $|\theta_1 - \theta_2|$ and, for $A_3 \subset U_0$, yields

$$\begin{aligned} P\Big\{u_0^{\delta,\rho}(l\delta+q\rho+\theta_1) = a_1\Big| u_0^{\delta,\rho}(s+\theta_1), s < l\delta+q\rho;&\\ u_0^{\delta,\rho}(l\delta+q\rho+2\theta_1-\theta_2) \in A_3; \bar{r}_0(\cdot), w(\cdot)\Big\}&\\ = \frac{\bar{r}_0((A_3, a_1), [l\delta+\theta_1-\theta_2-\delta, l\delta+\theta_1-\theta_2))}{\bar{r}_0((A_3, U_0), [l\delta+\theta_1-\theta_2-\delta, l\delta+\theta_1-\theta_2))}.& \end{aligned} \tag{5.5}$$

The key to the proof below is the fact that the construction implies that, in each interval $[l\delta, l\delta+\delta)$ and conditioned on $\{u_0^{\delta,\rho}(n\rho), n\rho < l\delta; \bar{r}_0(\cdot), w(\cdot)\}$, the random variables $\{u_0^{\delta,\rho}(l\delta+q\rho), q\rho < \delta\}$ are mutually independent and identically distributed. Define $\bar{u}_0^{\delta,\rho}(l\delta+q\rho) = (u_0^{\delta,\rho}(l\delta+q\rho+\theta_1), u_0^{\delta,\rho}(l\delta+q\rho+\theta_2))$. Let $r_0^{\delta,\rho}(\cdot)$ and $\bar{r}_0^{\delta,\rho}(\cdot)$ denote the relaxed control representations of $u_0^{\delta,\rho}(\cdot)$ and $\bar{u}_0^{\delta,\rho}(\cdot)$, resp. The following fact is implied by the construction. If $A_3 = U_0$ in (5.5), then $u_0^{\delta,\rho}(l\delta+q\rho+2\theta_1-\theta_2)$ is averaged out and (5.5) takes the value $r_0(a_1, [l\delta+\theta_1-\delta, l\delta+\theta_1))/\delta$.

We will need the joint probability of the controls at the two delays, conditioned on just $(\bar{r}_0(\cdot), w(\cdot))$. This is obtained by multiplying (5.4) by the probability that $u_0^{\delta,\rho}(l\delta + q\rho + \theta_1) = a_1$, conditioned on $(\bar{r}_0(\cdot), w(\cdot))$. To do this multiply the second and third lines of (5.4) by (5.5) with $A_1 = \{a_1\}$ and $A_3 = U_0$, yielding

$$P\left\{u_0^{\delta,\rho}(l\delta+q\rho+\theta_1) = a_1, u_0^{\delta,\rho}(l\delta+q\rho+\theta_2) = a_2 \middle| \bar{r}_0(\cdot), w(\cdot)\right\} = \bar{r}_0((a_1,a_2),[l\delta-\delta,l\delta))/\delta. \quad (5.6)$$

The class of admissible controls. By Theorems 5.1 and 5.3, this can be considered to be $(\mathcal{R}_0, \overline{\mathcal{R}}_0)$ in that it is closed under weak convergence (together with the associated path processes), any control in it can be approximated by a piecewise-constant control, and the infimum of the costs over this class equals that over the piecewise-constant controls.

Theorem 5.3. *Suppose that $\rho/\delta = 1/q_0 \to 0$ and let $x^{\delta,\rho}(\cdot)$ and $z^{\delta,\rho}(\cdot)$ denote the solution and reflection processes, resp., under $(r_0^{\delta,\rho}(\cdot), \bar{r}_0^{\delta,\rho}(\cdot), w(\cdot))$. As $\delta \to 0$ and $\rho/\delta \to 0$,*

$$(x^{\delta,\rho}(\cdot), z^{\delta,\rho}(\cdot), r_0^{\delta,\rho}(\cdot), \bar{r}_0^{\delta,\rho}(\cdot), w(\cdot)) \Rightarrow (x(\cdot), z(\cdot), r_0(\cdot), \bar{r}_0(\cdot), w(\cdot)), \quad (5.7)$$

solving (3.1). *The costs converge to that for the limit process.*

Proof. Let $\phi(\cdot)$ be bounded, continuous, and real-valued. By the definition of $\bar{r}_0^{\delta,\rho}(\cdot)$, with $\alpha = (\alpha_1, \alpha_2)$,

$$\int_0^t \phi(\bar{u}_0^{\delta,\rho}(s), s)ds = \int_0^t \int_{U_0\times U_0} \phi(\alpha, s)\bar{r}_0^{\delta,\rho}(d\alpha\, ds). \quad (5.8)$$

We now show that, for any $T < \infty$, as $\delta, \rho/\delta \to 0$,

$$\sup_{t\le T} E\left|\int_0^t \int_{U_0\times U_0} \phi(\alpha, s)\left[\bar{r}_0^{\delta,\rho}(d\alpha\, ds) - \bar{r}_0(d\alpha\, ds)\right]\right| \to 0. \quad (5.9)$$

By subdividing the time interval and adding the pieces we can suppose that $\phi(\cdot)$ does not depend on s, and evaluate

$$\sum_{l=0}^{t/\delta-1} E\left|\int_{U_0\times U_0}\int_{l\delta}^{l\delta+\delta} \phi(\alpha)\bar{r}_0^{\delta,\rho}(d\alpha\, ds) - \int_{U_0\times U_0}\int_{l\delta-\delta}^{l\delta} \phi(\alpha)\bar{r}_0(d\alpha\, ds)\right|.$$

Using (5.8) in the form

$$\int_{l\delta}^{l\delta+\delta}\int_{U_0\times U_0} \phi(\alpha)\bar{r}_0^{\delta,\rho}(d\alpha\, ds) = \rho\sum_{q=0}^{q_0-1}\phi(\bar{u}_0^{\delta,\rho}(l\delta+q\rho)),$$

it is sufficient to evaluate

$$E\sum_{l=0}^{t/\delta-1}\left|\rho\sum_{q=0}^{q_0-1}\phi(\bar{u}_0^{\delta,\rho}(l\delta+q\rho)) - \int_{l\delta-\delta}^{l\delta}\int_{U_0\times U_0}\phi(\alpha)\bar{r}_0(d\alpha\, ds)\right|. \quad (5.10)$$

We will evaluate each bracketed term separately, using the fact that the controls in each interval $[l\delta, l\delta+\delta)$ are mutually independent and identically distributed, given the data to $l\delta$ and $(\bar{r}_0(\cdot), w(\cdot))$. This will be done in several steps by centering the summands about their conditional expectation and using martingale inequalities. The conclusion will be that $\phi(\bar{u}_0^{\delta,\rho}(l\delta + q\rho))$ can (asymptotically) be replaced by its conditional expectation given only $(\bar{r}_0(\cdot), w(\cdot))$, which is given by (5.6). This yields the result, since (5.10) is zero with this replacement. Let E_t denote the expectation, conditioned on $\{u_0^{\delta,\rho}(s), s\le t; \bar{r}_0(\cdot), w(\cdot)\}$.

For the first step, define the conditional expectation

$$\begin{aligned}&\phi_1(u_0^{\delta,\rho}(l\delta+q\rho+\theta_1)) = \\ &E_{\{l\delta+q\rho+\theta_2-\rho\}}\phi(u_0^{\delta,\rho}(l\delta+q\rho+\theta_1), u_0^{\delta,\rho}(l\delta+q\rho+\theta_2)) = \\ &\sum_{\alpha_2}\phi(u_0^{\delta,\rho}(l\delta+q\rho+\theta_1),\alpha_2)P\left\{u_0^{\delta,\rho}(l\delta+q\rho+\theta_2)=\alpha_2\middle| u_0^{\delta,\rho}(s), s<l\delta+q\rho+\theta_2; \bar{r}_0(\cdot), w(\cdot)\right\}.\end{aligned} \quad (5.11)$$

By (5.4), this depends only on $(u_0^{\delta,\rho}(l\delta+q\rho+\theta_1), \bar{r}_0(\cdot), w(\cdot))$.

Center the terms in the inner sum (that over q) in (5.10) about their conditional expectations in (5.11) to yield

$$\left|\sum_{q=0}^{q_0-1}\rho\left[\phi(\bar{u}_0^{\delta,\rho}(l\delta+q\rho)) - \phi_1(u_0^{\delta,\rho}(l\delta+q\rho+\theta_1))\right]\right|. \quad (5.12)$$

Owing to the centering about the conditional expectation, for $q = 0, \ldots, q_0-1$, the summands in (5.12) are martingale differences with variance $O(\rho^2)$. Thus the variance of (5.12) is $O(\rho\delta)$ and the mean of its absolute value is $O(\sqrt{\rho\delta})$. Since there are t/δ such terms in (5.10) and $\rho/\delta \to 0$, it follows that the limit of (5.10) equals that of

$$E\sum_{l=0}^{t/\delta-1}\left|\rho\sum_{q=0}^{q_0-1}\phi_1(u_0^{\delta,\rho}(l\delta+q\rho+\theta_1)) - \int_{l\delta-\delta}^{l\delta}\int_{U_0\times U_0}\phi(\alpha)\bar{r}_0(d\alpha\, ds)\right|. \quad (5.13)$$

Repeat the procedure of the first step by centering the summands in the inner sum in (5.13) about their conditional expectation given the data to $l\delta + q\rho + \theta_1 - \rho$. To do this, note that, for $q < q_0$, (5.5) implies that

$$\begin{aligned}&E_{\{l\delta+q\rho+\theta_1-\rho\}}\phi_1(u_0^{\delta,\rho}(l\delta+q\rho+\theta_1)) \\ &= E_{\{l\delta+q\rho+\theta_1+(\theta_1-\theta_2)\}}\phi_1(u_0^{\delta,\rho}(l\delta+q\rho+\theta_1)) \\ &= E_{\{l\delta+q\rho+\theta_1+(\theta_1-\theta_2)\}} \\ &\quad \phi(u_0^{\delta,\rho}(l\delta+q\rho+\theta_1), u_0^{\delta,\rho}(l\delta+q\rho+\theta_2)) \\ &\equiv \phi_2(u_0^{\delta,\rho}(l\delta+q\rho+\theta_1+(\theta_1-\theta_2)).\end{aligned} \quad (5.14)$$

By the argument for the first step, the limits are unchanged if we use $\phi_2(u_0^{\delta,\rho}(l\delta+q\rho+\theta_1+(\theta_1-\theta_2))$ in (5.13). The key line in (5.14) is the third, $E_{\{l\delta+q\rho+\theta_1+(\theta_1-\theta_2)\}}\phi(\bar{u}^{\delta,\rho}(l\delta+q\rho))$.

Continue with this procedure, working backward $|\theta_1-\theta_2|$ at a time, with the conditioning on the control for each step being pushed back in time by that amount, until the conditioning on the control is in the interval $[\theta_1, \theta_2)$ and the conditional probabilities are given by (5.6) and $\phi(\bar{u}_0^{\delta,\rho}(l\delta +$

$q\rho)$ in (5.10) is replaced by $E_{\{w(\cdot),\bar{r}_0(\cdot)\}}\phi(\bar{u}_0^{\delta,\rho}(l\delta + q\rho)$ which by (5.6) is

$$\sum_{\alpha} \phi(a)\bar{r}_0(\alpha, [l\delta - \delta, l\delta))/\delta$$
$$= \int_{l\delta-\delta}^{l\delta} \int_{U_0 \times U_0} \phi(\alpha)\bar{r}_0(d\alpha\, ds)/\delta.$$

With this expression replacing $\phi(\bar{u}_0^{\delta,\rho}(l\delta + q\rho))$ in (5.10), (5.10) equals zero.

Each step introduces a further error of order $O(\sqrt{\rho\delta})$ in the approximation of the bracketed term in (5.10), and there are at most $t/|\theta_2 - \theta_1| + 1$ steps. Since there are t/δ bracketed terms in (5.10), modulo an error of magnitude $O(\sqrt{\rho\delta})/\delta$, we have that (5.10), hence (5.9), goes to zero as $\delta, \rho \to 0$..

The convergence of the paths and costs follows from a proof like that of Theorems 5.1 and 5.2 and the details are omitted. ■

5.5. Both independent and dependent controls. Suppose that the function $b(\cdot)$ in (3.1) or (2.3) is replaced by $b_1(\cdot) + b_2(\cdot)$ where in $b_1(\cdot)$ the values of the controls can be chosen independently of each other as in Subsection 5.3, and $b_2(\cdot)$ has the form used in Subsection 5.4, with the two sets being independent of each other. Then, with $r(\cdot)$ denoting the relaxed control representation of all of the controls, and $\bar{r}(\cdot), \bar{r}_0(\cdot)$ as in Subsections 5.3 and 5.4, resp., the construction of the approximations is just that of Sections 5.3 and 5.4 used for the appropriate set of controls. The set of admissible controls is $(\mathcal{R}, \overline{\mathcal{R}}, \overline{\mathcal{R}}_0)$.

References

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

[2] P. Dupuis and H. Ishii. SDE's with oblique reflection on nonsmooth domains. *Ann. Probability*, 21:554–580, 1993.

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

[4] H.J.Kushner. Numerical algorithms for optimal controls for nonlinear stochastic systems with delays. *IEEE Trans on Automatic Control*, 55:2170–2176, 2010.

[5] H.J.Kushner. Numerical methods for the optimization of nonlinear stochastic delay systems, and an application to internet regulation. In *Proceedings of the 49th Conference on Decision and Control*. IEEE Press, New York, 2010.

[6] H.J.Kushner. Numerical methods for controls for nonlinear stochastic systems with delays and jumps: Applications to admission control. *Stochastics An International Journal of Probability and Statistics*, 83:277–310, 2011. also available at http://www.dam.brown.edu/lcds/publications.

[7] T.G. Kurtz. *Approximation of Population Processes*, volume 36 of *CBMS-NSF Regional Conf. Series in Appl. Math.* SIAM, Philadelphia, 1981.

[8] H.J. Kushner. *Heavy Traffic Analysis of Controlled Queueing and Communication Networks*. Springer-Verlag, Berlin and New York, 2001.

[9] H.J. Kushner. *Numerical Methods for Controlled Stochastic Delay Systems*. Birkhäuser, Boston, 2008.

[10] H.J. Kushner and P. Dupuis. *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer-Verlag, Berlin and New York, 1992. Second edition, 2001.