Numerical Approximations for Nonlinear Stochastic Systems With Delays

Harold J. Kushner*
Division of Applied Mathematics
Lefschetz Center for Dynamical Systems
Brown University
Providence RI 02912

August 15, 2004

Abstract

We extend the numerical methods of [10], known as the Markov chain approximation methods, to controlled general nonlinear delayed reflected diffusion models. The path and the control can both be delayed. For the no-delay case, the method covers virtually all models of current interest. The method is robust, the approximations have physical interpretations as control problems closely related to the original one, and there are many effective methods for getting the approximations, and for solving the Bellman equation for low-dimensional problems. These advantages carry over to the delay problem. It is shown how to adapt the methods for getting the approximations, and the convergence proofs are outlined for the discounted cost function. Extensions to all of the cost functions of current interest as well as to models with Poisson jump terms are possible. The paper is particularly concerned with representations of the state and algorithms that minimize the memory requirements.

Key words: Optimal stochastic control, numerical methods, delay stochastic equations, numerical methods for delayed controlled diffusions, Markov chain approximation method.

Introduction

The aim of this paper is to extend the numerical methods of [10], known as the Markov chain approximation methods, to controlled delayed diffusion models. We work with a general reflected controlled diffusion model with delays. The basic idea is the approximation of the control problem by a control problem

*This work was partially supported by contract DAAD19-02-1-0425 from the Army Research Office and National Science Foundation grant ECS 0097447

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
15 AUG 2004		**00-08-2004 to 00-08-2004**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Numerical Approximations for Nonlinear Stochastic Systems with Delays	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Brown University,Division of Applied Mathematics,182 George Street,Providence,RI,02912**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**34**	

with a suitable approximating Markov chain model, solve the Bellman equation for the approximation, and then prove the convergence of the optimal costs to that for the original problem as the approximation parameter goes to zero. The method is robust, the approximations have physical interpretations as control problems closely related to the original one, and there are many effective methods for getting the approximations, and for solving the Bellman equation if the dimension is not too large. For example, for the no-delay case, the code [5, 11] exploits multigrid, accelerated Gauss-Seidel, and approximation in policy space methods.

The applications in [10] cover virtually all of the usual (non-delay) models and cost functions, including jump-diffusions, controlled variance and jumps, singular controls, etc. But to focus on the issues that are different for the delay case and not excessively repeat the development in the reference, we will consider only the diffusion case without variance control. The extensions to the more general cases cited above adapt the methods in the reference similarly to what is done for the simpler case treated here.

Models for many physical problems have reflecting boundaries. They occur naturally in models arising in queueing/communications systems [8], where the state space is often bounded owing to the finiteness of buffers and nonnegativity of their content, and the internal routing determines the reflection directions on the boundary. Numerical analysis for state-space problems is usually done in a bounded region. A standard way of bounding a state space is to impose a reflecting boundary, if one does not exist already. One selects the region so that the boundary plays a minor role. Alternatively, one can stop the process on exit from a bounded given set. In order to focus on the concepts that are new and essential in the delay case, most of the development will be for one-dimensional problems. As in [10] dimensionality is an issue only in that the computational time increases very rapidly as the dimension increases. Some comments on higher dimensional models are given at the end of Section 5.

Reference [7] has a brief discussion showing how the Markov chain approximation applies to uncontrolled delay equations. Only the simplest procedure was discussed, and there was no concern with the control problem, or with computationally efficient representations of the state. There is an extensive literature on the delayed linear system, quadratic cost, and white Gaussian noise case [2, 4, 9, 12, 15]. As for the no-delay case, this is essentially the same whether there is a driving noise or not. The problem reduces to the study of an abstract Ricatti equation. The paper [15] develops a "dual variable" approach to the problem where the control and the path variables are delayed. The development depends heavily on the linear structure and as yet has not been extended to the general nonlinear stochastic reflected diffusion problem.

The next section introduces the model and assumptions. In much of the development, for pedagogical purposes we divide the discussion into a part where only the path is delayed in the dynamics and a part where both the control and path are delayed. The cost function is confined to the discounted case. The existence of an optimal control is proved in Section 2. This is of interest for its own sake, but is also an introduction to the weak convergence methods used

for the proof of convergence of the numerical procedure as the approximation parameter goes to zero.

The basic ideas and proofs of the Markov chain approximation method are extensions of those for the no-delay case in [10], yet are not obvious. Due to this a review of the no-delay case is useful to set the stage and refresh familiarity with the basic concepts. This is done in Section 3. The fundamental assumption required for the convergence of the numerical procedure is the so-called "local consistency condition" [10]. This says no more than that the conditional mean change (resp, variance) in the state of the approximating chain is proportional to the drift (resp, covariance), modulo small errors. This would seem to be a minimal condition. In general, it need not hold everywhere (see. e.g., [10, Section 5.5]). To help focus ideas in the later discussion of the delay system, a simple example of construction of the chain is given and various related matters discussed. There are two types of construction that are of interest, called the "explicit" and "implicit" methods, owing to the similarity of one particular way of constructing them to methods of the same names for solving parabolic PDE's. Each has an interesting role to play for the delay model. In Section 4, we extend the concepts of Section 3 to the delay case. The notion of local consistency is still fundamental. The approximating chains are constructed almost exactly as they are for the no-delay case. The proofs of convergence in [10] are purely probabilistic, being based on weak convergence methods. The idea is to interpolate the chain to a continuous-time process in a suitable manner, show that the Bellman equation for the interpolation is the same as for the chain, and then show that the interpolated processes converge to an optimal diffusion as the approximating parameter goes to zero. The interpolations of interest are introduced and the convergence theorems are stated in Section 4. We try to bring the delay case into a form where the results of [10] can be appealed to for the completion of the algorithm or proof. The proof of convergence is in Section 8, where mainly the parts of the convergence proof that are different for the delay case are given. We have tried to present the minimal details that yield a coherent picture of the convergence proofs.

The state of the problem, as needed for the numerical procedure, consists of a segment of the path (over the delay interval) and of the control path as well (if the control is also delayed). This can consume a lot of memory. Section 5 is concerned with efficient representations of the state data for chains constructed by the "explicit" method. Sections 6 and 7 are concerned with the "implicit"method," which can be advantageous as far as memory requirements are concerned. In these sections, attention is confined to to the case where only the path is delayed. If the control is also delayed, then the problem is more complicated and reasonably efficient representations are not yet available.

1 The Model

The model. The maximum delay is denoted by $\tau > 0$. The solution path is denoted by $x(\cdot)$. Let $\mathcal{U}$ denote the compact control-value space. The controls $u(\cdot)$ are $\mathcal{U}$-valued, measurable and nonanticipative with respect to the driving Wiener process (this defines the set of admissible controls). $\bar{x}(t,\theta)$ is used to denote the path segment $x(t+\theta), \theta \in [-\tau, 0]$, $\tau > 0$, and we write $\bar{x}(t) = \bar{x}(t,\cdot)$. The function $\bar{u}(t)$ is defined analogously from $u(\cdot)$. The solution $x(t)$ is confined to a finite interval $G = [0, B]$ by reflection The reflection process is denoted by $z(\cdot)$. Its purpose is to assure that $x(t)$ does not escape from the interval G, and is the minimum "force" necessary. For more information on reflected (non-delay) SDE's see [10] or [8].

Delay in the path only. First consider systems where the control is not delayed and we use the reflected controlled delayed stochastic differential equation model, where $w(\cdot)$ is a standard Wiener process:[1]

$$dx(t) = b(\bar{x}(t), u(t))dt + \sigma(\bar{x}(t))dw(t) + dz(t), \quad \bar{x}(0) \text{ given.} \tag{1.1}$$

We can write $z(t) = y_1(t) - y_2(t)$, where $y_1(\cdot)$ (resp. $y_2(\cdot)$) is a continuous and nondecreasing processes that can increase only at t where $x(t) = 0$ (resp., when $x(t) = B$). Except for the LQG problem (without a reflection term) (see., e.g., [2]), little is known about the control of such systems. An example of (1.1) is

$$dx(t) = b_1(x(t), u(t))dt + b_2(x(t-\tau), u(t))dt + \sigma(x(t))dw(t) + dz(t).$$

For a set S in some metric space and $t_1 < t_2$, let $D[S; t_1, t_2]$ denote the set of S-valued functions that are right continuous on $[t_1, t_2)$, have left-hand limits on $(t_1, t_2]$, and with the Skorohod topology [1, 3]. If S is the set of real numbers, then we write just $D[t_1, t_2]$, or $D[t_1, \infty)$ if $t_2 = \infty$. Since $b(\cdot)$ depends on the "segment" of the $x(\cdot)$-process over an interval of length τ, its domain is a function space and we need to define the space of such segments. In work on the mathematics of delay equations it is common to use the path spaces $C[-\tau, 0]$ (with the sup norm topology) or $L_2[-\tau, 0]$. Any of these could be used here. But the Skorohod space $D[-\tau, 0]$ is more appropriate for the approximation and weak convergence analysis of concern and involves no loss of generality. If the model is extended to include a Poisson-type jump term, then the use of $D[-\tau, 0]$ is indispensable. Note that if $f^n(\cdot)$ converges to $f(\cdot)$ in $D[t_1, t_2]$ and $f(\cdot)$ is continuous, then the convergence is uniform on any finite interval.

We will use $\hat{x}, \hat{y}$ (or $\hat{x}(\cdot), \hat{y}(\cdot)$) to denote the canonical point in $D[G; -\tau, 0]$. A shortcoming of the Skorohod topology is that the function defined by $f(\hat{x}) = \hat{x}(t)$, for any fixed $t \in [-\tau, 0]$, is not continuous (it is measurable). But it is

[1] By adapting the techniques in [10], a driving Poisson jump process and controlled variance can also be treated, as can singular controls, boundary absorption and optimal stopping. But here we aim to concentrate on the issues arising due to the delay without overly complicating the development.

continuous at all points $\hat{x}(\cdot)$ that are continuous functions. In our case, all the solution paths $x(\cdot)$ will be continuous. The following assumption covers the common case where $b(\bar{x}(t), \alpha) = \sum_i b_i(x(t-\tau_i), \alpha), 0 \le \tau_i \le \tau$, where the $b_i(\cdot)$ are continuous.

A1.1. $b(\cdot)$ is bounded and measurable and is continuous on $D[G; -\tau, 0] \times \mathcal{U}$ at each point $(\hat{x}(\cdot), \alpha)$ such that $\hat{x}(\cdot)$ is continuous. The function $\sigma(\cdot)$ is bounded and measurable and is continuous on $D[G; -\tau, 0]$ at each point $\hat{x}(\cdot)$ that is continuous.

Relaxed controls. For purposes of proving approximation and limit theorems, it is usual and very convenient to work in terms of relaxed controls. Recall the definition of a relaxed control $m(\cdot)$ [10]. It is a measure on the Borel sets of $\mathcal{U} \times [0, \infty)$, with $m(A \times [0, \cdot])$ being measurable and nonanticipative with respect to $w(\cdot)$ for each Borel $A \in \mathcal{U}$, and satisfying $m(\mathcal{U} \times [0, t]) = t$. Write $m(A, t) = m(A \times [0, t])$. The left-hand derivative[2] $m'(d\alpha, t) = \lim_{0<\delta\to 0}[m(d\alpha, t) - m(d\alpha, t-\delta)]/\delta$ is defined for almost all (ω, t). By the definitions, $m(d\alpha\, ds) = m'(d\alpha, s)ds$. For $0 \le v \le \tau$, we write $m(d\alpha, ds - v)$ for $m(d\alpha, s - v) - m(d\alpha, s - ds - v)$. The weak topology is used on the relaxed controls. Thus $m^n(\cdot)$ converges to $m(\cdot)$ if and only if $\int\int \phi(\alpha, s)m^n(d\alpha\, ds) \to \int\int \phi(\alpha, s)m(d\alpha\, ds)$ for all continuous functions $\phi(\cdot)$ with compact support. With this topology, the space of relaxed controls is compact. An ordinary control $u(\cdot)$ can be written as the relaxed control $m(\cdot)$ defined by its derivative $m'(A, t) = I_{\{u(t)\in A\}}$, where I_K is the indicator function of the set K. Then $m(A, t)$ is the amount of time that the control takes values in the set A by time t.

Rewriting (1.1) in terms of relaxed controls yields

$$\begin{aligned} x(t) &= x(0) + \int_0^t \int_{\mathcal{U}} b(\bar{x}(s), \alpha)m(d\alpha\, ds) + \int_0^t \sigma(\bar{x}(s))dw(s) + z(t) \\ &= x(0) + \int_0^t \int_{\mathcal{U}} b(\bar{x}(s), \alpha)m'(d\alpha, s)ds + \int_0^t \sigma(\bar{x}(s))dw(s) + z(t). \end{aligned} \tag{1.2}$$

Delays in the control. We will also consider the problem where the control as well as the path is delayed. Let $\mathcal{B}[\mathcal{U}; -\tau, 0]$ be the space of measurable functions on $[-\tau, 0]$ with values in $\mathcal{U}$, and let $\hat{u}(\cdot)$ denote a canonical element of $\mathcal{B}[\mathcal{U}; -\tau, 0]$. Then the dynamical term $b(\cdot)$ becomes a function of both $\hat{x}, \hat{u}$. Depending on the applications of interest, there are a variety of choices for the way that the control appears in $b(\cdot)$. We will use the following quite general assumption, where $\bar{u}(t)$ denotes the function $u(t + \theta), \theta \in [-\tau, 0]$.

A1.2. Let $\mu(\cdot)$ be a bounded measure on the Borel sets of $[-\tau, 0]$ and let $b(\cdot)$ be a bounded measurable function on $D[G; -\tau, 0] \times \mathcal{U} \times [-\tau, 0]$. For each $v \in [-\tau, 0]$,

[2] In [10] m_t was used to denote the derivative. But this notation would be confusing in the context of the notation required to represent the various delays in this paper.

$b(\hat{x}, \alpha, v)$ is continuous in $(\hat{x}, \alpha)$ at each point $\hat{x}$ that is a continuous function. For ordinary controls $u(\cdot)$, the drift term at time t is assumed to have the form (replacing $b(\bar{x}(t), u(t))$ in (1.1))

$$\bar{b}(\bar{x}(t), \bar{u}(t)) = \int_{-\tau}^{0} b(\bar{x}(t), u(t+v), v)\mu(dv).$$

For a relaxed control $m(\cdot)$, the integral of the drift term then has the form

$$\begin{aligned} &\int_0^t \int_{-\tau}^0 \int_{\mathcal{U}} b(\bar{x}(s), \alpha, v) m'(d\alpha, s+v)\mu(dv)ds \\ &\qquad = \int_{-\tau}^0 \left[\int_0^t \int_{\mathcal{U}} b(\bar{x}(s), \alpha, v) m(d\alpha, ds+v)\right] \mu(dv). \end{aligned} \tag{1.3}$$

An example of the general form covered by (A1.2) is, for $0 \le \tau_i \le \tau$,

$$dx(t) = x(t)x(t-\tau_1)u(t-\tau_2)dt + u^2(t-\tau_3)dt + b_0(\bar{x}(t))dt + \sigma(\bar{x}(t))dw + dz(t),$$

in which case $\mu(\cdot)$ is concentrated on the two points $\{-\tau_2, -\tau_3\}$. What is not covered are "cross" terms in the control such as $u(t-\tau_1)u(t-\tau_2)$, where $\tau_1 \neq \tau_2$. The full system equation is

$$x(t) = x(0) + \int_{-\tau}^0 \left[\int_0^t \int_{\mathcal{U}} b(\bar{x}(s), \alpha, v) m(d\alpha, ds+v)\right] \mu(dv) + \int_0^t \sigma(\bar{x}(s))dw(s) + z(t), \tag{1.4}$$

and the initial data is $(\bar{x}(0), \bar{u}(0))$. Let $\bar{m}(t)$ denote the segment of $m(\cdot)$ on $[t-\tau, t]$.

Weak-sense solutions. If $w(\cdot)$ is a Wiener process on $[0, \infty)$ and $m(\cdot)$ is a relaxed control on the same probability space and it is defined on either $[-\tau, \infty)$ or $[0, \infty)$, and is nonanticipative with respect to $w(\cdot)$, then we say that the pair is admissible or, if $w(\cdot)$ is understood, that $m(\cdot)$ is admissible. Suppose that, given any admissible pair $w_1(\cdot), m_1(\cdot)$, and $\bar{x}_1(0)$ defined on the same probability space, $(m_1(\cdot), \bar{x}_1(0))$ is nonanticipative with respect to $w_1(\cdot)$, and there is a probability space on which is defined a set $(x(\cdot), w(\cdot), m(\cdot), z(\cdot))$ solving (1.1) or (1.4), where $(x(\cdot), m(\cdot), z(\cdot))$ is nonanticipative with respect to the Wiener process $w(\cdot)$, $(w(\cdot), m(\cdot), \bar{x}(0))$ has the same probability law as $(w_1(\cdot), m_1(\cdot), \bar{x}_1(0))$, and the probability law of the solution set does not depend on the probability space. Then we say that there is a solution in the weak sense [6]. If the control is delayed, then it will be defined on the interval $[-\tau, \infty)$. Thus the nonanticipativity of $(m(\cdot), \bar{x}(0))$ implies the independence of $(\bar{m}(0), \bar{x}(0))$ and $w(\cdot)$. Such independence will always hold. We always assume the following condition.

A1.3. There is a weak-sense unique weak-sense solution to (1.1) and (1.4) for each admissible relaxed control and initial data.

The techniques that are used to prove existence and uniqueness for the no-delay problem can be adapted to the delay problem. For example, use (A1.2) and the Lipschitz condition $|b(\hat{x},\alpha,v)-b(\hat{y},\alpha,v)| \le K \sup_{-\tau\le s\le 0}|x(s)-y(s)|$, and a standard Picard iteration. Alternatively, one can use the Girsanov measure transformation methods [8, 10]. See also [13, Section 1.7] for the uncontrolled problem.

The discounted cost function. To simplify the discussion, throughout the paper we focus on a discounted cost function. Let $\beta > 0$, let $c = (c_1, c_2)$ be a given constant, and let $E^m_{\hat{x}}$ denote the expectation under the initial condition $\hat{x} = \bar{x}(0)$, when the relaxed control $m(\cdot)$ is used. Then the cost function for (1.1) is

$$W(\hat{x},m) = E^m_{\hat{x}} \int_0^\infty \int_{\mathcal{U}} e^{-\beta t}\left[k(\bar{x}(t),\alpha)m'(d\alpha,t)dt + c'dy(t)\right], \tag{1.5}$$

$$V(\hat{x}) = \inf_m W(\hat{x},m),$$

where the inf is over the admissible relaxed controls, and $k(\cdot)$ is assumed to satisfy the conditions on $b(\cdot)$ in (1.1).

For a relaxed control $m(\cdot)$, let $\bar{m}(0)$ denote the segment $m(\cdot,s), s\in[-\tau,0]$. Write $\hat{m}$ for the canonical value of $\bar{m}(0)$. For (1.4), the cost function is, where $k(\cdot)$ is assumed to satisfy the conditions on $b(\cdot)$ in (A1.2),

$$W(\hat{x},m) = E^m_{\hat{x}} \int_0^\infty \int_{-\tau}^0 \int_{\mathcal{U}} e^{-\beta t}\left[k(\bar{x}(t),\alpha,v)m'(d\alpha,t+v)\mu(dv)dt + c'dy(t)\right], \tag{1.6}$$

$$V(\hat{x},\hat{m}) = \inf_m W(\hat{x},m),$$

where the infimum is over all relaxed controls with initial segments $\bar{m}(0) = \hat{m}$.[3] Recall that, in our notation, for $v \ge 0$, $m'(d\alpha, t-v)dt = m(d\alpha, dt-v)$.

2 Preliminary Results: Existence of an Optimal Control

Theorem 2.1 establishes the existence of an optimal relaxed control. Since (1.1) is a special case of (1.4), we work with (1.4). The proof of existence closely follows the standard procedure for the no-delay problem, say that of [10, Theorem 10.2.1], and we will only outline the procedure and note the differences.

[3] Since we are working with weak-sense solutions, the Wiener process might not be fixed. For example, if Girsanov measure transformation methods are used, then the Wiener process will depend on the control. Then the inf in (1.5) or (1.6) should be over all admissible pairs $(m(\cdot), w(\cdot))$, with the given initial data. But to simplify the notation, we write simply $\inf_m$. This is essentially a theoretical issue. The numerical procedures give feedback controls and all that we need to know is that there is an optimal value function to which the approximating values will converge.

Theorem 2.2 says that approximations to the controls give approximations to the cost, and the proof is nearly identical to that of Theorem 2.1. Theorem 2.3 asserts that the use of relaxed controls does not affect the infimum of the costs. The no-delay form is [10, Theorem 10.1.2] and the proof is omitted since the adjustments for the present case are readily made, given the comments in the proof of Theorem 2.1.

Theorem 2.1 *Assume* (A1.2)–(A1.3). *Then there is an optimal control for any fixed initial condition* $\hat{x}, \hat{m}$, *where* $\hat{x}$ *is continuous on* $[-\tau, 0]$. *I.e., there is a set* $(x(\cdot), w(\cdot), m(\cdot), z(\cdot))$ *solving* (1.4), *where* $(x(\cdot), m(\cdot), z(\cdot))$ *is nonanticipative with respect to the Wiener process* $w(\cdot)$, $\bar{m}(0) = \hat{m}$, $\bar{x}(0) = \hat{x}$, *and* $W(\hat{x}, m) = V(\hat{x}, \hat{m})$.

Proof. Let $m^n(\cdot)$ be a minimizing sequence of relaxed controls, with associated solutions $x^n(\cdot), z^n(\cdot)$, with $\bar{x}^n(0) = \hat{x}$, Wiener processes $w^n(\cdot)$, and $\bar{m}^n(0) = \hat{m}$. Write $z^n(\cdot) = y_1^n(\cdot) - y_2^n(\cdot)$. Thus

$$x^n(t) = \hat{x}(0)+ \int_{-\tau}^{0} \left[\int_0^t \int_{\mathcal{U}} b(\bar{x}^n(s), \alpha, v) m^n(d\alpha, ds + v) \right] \mu(dv) + \int_0^t \sigma(\bar{x}^n(s)) dw^n(s) + z^n(t). \tag{2.1}$$

The sequence $(x^n(\cdot), m^n(\cdot), y^n(\cdot), w^n(\cdot))$ is tight and all weak-sense limit processes are continuous, as follows. The tightness of the $(w^n(\cdot), m^n(\cdot))$ and the continuity of their weak-sense limits are obvious, as is the Wiener property of any weak-sense limit of $w^n(\cdot)$. The processes defined by the ordinary and stochastic integral terms of (2.1) are also tight and all weak-sense limits are continuous. The tightness and asymptotic continuity of the $y^n(\cdot)$ can be proved by contradiction. If it is false, then there will be a jump of $x^n(\cdot)$, asymptotically, into the interior of $[0, B]$, which is impossible, since the $y^n(\cdot)$ can change only on the boundary. Thus the asymptotic continuity assertion holds for $(x^n(\cdot), y^n(\cdot))$.

Now, take a weakly convergent subsequence with limit $(x(\cdot), m(\cdot), y(\cdot), w(\cdot))$, index it by n also, and use the Skorohod representation [3, page 102], so that we can assume that the convergence is w.p.1 in the topologies of the spaces of concern. By the weak convergence, we must have $x(t) \in [0, B]$, and that $y_1(\cdot)$ (resp., $y_2(\cdot)$) can change only at t where $x(t) = 0$ (resp., where $x(t) = B$). Since $\sup_{s \le t} |x^n(s) - x(s)| \to 0$ for each $t > 0$, $\bar{x}^n(t) \to \bar{x}(t)$, uniformly on any finite interval, all w.p.1. Then, (A1.2) implies that, for all $v \in [-\tau, 0]$,

$$\sup_{s \le t, \alpha} |b(\bar{x}^n(s), \alpha, v) - b(\bar{x}(s), \alpha, v)| \to 0$$

w.p.1, and also for $\sigma(\cdot), k(\cdot)$ replacing $b(\cdot)$. The last sentence and the continuity and boundedness assumptions (A1.2) yield

$$\int_0^t \int_{\mathcal{U}} b(\bar{x}^n(s), \alpha, v) m^n(d\alpha, ds - v) \to \int_0^t \int_{\mathcal{U}} b(\bar{x}(s), \alpha, v) m(d\alpha, ds - v)$$

for all $t \geq 0, v \in [-\tau, 0]$, w.p.1. From this it follows that the first integral in (2.1) converges to the process obtained when the superscript n is dropped.

Nonanticipativity is shown as follows, also following the reference. Let $g_j(\cdot), j \leq J$, be continuous functions with compact support and write

$$\langle m, g_j \rangle (t) = \int_0^t \int_{\mathcal{U}} g_j(\alpha, s) m(d\alpha \, ds).$$

For arbitrary $t > 0$ and integer $I > 0$, let $s_i \leq t$ for $i \leq I$, and let $h(\cdot)$ be an arbitrary bounded and continuous function. By the nonanticipativity for each n,

$$\begin{aligned} Eh\left(x^n(s_i), w^n(s_i), y^n(s_i), \langle m^n, g_j \rangle (s_i), i \leq I, j \leq J\right) \times \\ (w^n(t+\tau) - w^n(t)) = 0. \end{aligned} \tag{2.2}$$

By the weak convergence and the continuity of the limit processes, (2.2) holds with the superscript n dropped. Now, the arbitrariness of $h(\cdot), I, J, s_i, t, g_j(\cdot)$ implies that $w(\cdot)$ is a martingale with respect to the sigma-algebra generated by $(x(\cdot), w(\cdot), z(\cdot), m(\cdot))$. Hence the nonanticipativity of the limit processes.

The convergence of the stochastic integral is obtained by an approximation argument. For a measurable function $f(\cdot)$ and $\kappa > 0$, let $f_\kappa(\cdot)$ be the approximation that takes the value $f(n\kappa)$ on $[n\kappa, n\kappa + \kappa)$. Then, by the weak convergence,

$$\int_0^t \sigma(\bar{x}_\kappa^n(s)) dw^n(s) \to \int_0^t \sigma(\bar{x}_\kappa(s)) dw(s),$$

The left side can be made arbitrarily close to the stochastic integral in (2.1), in the mean square sense, uniformly in n, by choosing κ small enough. By the nonanticipativity, as $\kappa \to 0$ the right hand term converges to the stochastic integral with $\bar{x}(\cdot)$ replacing $\bar{x}_\kappa(\cdot)$. Finally, by the weak convergence and the minimizing property of $m^n(\cdot)$, $W(\hat{x}, m^n) \to W(\hat{x}, m) = V(\hat{x})$, the infimum of the costs. ■

Theorem 2.2. *Assume* (A1.2)–(A1.3). *Let admissible* $(\bar{x}^n(0), \bar{m}^n(0))$ *converge weakly to* $(\hat{x}, \hat{m})$. *Then* $V(\bar{x}^n(0), \bar{m}^n(0)) \to V(\hat{x}, \hat{m})$.

The next theorem asserts that the use of relaxed controls does not change the minimal values. See [10, Theorem 10.1.2] for the no-delay case. The proof depends on the fact that for any relaxed control $m(\cdot)$ one can find a sequence of ordinary controls $u^n(\cdot)$, each taking a finite number of values in $\mathcal{U}$, such that $(\hat{x}, m^n(\cdot), w(\cdot))$ converges weakly to $(\hat{x}, m(\cdot), w(\cdot))$ where $m^n(\cdot)$ is the relaxed control representation of $u^n(\cdot)$.

Theorem 2.3. *Assume* (A1.2)–(A1.3). *Fix the initial control segment* $u_1(\theta), \theta \in [-\tau, 0]$, *and let* $\bar{m}_1(0)$ *be the relaxed control representation of this ordinary control segment* $\bar{u}_1(0)$. *Then*

$$\inf_{u(\cdot)} W(\hat{x}, u) = \inf_m W(\hat{x}, m),$$

where the $\inf_u$ *(resp.,* $\inf_m$*) is over all controls (resp., relaxed controls) with initial segments* $\bar{u}_1(0)$ *(resp.,* $\bar{m}_1(0)$*.*

3 A Markov Chain Approximation: The No-Delay Case

Local consistency. In this section we review the basic ideas for the no-delay case as preparation for the treatment of the delay case in the next section. Keep in mind that the restriction to one dimension is for expository simplicity only. The numerical method is based on a Markov chain approximation to the diffusion [10]. For an approximation parameter $h > 0$, and B assumed to be an integral multiple of h, discretize the interval $[0, B]$ as $G_h = \{0, \ldots, B-h, B\}$. At 0 or B, the process (1.1) or (1.4) is still a diffusion and the drift or stochastic integral terms might try to force it out of $[0, B]$. But then the reflection process cancels those effects and prevents exit. For the approximation, the analog is to still approximate the diffusion at 0 and B, as at any interior point of G_h. But if the chain goes from 0 to $-h$ or from B to $B+h$, it will be *immediately* reflected back to G_h. The full state space, including the reflecting boundary points, is denoted by $G_h^+ = \{-h, 0, h, \ldots, B, B+h\}$. Let $\mathcal{U}^h$ be a finite set such that the Hausdorff distance between $\mathcal{U}^h$ and $\mathcal{U}$ goes to zero as $h \to 0$. Let $\{\xi_n^h, n < \infty\}$ be a controlled discrete parameter Markov chain on the discrete state space G_h^+ with transition probabilities denoted by $p^h(x, y|\alpha)$. The α is the control parameter and takes values in $\mathcal{U}^h$. We use u_n^h to denote the random variable which is the actual control action for the chain at discrete time n. In addition, suppose that we have an "interpolation interval" $\Delta t^h(x, \alpha)$ such that $\sup_{x,\alpha} \Delta t^h(x, \alpha) \to 0$ as $h \to 0$, but $\Delta t^h(x, \alpha) > 0$ for each $h > 0$ and $x \in G_h$. As will be seen, getting such an interval is always an automatic byproduct of getting the transition probabilities. Define $\Delta t_n^h = \Delta t^h(\xi_n^h, u_n^h)$.

The distribution of ξ_{n+1}^h, conditioned on $\xi_i^h, u_i^h, i \le n$, will depend only on ξ_n^h, u_n^h, and not on n otherwise. Thus, let $E_x^{h,\alpha}$ denote the conditional expectation given $u_0^h = \alpha, \xi_0^h = x$. Define $\Delta\xi_n^h = \xi_{n+1}^h - \xi_n^h$ and the martingale difference[4]

$$\beta_n^h = \left[\Delta\xi_n^h - E\left\{\Delta\xi_n^h \middle| \xi_i^h, u_i^h, i \le n\right\}\right] I_{\{\xi_n^h \in G_h\}}.$$

The key condition for convergence of the numerical procedure is the following "local consistency" condition for $\xi_n^h, \Delta t_n^h$. The equalities define $b^h(\cdot), a^h(\cdot)$. For

[4]Here and in the sequel, when we say that some process derived from the chain is a martingale or martingale difference, the relevant filtration is that generated by the path and control data.

$x \in G_h$ we require:

$$\begin{aligned}
&E_x^{h,\alpha}\Delta\xi_0^h = b^h(x,\alpha)\Delta t^h(x,\alpha) = b(x,\alpha)\Delta t^h(x,\alpha) + o(\Delta t^h(x,\alpha)),\\
&E_x^{h,\alpha}\beta_0^h[\beta_0^h]' = a^h(x)\Delta t^h(x,\alpha) = a(x)\Delta t^h(x,\alpha) + o(\Delta t^h(x,\alpha)),\\
&a(x) = \sigma(x)\sigma'(x),\\
&\sup_{\omega,n}|\xi_{n+1}^h - \xi_n^h| \xrightarrow{h} 0, \quad \sup_{x,\alpha}\Delta t^h(x,\alpha) \to 0.
\end{aligned} \tag{3.1}$$

Set $p^h(-h,0|\alpha) = p^h(B+h,B|\alpha) = 1$ and $\Delta t^h(-h,\alpha) = \Delta t^h(B+h,\alpha) = 0$, so that the states $-h, B+h$ are instantaneously reflecting.

An example of construction of the transition probabilities. Straightforward and automatic procedures for getting the $p^h(\cdot)$ are discussed in $\Delta t^h(\cdot)$ [10]. To facilitate understanding the general method of adapting the algorithms to the delay case in the next section, we will work with one particular simple approximation. This method is used *for illustrative purposes only* and any of the procedures for getting the approximating chain in [10] could be used instead. The method to be discussed gets suitable approximations whether or not the $W(\cdot)$ introduced below has any derivatives. Its use of finite differences is only a formal device. The proofs of convergence are all probabilistic, somewhat analogous to that of Theorem 2.1. Let $W(\cdot)$ be a purely formal solution to the PDE

$$\frac{1}{2}\sigma^2 W_{xx}(x) + bW_x(x) + k(x,\alpha) = 0,$$

where $W_x(\cdot)$ and $W_{xx}(\cdot)$ denote the first and second derivatives with respect to x. Suppose that $\sigma^2(x) \geq h|b(x,\alpha)|$ for all x, α, and use the finite difference approximations

$$\begin{aligned}
f_{xx}(x) &\to \frac{f(x+h) + f(x-h) - 2f(x)}{h^2},\\
f_x(x) &\to \frac{f(x+h) - f(x-h)}{2h}.
\end{aligned} \tag{3.2}$$

For $x \in G_h$, this leads to the finite difference approximation

$$\begin{aligned}
&W^h(x) =\\
&p^h(x,x+h|\alpha)W^h(x+h) + p^h(x,x-h|\alpha)W^h(x-h) + \Delta t^h(x,\alpha)k(x,\alpha),
\end{aligned}$$

where

$$p^h(x,x\pm h|\alpha) = \frac{\sigma^2(x) \pm hb(x,\alpha)}{2\sigma^2(x)}, \quad \Delta t^h(x,\alpha) = \frac{h^2}{\sigma^2(x)}, \quad x \in G_h. \tag{3.3}$$

Condition (3.1) clearly holds, so that the terms in (3.3) can serve as the transition probabilities and interpolation interval for the approximating chain.[5] In

[5] If $\sigma^2(x) < h|b(x,\alpha)|$ at some x, α, then one-sided difference approximations can be used there to get the appropriate $p^h(\cdot), \Delta t^h(\cdot)$ [10].

the delay case, the correct functional dependence of $b(\cdot), \sigma(\cdot), k(\cdot)$ on the path segment will need to be used.

The Bellman equation. After getting the approximating chain, one approximates the cost and writes the Bellman equation, which is

$$V^h(x) = \inf_{\alpha \in \mathcal{U}^h} \left\{ e^{-\beta \Delta t^h(x,\alpha)} \sum_y p^h(x, y|\alpha) V^h(y) + k(x,\alpha) \Delta t^h(x,\alpha) \right\}, \quad x \in G_h, \tag{3.4a}$$

and, for the boundary states,

$$V^h(-h) = V^h(0) + c_1 h, \quad V^h(B+h) = V^h(B) + c_2 h. \tag{3.4b}$$

The form (3.4a) shows why the $\Delta t^h(x,\alpha)$ were called interpolation intervals. The procedure just used might be referred to henceforth as the "explicit" method, in analogy to the procedure of the same name based on the approximation (3.2) for solving parabolic PDE's by finite differences. A special case of the results in [10, Chaper 10] is that $V^h(x) \to V(x)$, the minimal cost, as $h \to 0$.

A continuous-time approximating Markov process. The following fact will be useful later. Suppose that the ξ_n^h were replaced by a *continuous-time* Markov chain $\psi^h(\cdot)$ with transition probabilities $p^h(x, y|\alpha)$ and whose mean holding times, when in state x with control value α used, are $\Delta t^h(x,\alpha)$. Let the cost function be just the no-delay form of (1.5). Then the Bellman equation is still (3.4), modulo an asymptotically negligible difference in the cost rate and discount factor [10, Section 4.3]. Thus, either model (ξ_n^h or $\psi^h(\cdot)$) could be used to study the convergence of the numerical procedure.

Constant interpolation interval. For simplicity of coding and considerations of memory requirements in the next section, it will be useful to have $\Delta t^h(\cdot)$ not depending on the state or control. This is easily arranged and the desired transition probabilities and interpolation interval are readily obtained from the $p^h(\cdot), \Delta t^h(\cdot)$ above, as follows. Define the new interpolation interval $\overline{\Delta}^h = \inf_{\alpha, \xi \in G_h} \Delta t^h(\xi, \alpha)$. The possibility that $\overline{\Delta}^h < \Delta t^h(x,\alpha)$ at some x, α is compensated for by allowing the state x to communicate with itself at that point. Let $\bar{p}^h(x, y|\alpha)$ denote the new transition probabilities. Conditioned on the event that a state does not communicate with itself on the current transition, the transition probabilities are as in (3.3). Thus, the general formula for getting them from the $p^h(\cdot)$ is ([10, Section 7.7])

$$\begin{aligned} \bar{p}^h(x, y|\alpha) &= p^h(x, y|\alpha)(1 - \bar{p}^h(x, x|\alpha)), \quad \text{for } x \neq y, \\ \bar{p}^h(x, x|\alpha) &= 1 - \frac{\overline{\Delta}^h}{\Delta t^h(x,\alpha)}. \end{aligned} \tag{3.5}$$

Continuous-time interpolations. The proofs of convergence in [10] depend on continuous-time interpolations of the ξ_n^h process. The simplest interpolation,

called $\xi^h(\cdot)$, is piecewise constant with intervals Δt_n^h. For any t, the interval $[t,t)$ is considered to be empty. Define $t_n^h = \sum_{i=0}^{n-1} \Delta t_i^h$. and define $\xi^h(t) = \xi_n^h$ and $u_\xi^h(t) = u_n^h$ for $t \in [t_n^h, t_{n+1}^h)$. Suppose that $\xi_n^h = -h$ or $B+h$, a reflecting state. Then $\Delta t_n^h = 0$ and the interval $[t_n^h, t_n^h + \Delta t_n^h)$ is empty. This implies the important fact that *the values of the reflecting states are ignored in constructing the continuous-time interpolation.* This will always be the case. Let $m_\xi^h(\cdot)$ denote the relaxed control representation of $u_\xi^h(\cdot)$. It is constant on the intervals $[t_n^h, t_{n+1}^h)$ and the derivative on that interval is defined by $m^{h\prime}(A,t) = I_{\{u_n^h \in A\}}$. Then, by (3.1),

$$\xi_{n+1}^h = \xi_n^h + b^h(\xi_n^h, u_n^h)\Delta t_n^h + \beta_n^h + \delta z_n^h, \tag{3.6}$$

where $\delta z_n^h = \delta y_{1,n}^h - \delta y_{2,n}^h$, where $\delta y_{1,n}^h = hI_{\{\xi_n^h = -h\}}$ and $\delta y_{2,n}^h = hI_{\{\xi_n^h = B+h\}}$. For each $t \geq 0$, define the stopping time, where $\sum_0^{-1} = 0$,

$$d^h(t) = \max\left\{ n : \sum_{i=0}^{n-1} \Delta t_i^h = t_n^h \leq t \right\}.$$

Note that $d^h(t)$ will *never* be the index of a reflecting state, since the time intervals for those are zero. Then

$$\xi^h(t) = \xi^h(0) + \sum_{i=0}^{d^h(t)-1} b^h(\xi_i^h, u_i^h)\Delta t_i^h + B^h(t) + z_\xi^h(t), \tag{3.7}$$

where

$$B^h(t) = \sum_{i=0}^{d^h(t)-1} \beta_i^h, \quad z_\xi^h(t) = \sum_{i=0}^{d^h(t)-1} \delta z_i^h.$$

In interpolated and relaxed control form,

$$\xi^h(t) = x(0) + \int_0^t \int_{\mathcal{U}^h} b^h(\xi^h(s), \alpha) m_\xi^h(d\alpha\, ds) + B^h(t) + z_\xi^h(t).$$

Although the fact will not be used in the sequel, it is interesting to note that $B^h(\cdot)$ is an "approximation to a stochastic integral in the following sense. There are martingale differences δw_n^h whose continuous-time interpolation (intervals Δt_n^h) converges weakly to a standard Wiener process and $\beta_n^h \approx \sigma(\xi_n^h)\delta w_n^h$ in the sense that [7, Section 6.6]

$$\sum_{n=1}^{d^h(t)-1} \beta_n^h = \sum_{n=1}^{d^h(t)-1} \sigma(\xi_n^h)\delta w_n^h + \text{asymptotically negligible error.}$$

There is another continuous-time interpolation that will be important. Recall the comment in the paragraph below (3.4) concerning the equivalence of the Bellman equations for the sequence ξ_n^h and the continuous-time Markov process $\psi^h(\cdot)$. In the proofs of convergence in [10, Chapter 10], this latter process was

used since it simplified the proof. It will also be used here for that purpose. For the delay case, both interpolations $\xi^h(\cdot)$ and $\psi^h(\cdot)$ play a role due to the presence of the segment $\bar{x}(t)$ in the dynamics and the need to approximate it in the dynamics of the approximating chain. Let us now formalize the definition of $\psi^h(\cdot)$. Let ν_n^h be i.i.d. exponentially distributed variables with unit mean, and independent of $\{\xi_n^h, u_n^h\}$. Define the intervals $\Delta\tau_n^h = \nu_n^h \Delta t_n^h$. Then $\psi^h(\cdot)$ denotes the continuous-time interpolation of the sequence $\{\xi_n^h\}$ with intervals $\{\Delta\tau_n^h\}$. Let $u_\psi^h(\cdot)$ denote the continuous-time interpolation of the controls $\{u_n^h\}$ with intervals $\{\Delta\tau_n^h\}$, and let $m_\psi^h(\cdot)$ be its relaxed control representation.[6] Analogously, let $M^h(\cdot)$ (resp., $z_\psi^h(\cdot), y_{i,\psi}^h(\cdot)$) be the continuous-time interpolation of $\{\beta_n^h\}$ (resp., of $\{\delta z_n^h, \delta y_{i,n}^h\}$) with intervals $\{\Delta\tau_n^h\}$. Write $y_\psi^h(\cdot) = (y_{1,\psi}^h(\cdot), y_{2,\psi}^h(\cdot))$ and $z_\psi^h(\cdot) = y_{1,\psi}^h(\cdot) - y_{2,\psi}^h(\cdot)$.

By the definition, $\psi^h(\cdot)$ is a continuous-time controlled Markov chain. We have, where now all processes are interpolated with the intervals $\Delta\tau_n^h$,

$$\psi^h(t) = x(0) + \int_0^t \int_{\mathcal{U}^h} b^h(\psi^h(s), \alpha) m_\psi^h(d\alpha\, ds) + M^h(t) + z_\psi^h(t). \tag{3.8}$$

The process $M^h(\cdot)$ can be represented as [10, Section 10.4.1]

$$M^h(t) = \int_0^t \sigma^h(\psi^h(s)) dw^h(s) = \int_0^t \sigma(\psi^h(s)) dw^h(s) + \epsilon^h(t),$$

where $\epsilon^h(\cdot)$ converges weakly to the zero process and $w^h(\cdot)$ is a martingale with quadratic variation It and which converges weakly to a standard Wiener process. In the proofs in [10, Chapter 10] it is shown that $\psi^h(\cdot)$ converges to an optimal limit process, with optimal cost $V(x) = \lim_h V^h(x)$. The analog of these facts will also be true when there are delays.

The interpolations $\xi^h(\cdot)$ and $\psi^h(t)$ are asymptotically identical. They are both continuous-time scalings of the basic chain ξ_n^n, and the scalings are asymptotically identical. This fact was not needed in the proof of the classical no-delay case, but will be important in treating the delay case. The following theorem holds for the delay case in the next section as well.

Theorem 3.1. *Recall the definition of $d^h(\cdot)$ below* (3.6). *Then, for each $t > 0$,*

$$\lim_{h\to 0} E \sup_{s\le t} \left[\sum_{i=0}^{d^h(s)} (\Delta\tau_i^h - \Delta t_i^h) \right]^2 = 0. \tag{3.9}$$

Proof. Owing to the mutual independence of the exponential random variables $\{\nu_n^h\}$ and their independence of everything else, the discrete parameter process

[6] In [10], these were called just $u^h(\cdot)$ and $m^h(\cdot)$. But for the delay case both interpolations $\xi^h(\cdot)$ and $\psi^h(\cdot)$ are used and we need to distinguish them.

$L_n = \sum_{i=0}^{n}(\Delta\tau_i^h - \Delta t_i^h)$ is a martingale. Hence, the conditional expectation of the squared term in (3.9) given the $\{\Delta t_i^h\}$ equals

$$E\left[\sum_{i=0}^{d^h(t)}\left[\Delta\tau_i^h - \Delta t_i^h\right]^2 \middle| \Delta t_i^h, i < \infty\right] = \sum_{i=0}^{d^h(t)}[\Delta t_i^h]^2 \le (t + \sup_n \Delta t_n^h)\sup_n \Delta t_n^h \xrightarrow{h} 0,$$

which yields the desired result since $E\sup_{j\le n}|L_j|^2 \le 4E|L_n|^2$. ■

The "implicit" approximation. An alternative to the method of obtaining a Markov chain approximation that was illustrated by the use of (3.2) uses what has been called an "implicit" method of approximation [10, Chapter 12], again owing to its similarity to the implicit method of dealing with numerical solutions of parabolic PDE's via finite differences, although it is used here only as a *heuristic guide* to the construction and no assumption concerning differentiability is made. The fundamental difference between the explicit and implicit approaches to the Markov chain approximation lies in the fact that in the former the time variable is treated differently than the state variables: It is a true "time" variable, and its value increases by either (depending on the interpolation used) Δt_n^h or $\nu_n^h \Delta t_n^h$ at step n. In the implicit approach, the time variable is treated as just another state variable. It is discretized in the same manner as are the other state variables: For the no-delay case, the approximating Markov chain has a state space that is a discretization of the $(x,t)-$space, and the component of the state of the chain that comes from the original time variable does not necessarily increase its value at each step. The idea is analogous when there are delays, and leads to some interesting and possibly more efficient numerical schemes. Let $\delta > 0$ be the discretization level for the time variable. For the no-delay case, the "implicit procedure" analog of the transition probabilities of (3.3), obtained via use of (3.2), would start with the finite difference approximations of the form

$$\begin{aligned} f_t(x,t) &\to \frac{f(x,t+\delta) - f(x,t)}{\delta} \\ f_x(x,t) &\to \frac{f(x+h,t) - f(x-h,t)}{2h} \\ f_{xx}(x,t) &\to \frac{f(x+h,t) + f(x-h,t) - 2f(x,t)}{h^2}. \end{aligned} \tag{3.10}$$

Note that the approximations for f_x and f_{xx} are made at t and not $t+\delta$. Denote the chain by $\zeta_n^{h,\delta} = (\phi_n^{h,\delta}, \xi_n^{h,\delta})$= (time, space) variables. [7]

The general implicit method. Analogously to the method of going from the formal approximation to the PDE above (3.2) to (3.3), the transition probabilities and interpolation interval can be determined by substituting (3.10) into the

[7] Again, if $\sigma^2(x) < h|b(x,\alpha)|$ at some x, α, then one-sided difference approximations can be used there to get the appropriate $p^h(\cdot), \Delta t^h(\cdot)$ [10].

PDE $W_t(x,t) + [\sigma^2/2]W_{xx}(x,t) + bW_x(x,t) + k(x,t) = 0$ and collecting coefficients. But there is a general method that starts with the $p^h(\cdot), \Delta t^h(\cdot)$. Suppose that at the current step the time variable does not advance. Then, conditioned on this event and on the value of the current spatial state, the distribution of the next spatial state is just the $p^h(x, y|\alpha)$ used previously. So one needs only determine the conditional probability that the time variable advances, conditioned on the current state. This is obtained by a "local consistency" argument and no matter how the $p^h(\cdot)$ were derived, the (no-delay) transition probabilities $p^{h,\delta}(\cdot)$ and interpolation interval $\Delta t^{h,\delta}(\cdot)$ for the implicit procedure can be determined from the $p^h(\cdot), \Delta t^h(\cdot)$ by the formulas [10, Section 12.4], for $x \in G_h$,

$$\begin{aligned} p^h(x, y|\alpha) &= \frac{p^{h,\delta}(x, n\delta; y, n\delta|\alpha)}{1 - p^{h,\delta}(x, n\delta; x, n\delta + \delta|\alpha)}, \\ p^{h,\delta}(x, n\delta; x, n\delta + \delta|\alpha) &= \frac{\Delta t^h(x,\alpha)}{\Delta t^h(x,\alpha) + \delta}, \\ \Delta t^{h,\delta}(x,\alpha) &= \frac{\delta \Delta t^h(x,\alpha)}{\Delta t^h(x,\alpha) + \delta}. \end{aligned} \tag{3.11}$$

These formulas hold provided only that no state communicates with itself under the $p^h(\cdot)$. The reflecting states $x = -h$ and $B + h$ are treated as before. In the no-delay case, the implicit procedure was used in [10] mainly to deal with control problems that were defined over a fixed finite time interval. It will be used in a quite different way in the delay case.

4 The System With Delays: Consistency and Convergence

Local consistency conditions; delay in path only. The approach is analogous to what was done for the no-delay case. The main issues concern accounting for the fact that $b(\cdot), \sigma(\cdot)$ and $k(\cdot)$ depend on the solution path over an interval of length τ. We will construct a controlled process $\xi_n^h, n \geq 0$, and interpolation intervals $\Delta t_n^h, n \geq 0$, in much the same way as was done in Section 3. Details of a construction analogous to that of (3.2) and (3.3) are in the next section. The initial condition $\bar{x}(0)$ for (1.1) is an arbitrary continuous function. The numerics work on a discrete space. so this function will have to be approximated. The exact form of the approximation is not important at this point, and we simply assume that we use a sequence $\bar{\xi}_0^h \in D[G_h; -\tau, 0]$, that is piecewise constant and that converges to $\bar{x}(0)$ uniformly on $[-\tau, 0]$.

Given the $\xi_n^h, \Delta t_n^h$, define $t_n^h = \sum_{i=0}^{n-1} \Delta t_i^h$. Define $\xi^h(\cdot)$ such that on $[0, \infty)$, it is the continuous-time interpolation of $\{\xi_n^h\}$ with intervals $\{\Delta t_n^h\}$, as in Section 3, and the segment on $[-\tau, 0]$ is $\bar{\xi}_0^h$, where $\xi_0^h = \bar{\xi}_0^h(0)$. Define

$$\bar{\xi}^h(t, \theta) = \xi^h(t + \theta), \quad \text{for } \theta \in [-\tau, 0],$$

and set $\bar{\xi}_n^h(\cdot) = \bar{\xi}^h(t_n^h, \cdot)$, the interpolated path on $[t_n^h - \tau, t_n^h]$. We will also write $\bar{\xi}_n^h = \bar{\xi}_n^h(\cdot)$. Note that $\bar{\xi}_n^h(\cdot)$ is discontinuous at $t = 0$. The value there will be ξ_n^h if that is in G_h. If ξ_n^h is a reflecting state, then the value is the closest one on G_h. We need to define a path segment that plays the role of $\bar{x}(t)$. There is some flexibility in the way that this approximation is constructed from the $\{\xi_n^h\}$. The choice influences the computational complexity, and we return to this issue later. Until further notice, we use $\bar{\xi}_n^h$.

Analogously to the no-delay case, the chain and intervals are assumed to satisfy the following properties (an example is given below). There is a function $\Delta t^h(\cdot)$ on $D[G_h; -\tau, 0] \times \mathcal{U}^h$ such that $\Delta t_n^h = \Delta t^h(\bar{\xi}_n^h, u_n^h)$ where u_n^h is the control applied at time n. Recall the definitions of $\Delta\xi_n^h$ and of β_n^h above (3.1). The distribution of ξ_{n+1}^h, given the initial data and $\xi_i^h, u_i^h, i \le n$, will depend only on $\bar{\xi}_n^h, u_n^h$ and not otherwise on n, analogously to the case in Section 3. Thus, let $E_{\hat{\xi}}^{h,\alpha}$ denote the expectation under control value $u_0^h = \alpha$ and $\bar{\xi}_0^h = \hat{\xi} \in D[G_h; -\tau, 0]$ and is piecewise constant. Analogously to the no-delay case, local consistency for $\xi_n^h, \Delta t_n^h$ is said to hold if, for $\xi_0^h = \hat{\xi}(0)$, ((4.1) defines $b^h(\cdot), a^h(\cdot)$)

$$\begin{aligned}
&E_{\hat{\xi}}^{h,\alpha}\Delta\xi_0^h = b^h(\hat{\xi},\alpha)\Delta t^h(\hat{\xi},\alpha) = b(\hat{\xi},\alpha)\Delta t^h(\hat{\xi},\alpha) + o(\Delta t^h(\hat{\xi},\alpha)),\\
&E_{\hat{\xi}}^{h,\alpha}\beta_0^h[\beta_0^h]' = a^h(\hat{\xi})\Delta t^h(\hat{\xi},\alpha) = a(\hat{\xi})\Delta t^h(\hat{\xi},\alpha) + o(\Delta t^h(\hat{\xi},\alpha)),\\
&a(\hat{\xi}) = \sigma(\hat{\xi})\sigma'(\hat{\xi}),\\
&\sup_{n,\omega}|\xi_{n+1}^h - \xi_n^h| \xrightarrow{h} 0, \quad \sup_{\hat{\xi},\hat{u}} \Delta t^h(\hat{\xi},\hat{u}) \xrightarrow{h} 0.
\end{aligned} \tag{4.1}$$

The reflecting boundary is treated exactly as for the no-delay case below (3.1). In particular, if $\xi_n^h = -h$ (resp., $B + h$), then $\xi_{n+1}^h = 0$ (resp., B) and the interpolation interval is zero.

Let $E_{\hat{\xi}}^{h,u^h}$ denote the expectation given initial condition $\hat{\xi} = \bar{\xi}_0^h$ and control sequence $u^h = \{u_n^h, n < \infty\}$. The cost function for the chain is

$$\begin{aligned}
W^h(\hat{\xi}, u^h) = E_{\hat{\xi}}^{h,u^h}\sum_{n=0}^{\infty} e^{-\beta t_n^h}\left[k(\bar{\xi}_n^h, u_n^h)\Delta t^h(\bar{\xi}_n^h, u_n^h) + c'\delta y_n^h\right],\\
V^h(\hat{\xi}) = \inf_{u^h} W^h(\hat{\xi}, u^h).
\end{aligned} \tag{4.2}$$

Let $y^h(\cdot)$ denote the continuous-time interpolation of $\{\delta y_n^h\}$ with intervals $\{\Delta t_n^h\}$. By [10, Theorem 11.1.3], for any $T < \infty$

$$\limsup_h \sup_t E\left|y^h(t+T) - y^h(t)\right|^2 < \infty, \quad \sup_t E\left|y(t+T) - y(t)\right|^2 < \infty. \tag{4.3}$$

This implies that the costs are well defined. Recall the definitions of the interpolations $\psi^h(\cdot), u_\psi^h(\cdot), M^h(\cdot), m_\psi^h(\cdot)$ and of $d^h(s) = \max\{n : t_n^h \le s\}$ in Section 3. Define $\tau_n^h = \sum_{i=0}^{n-1}\Delta\tau_i^h$, $d_\tau^h(s) = \max\{n : \tau_n^h \le s\}$ and set $q_\tau^h(s) = t_{d_\tau^h(s)}^h$. For $\xi_0^h = \bar{\xi}_0^h(0)$, we have

$$\psi^h(t) = \xi_0^h + \int_0^t \int_{\mathcal{U}^h} b^h(\bar{\xi}^h(q_\tau^h(s)), \alpha) m_\psi^h(d\alpha\, ds) + M^h(t) + z_\psi^h(t), \tag{4.4}$$

where $M^h(\cdot)$ is a martingale with quadratic variation process

$$\int_0^t [\sigma^h(\bar{\xi}^h(q_\tau^h(s)))]^2 ds.$$

Modulo an asymptotically negligible error due to the "continuous time" approximation of the discount factor, the cost function (4.2) can be written as

$$W^h(\hat{\xi}, u^h) = E_{\hat{\xi}}^{h,u^h} \int_0^\infty \int_{\mathcal{U}^h} e^{-\beta t} \left[k(\bar{\xi}^h(q_\tau^h(s)), \alpha) m_\psi^h(d\alpha\, ds) + c' dy_\psi^h(s) \right]. \quad (4.5)$$

The transition probabilities. The approach of the simple example in Section 3 or, indeed, of any of the methods in [10] for obtaining the transition probabilities for the no-delay case can be readily adapted to the delay case. In all cases in [10], the transition probability for the chain in the no-delay case can be represented as a ratio; i.e., for $x \in G_h$,

$$P\{\xi_1^h = y | \xi_0^h = x, u_0^h = \alpha\} = p^h(x, y|\alpha) = N^h(x, y, \alpha)/D^h(x, \alpha),$$

and $\Delta t^h(x, \alpha) = h^2/D^h(x, \alpha)$, where $N^h(\cdot), D^h(\cdot)$ are functions of $b(\cdot), \sigma^2(\cdot)$:

$$N^h(x, y, \alpha) = \bar{N}^h(b(x, \alpha), \sigma^2(x), y), \quad D^h(x, \alpha) = \bar{D}^h(b(x, \alpha), \sigma^2(x)).$$

For the delay case, for the same ratios, simply use the forms

$$P\{\xi_1^h = y | \bar{\xi}_0^h = \hat{\xi}, u_0^h = \alpha\} = \frac{\bar{N}^h(b(\hat{\xi}, \alpha), \sigma^2(\hat{\xi}), y)}{\bar{D}^h(b(\hat{\xi}, \alpha), \sigma^2(\hat{\xi}))}. \quad (4.6)$$

Consider, in particular, the delay case analog of the approach that led to (3.3). Suppose that $\sigma^2(\hat{\xi}) \geq h|b(\hat{\xi}, \alpha)|$. Write $\hat{\xi} = \bar{\xi}_0^h, \xi_0 = \xi_0^h = \hat{\xi}(0)$. and $P_{\hat{\xi}}^{h,\alpha}\{\xi_1^h = \xi_0 \pm h\} = P\{\xi_1^h = \xi_0 \pm h | \bar{\xi}_0^h = \hat{\xi}, u_0^h = \alpha\}$. The analog for the delay case is (which defines $\bar{N}^h(\cdot), \bar{D}^h(\cdot)$)

$$\begin{gathered} P_{\hat{\xi}}^{h,\alpha}\{\xi_1^h = \xi_0 \pm h\} = \frac{\sigma^2(\hat{\xi}) \pm h b(\hat{\xi}, \alpha)}{2\sigma^2(\hat{\xi})} = \frac{\bar{N}^h(b(\hat{\xi}, \alpha), \sigma^2(\hat{\xi}), \xi_0 \pm h)}{\bar{D}^h(b(\hat{\xi}, \alpha), \sigma^2(\hat{\xi}))}, \\ \Delta t^h(\hat{\xi}, \alpha) = \frac{h^2}{\sigma^2(\hat{\xi})}. \end{gathered} \quad (4.7)$$

The cost rate becomes $k(\hat{\xi}, \alpha)$. If $\sigma(\cdot)$ is a constant, then the intervals Δt_n^h are all h^2/σ^2. The following assumption obviously holds for our special example. It is unrestrictive in general.

A4.1. The transition probabilities and interpolation intervals are given in the form (4.6).

The proof of the next two theorems for the convergence of the numerical method is in Section 8.

Theorem 4.1. *Let $\xi_n^h, \Delta t_n^h$ be locally consistent with the model* (1.1)*, whose initial condition is $\bar{x}(0)$, a continuous function. Let $\bar{\xi}_0^h \in D[G_h; -\tau, 0]$ be any piecewise constant sequence that converges to $\bar{x}(0)$ uniformly on $[-\tau, 0]$. Assume* (A1.1), (A1.3), *and* (A4.1). *Then $V^h(\bar{\xi}_0^h) \to V(\bar{x}(0))$.*

Delay in the path and control. As for the case where only the path is delayed, one constructs a chain $\xi_n^h, n \geq 0$, and interpolation intervals $\Delta t_n^h, n \geq 0$. The initial data for (1.4) is $\bar{x}(0) = \hat{x}$ and $\bar{u}(0) \in L_2[\mathcal{U}; -\tau, 0]$. The initial control data for the chain is slightly different. For the process (1.4), either $u(s), s \in [-\tau, 0]$ or $u(s), s \in [-\tau, 0)$ will do for the initial control data. But for the chain, the control u_0^h is to be determined at time 0, and should not be given as part of the initial data. This fact accounts for some of the definitions below.

Let $\bar{\xi}_0^h(\cdot), \xi^h(\cdot), \bar{\xi}^h(\cdot)$ be as above Theorem 4.1. Let $\bar{u}_0^h$ be any piecewise constant sequence in $D[\mathcal{U}^h; -\tau, 0]$ whose intervals are those of the $\bar{\xi}_0^h$, and that converges to $\bar{u}(0)$ in the L_2-sense. Let $u_\xi^h(\cdot)$ denote the function on $[-\tau, \infty)$ that equals $\bar{u}_0^h$ on $[-\tau, 0)$, and on $[0, \infty)$ equals the continuous-time interpolation of the u_n^h with intervals Δt_n^h. Let $\bar{u}_n^h(s), s \in [-\tau, 0]$, denote the segment of $u_\xi^h(\cdot)$ on $[t_n^h - \tau, t_n^h]$: $\bar{u}_n^h(\theta) = u^h(t_n^h + \theta), \theta \in [-\tau, 0]$. Let $\hat{u}_n^h$ denote the segment on the half open interval $[t_n^h - \tau, t_n^h)$.

Recall the definition of $\bar{b}(\cdot)$ in (A1.2). The distribution of ξ_{n+1}^h, given the initial data and $\xi_i^h, u_i^h, i \leq n$, depends only on $\bar{\xi}_n^h, \bar{u}_n^h$ and not on n otherwise. Thus, let $E_{\hat{\xi}}^{h,\bar{u}_0}$ denote the expectation given $\bar{\xi}_0^h = \hat{\xi}, \bar{u}_0^h = \bar{u}_0$. The local consistency condition for the chain and interpolation intervals is that there exists a function $\Delta t^h(\cdot)$ such that $\Delta t_n^h = \Delta t^h(\bar{\xi}_n^h, \bar{u}_n^h)$ and

$$\begin{aligned}
&E_{\hat{\xi}}^{h,\bar{u}_0} \Delta \xi_0^h = \bar{b}^h(\hat{\xi}, \bar{u}_0) \Delta t^h(\hat{\xi}, \bar{u}_0) = \bar{b}(\hat{\xi}, \bar{u}_0) \Delta t^h(\hat{\xi}, \bar{u}_0) + o(\Delta t^h(\hat{\xi}, \bar{u}_0)), \\
&E_{\hat{\xi}}^{h,\bar{u}_0} \beta_0^h [\beta_0^h]' = a^h(\hat{\xi}) \Delta t^h(\hat{\xi}, \bar{u}_0) = a(\hat{\xi}) \Delta t^h(\hat{\xi}, \bar{u}_0) + o(\Delta t^h(\hat{\xi}, \bar{u}_0)), \\
&a(\hat{\xi}) = \sigma(\hat{\xi}) \sigma'(\hat{\xi}), \\
&\sup_{n,\omega} |\xi_{n+1}^h - \xi_n^h| \xrightarrow{h} 0, \quad \sup_{\hat{\xi}, \bar{u}_0} \Delta t^h(\hat{\xi}, \hat{u}) \xrightarrow{h} 0.
\end{aligned} \tag{4.8}$$

The reflecting boundary is treated exactly as it was when only the path is delayed.

Now extend the definition of $u_\psi^h(\cdot)$ to the interval $[-\tau, \infty)$ by letting it equal $\bar{u}_0^h(s)$ for $s \in [-\tau, 0)$, and let $m_\psi^h(\cdot)$ denote the relaxed control representation of $u_\psi^h(\cdot)$. For $\xi_0^h = \bar{\xi}_0^h(0)$, (4.4) is replaced by

$$\psi^h(t) = \xi_0^h + \int_{-\tau}^0 \left[\int_0^t \int_{\mathcal{U}^h} b^h(\bar{\xi}^h(q_\tau^h(s)), \alpha, v) m_\psi^h(d\alpha, ds + v) \right] \mu(dv) + M^h(t) + z_\psi^h(t). \tag{4.9}$$

Let $\hat{u}$ denote the restriction of the canonical value $\bar{u}_0$ to the half open interval $[-\tau, 0)$. Let $E_{\hat{\xi}}^{h,\hat{u},u^h}$ denote the expectation under initial data $\bar{\xi}_0^h = \hat{\xi}$ and control sequence $u^h = \{u_n^h, n \geq 0\}$, with initial control segment (on $[-\tau, 0)$) $\hat{u}$. Recalling

the definition of $\bar{k}(\cdot)$ in A1.2, the cost function can be written as

$$W^h(\hat{\xi},\hat{u},u^h) = E^{h,\hat{u},u^h}_{\hat{\xi}} \sum_{n=0}^{\infty} e^{-\beta t^h_n} \left[\bar{k}(\bar{\xi}^h_n,\bar{u}^h_n)\Delta t^h(\bar{\xi}^h_n,\bar{u}^h_n) + c'\delta y^h_n\right], \qquad V^h(\hat{\xi},\hat{u}) = \inf_{\{u^h_n\}} W^h(\hat{\xi},\hat{u},u^h), \tag{4.10}$$

In integral and relaxed control form, and modulo an asymptotically negligible error due to the approximation of the discount factor, (4.10) equals

$$W^h(\hat{\xi},\hat{u},u^h) = E^{h,\hat{u},u^h}_{\hat{\xi}} \int_{-\tau}^{0} \left[\int_0^{\infty}\int_{\mathcal{U}^h} e^{-\beta t}\left[b^h(\bar{\xi}^h(q^h_\tau(t)),\alpha,v)m^h_\psi(d\alpha,dt-v) + c'dy^h_\psi(t)\right]\mu(dv)\right]. \tag{4.10}$$

In general, the transition probabilities are given by a ratio as in (4.6) and we formalize this as follows.

A4.2. The transition probabilities and interpolation intervals are given in the form (4.6), where $\bar{b}(\cdot)$ replaces $b(\cdot)$.

Theorem 4.2. *Let $\xi^h_n, \Delta t^h_n$ be locally consistent with* (1.4)*, which has the initial data $\bar{x}(0)$, a continuous function, and $\bar{u}(0) \in L_2[\mathcal{U};-\tau,0]$. Let $\bar{\xi}^h_0 \in D[G_h;-\tau,0]$ be piecewise constant, and converge to $\bar{x}(0)$ uniformly on $[-\tau,0]$. Let $\bar{u}^h_0 \in D[\mathcal{U}^h;-\tau,0]$ be piecewise constant with the same intervals as $\bar{\xi}^h_0$, have values $\bar{u}^h_0(\theta) \in \mathcal{U}^h$, and converge to $\bar{u}(0)$ in the sense of L_2. Let $\hat{u}^h$ denote the segment of $\bar{u}^h_0$ on $[-\tau,0)$. Assume* (A1.2), (A1.3) *and* (A4.2)*. Then $V^h(\bar{\xi}^h_0,\hat{u}^h) \to V(\bar{x}(0),\bar{u}(0))$.*

5 Computational Procedures

The Bellman equation. Path only delayed. Let $\overline{\Delta}^h = \inf_{\alpha,\hat{\xi}} \Delta t^h(\hat{\xi},\alpha)$, where $\alpha \in \mathcal{U}^h, \hat{\xi} \in D[G_h;-\tau,0]$, and suppose (w.l.o.g.) that $\tau/\overline{\Delta}^h = K^h$ is an integer. The interpolated time interval $[t^h_n - \tau, t^h_n]$ is covered by at most K^h intervals of length $\bar{\Delta}^h$. The $\bar{\xi}^h_n$ can be represented in terms of a finite state Markov process as follows. Recall that the reflection states do not appear in the construction of $\xi^h(\cdot)$. Let $\xi^h_{n,i}, i > 0$, denote the ith nonreflection state before time n, and $\Delta t^h_{n,i}$ the associated interpolation interval. We can represent $\bar{\xi}^h_n$ in terms of $\{(\xi^h_{n,K^h},\Delta t^h_{n,K^h}),\ldots,(\xi^h_{n,1},\Delta t^h_{n,1}),\xi^h_n\}$. This new representation is clearly a $(2K^h+1)$-dimensional controlled Markov chain. Let $\hat{\xi}$ be an arbitrary element of $D[G_h;-\tau,0]$ that is piecewise constant and $\hat{\xi}(0) = \xi_0$. Then, if there are no delays in the control the Bellman equation for the process defined by

either this chain or (4.9) with cost (4.2) can be written as

$$V^h(\hat{\xi}) = \inf_{\alpha\in\mathcal{U}^h}\left[e^{-\beta\Delta t^h(\hat{\xi},\alpha)}\sum_{\pm}P^{h,\alpha}_{\hat{\xi}}\{\xi^h_1=\xi_0\pm h\}V^h(\hat{y}^{\pm})+k(\hat{\xi},\alpha)\Delta t^h(\hat{\xi},\alpha)\right], \tag{5.1a}$$

The terms $\hat{y}^{\pm}$ denote the functions on $[-\tau,0]$ with values

$$\hat{y}^{\pm}(\theta)=\hat{\xi}(\theta+\Delta t^h(\hat{\xi},\alpha)),\ -\tau\le\theta<-\Delta t^h(\hat{\xi},\alpha),$$

$$\hat{y}^{\pm}(\theta)=\xi_0,\ -\Delta t^h(\hat{\xi},\alpha)\le\theta<0,\quad \hat{y}^{\pm}(0)=\xi_0\pm h.$$

If $\hat{\xi}(0)=\xi_0=-h$, and the other values of $\hat{\xi}(\cdot)$ are in G_h, then, $\Delta t^h(\hat{\xi},\alpha)=0$ and

$$V^h(\hat{\xi})=V^h(\hat{y}^{+})+c_1h. \tag{5.1b}$$

If $\hat{\xi}(0)=\xi_0=B+h$, and the other values of $\hat{\xi}(\cdot)$ are in G_h, then, $\Delta t^h(\hat{\xi},\alpha)=0$ and

$$V^h(\hat{\xi})=V^h(\hat{y}^{-})+c_2h. \tag{5.1c}$$

Owing to the contraction due to the discounting, there is a unique solution to (5.1).

Simplifying the state representation. Path only delayed. If the interpolation interval $\Delta t^h(\hat{\xi},\alpha)$ is not constant, then the construction of the $\bar{\xi}^h_n$ requires that we keep a record of the values of both the $\xi^h_i,\Delta t^h_i$, for the indices i that contribute to the construction. The use of constant interpolation intervals simplifies this problem. Suppose that the intervals are constant with value $\overline{\Delta}^h$). It is then apparent from the form of the Bellman equation (5.1) that the state space for the control problem for the approximating chain consists of functions $\hat{\xi}(\cdot)$ that are constant on $[-\tau+i\overline{\Delta}^h,-\tau+i\overline{\Delta}^h+\overline{\Delta}^h)$, $i\le M$, with values in G_h there and with $\hat{\xi}(0)\in G_h^+$. In addition, $\bar{\xi}^h_n$ is a piecewise constant interpolation of the K^h+1 values $\hat{\xi}^h_n\equiv(\xi^h_{n,K^h},\cdots,\xi^h_{n,1},\xi^h_n)$ and we can identify $\bar{\xi}^h_n$ with this vector without ambiguity. If $\xi^h_n\in G_h$, then $\hat{\xi}^h_{n+1}=(\xi^h_{n,K^h-1},\ldots,\xi^h_{n,1},\xi^h_n,\xi^h_{n+1})$. If $\xi^h_n=-h$, then $\hat{\xi}^h_{n+1}=(\xi^h_{n,K^h},\cdots,\xi^h_{n,1},0)$, and analogously if $\xi^h_n=B+h$. Thus the full state vector is (K^h+1)-dimensional and the maximum number of possible values can be very large, up to $(B/h+3)(B/h+1)^{K^h}$.

The analog of the procedure (3.5) for getting an approximating chain with a constant interpolation interval is obvious. Let $\bar{P}$ denote the transition probabilities for the constant interpolation interval case. For the delay case with $\hat{\xi}(0)=\xi^h_0=\xi_0\in G_h$, and $\xi^h_1=\xi_0+\{\pm h,0\}$, use

$$\begin{aligned}\bar{P}^{h,\alpha}_{\hat{\xi}}\{\xi^h_1=\xi_0\pm h\}&=P^{h,\alpha}_{\hat{\xi}}\{\xi^h_1=\xi_0\pm h\}(1-\bar{P}^{h,\alpha}_{\hat{\xi}}\{\xi^h_1=\xi_0\}),\\ \bar{P}^{h,\alpha}_{\hat{\xi}}\{\xi^h_1=\xi_0\}&=1-\frac{\overline{\Delta}^h}{\Delta t^h(\hat{\xi},\alpha)}.\end{aligned} \tag{5.2}$$

Simplification of the state space. We only need to keep track of ξ_n^h and the differences between successive components of $\hat{\xi}_n^h$. This gives the representation

$$\hat{\xi}_n^h = (c_{n,K^h}^h, \cdots, c_{n,1}^h, \xi_n^h), \quad \text{where, for } 1 < i \le M_h,\ c_{n,i}^h = \xi_{n,i}^h - \xi_{n,i-1}^h \quad (5.3)$$

and $c_{n,1}^h = \xi_n^h - \xi_{n,1}^h$. Suppose now that $\sigma^2(\cdot)$ is constant and non-zero and that (4.7) is used, so that $\Delta t^h(\hat{\xi}, \alpha) = h^2/\sigma^2$ is constant. Then the $c_{n,i}^h$ take at most two values and the number of values in the state space is reduced to $(B/h + 3)2^{K^h}$. The two values and the reconstruction of the $\xi_{n,i}^h$ from them are easily determined by an iterative procedure. For example, if $\xi_n^h = -h$, then $\xi_{n,1}^h = 0$. If $\xi_n^h = 0$ then $\xi_{n,1}^h \in \{0, h\}$. If ξ_n^h is not a reflecting or boundary value then $\xi_{n,1}^h = \xi_{n-1}^h = \xi_n^h \pm h$. If $\xi_{n,i}^h = 0$, then $\xi_{n,i-1}^h \in \{0, h\}$. If $\xi_{n,i}^h$ is not a boundary value then $\xi_{n,i-1}^h \in \{-h, h\}$, and so forth.

If $\sigma^2(\cdot)$ is not constant, then use the form (5.2) to get a chain with a constant interval. Now $\xi_{n+1}^h - \xi_n^h \in \{-h, 0, h\}$, each of the $c_{n,i}^h$ can take as much as three values and we have at most $(B/h+3)3^{K^h}$ values in the state space. The approach in the next section uses fewer intervals to cover $[-\tau, 0]$ and has the promise of being more efficient in terms of memory requirements.

Delay in path and control. Now suppose that both the control and the path are delayed. The memory state at time n for the discrete-time dynamic program is $\bar{\xi}_n^h, \hat{u}_n^h$, the value of $\xi^h(\cdot)$ on the closed interval $[t_n^h - \tau, t_n^h]$, together with the path of the interpolated control (intervals Δt_n^h) on the half open interval $[t_n^h - \tau, t_n^h)$. The canonical value of $\hat{u}_0^h$ is $\hat{u}$. The transition probabilities at time n depend on the memory variables and the new control value u_n^h. Thus, write $(\hat{u}, \alpha)$ for the canonical value of the control on $[-\tau, 0]$, where α denotes the value at time 0. We can then use terms such as $\bar{b}(\hat{\xi}, \hat{u}, \alpha)$ without ambiguity. The form (4.6) still applies, with $\bar{b}(\hat{\xi}, \hat{u}, \alpha)$ used in lieu of $b(\hat{\xi}, \alpha)$.

Analogously to what was done at the beginning of the section for the case where the control is not delayed, the memory variables can be imbedded into a Markov process, with values at time n

$$\left\{(\xi_{n,K^h}^h, u_{n,K^h}^h, \Delta t_{n,K^h}^h), \ldots, (\xi_{n,1}^h, u_{n,1}^h, \Delta t_{n,1}^h), \xi_n^h\right\}.$$

Then for the transition probabilities that are the analogs of those in (4.7) for the present case, the analog of (5.1a) is

$$V^h(\hat{\xi}, \hat{u}) = \inf_{\alpha \in \mathcal{U}^h} \left[e^{-\beta \Delta t^h(\hat{\xi}, \hat{u}, \alpha)} \sum_{\pm} P_{\hat{\xi}, \hat{u}}^{h,\alpha}\{\xi_1^h = \xi_0 \pm h\} V^h(\hat{y}^{\pm}, \hat{u}_\alpha) + k(\hat{\xi}, \hat{u}, \alpha)\Delta t^h(\hat{\xi}, \hat{u}, \alpha) \right], \quad (5.4)$$

where the following definitions are used. $P_{\hat{\xi}, \hat{u}}^{h,\alpha}$ denotes the transition probability with memory variables $\hat{\xi}, \hat{u}$ and new control value α used. The $\hat{y}^{\pm}$ denotes the

new "path memory sections" defined below (5.1a), with $\Delta t^h(\hat{\xi}, \hat{u}, \alpha)$ used in lieu of $\Delta t^h(\hat{\xi}, \alpha)$. The new "control memory section" $\hat{u}_\alpha$ is defined by

$$\hat{u}_\alpha(\theta) = \hat{u}(\theta + \Delta t^h(\hat{\xi}, \hat{u}, \alpha)),\ -\tau \le \theta < -\Delta t^h(\hat{\xi}, \hat{u}, \alpha),$$
$$\hat{u}_\alpha(\theta) = \alpha,\ -\Delta t^h(\hat{\xi}, \hat{u}, \alpha) \le \theta < 0.$$

The reflecting states are treated as for the no-delay case. Because of the contraction due to the discounting there is a unique solution to (5.4).

We can use the more efficient representation (5.3) for the path variable. But the total memory requirements might still be too large, unless $\mathcal{U}$ itself can be effectively "approximated" by only a few values.

A comment on higher-dimensional problems. We have concentrated on one-dimensional models. But the ideas concerning approximation and the convergence results all extend to quite general higher-dimensional problems. The solution to the reflected diffusion is confined to a compact region G by a boundary reflection. The conditions on G and the reflection directions are exactly as in [10] and there is no need to say more. The Markov chain approximations for the delay problem in higher dimensions adapts the methods of the reference in the same way as was done for the one-dimensional problem considered in this paper, using the comments of this paper as a guide to the substitutions. For the no-delay problem, the required memory grows rapidly as the dimension increases, and that also holds here. Two-dimensional problems are feasible at present.

Representations analogous to (5.3) can also be used for the higher-dimensional problem. Consider a two-dimensional problem in a box $[0, B_1] \times [0, B_2]$, with the same path delay in each coordinate, no control delay, and discretization level h in each coordinate. The ξ_n^h in (5.3) is replaced by vector containing the current two-dimensional value of the chain. The difference $c_i = \xi_{n,i}^h - \xi_{n,i-1}^h$ is now a two dimensional vector. The values can be computed iteratively, as for the one-dimensional case, but the details will not be presented here.

6 The Implicit Approximation: Path Only Delayed

Let $\delta > 0$ with $h^2/\delta \to 0$ as $h \to 0, \delta \to 0$, and suppose that $\tau/\delta = Q$ is an integer. The process $\zeta_n^{h,\delta} = (\phi_n^{h,\delta}, \xi_n^{h,\delta}) =$ (temporal variable, spatial variable) whose transition probabilities were defined by (3.11) leads to some intriguing possibilities for efficient representation of the memory data for the delay problem. Recall that either the spatial variable $\xi_n^{h,\delta}$ changed or the time variable $\phi_n^{h,\delta}$ advanced at each iteration, but not both. We will construct the analog of $\zeta_n^{h,\delta}$ for the delay case. There are several choices for the time scale of the continuous-time interpolations. One can use the analog of the $\Delta t^{h,\delta}$ defined in (3.11), and proceed as in the last section. Another possibility, which we will pursue, is to let the value of $\phi_n^{h,\delta}$ determine the interpolation. More precisely,

at the nth step the interpolated time would be $\phi_n^{h,\delta}$ and not t_n^h. Then the time variable for the interpolation does not necessarily advance at each step of the chain. It will be seen that both interpolations are the same asymptotically.

Suppose that $\bar{\xi}_n^{h,\delta}$ denotes the part of the path that represents the memory state at iterate n for the chain. It will be defined precisely after writing the transition probabilities. Now, adapting the procedure that led to (3.11) to the delay case yields the transition probabilities and interpolation intervals for the $\zeta_n^{h,\delta} = (\phi_n^h, \xi_n^{h,\delta})$ process in terms of those for the ξ_n^h process as:

$$\begin{aligned}
&P\left\{\xi_1^{h,\delta} = \xi_0 \pm h \middle| \bar{\xi}_0^{h,\delta} = \hat{\xi}, \phi_0^{h,\delta} = \phi_0, \xi_0^{h,\delta} = \xi_0, u_0^{h,\delta} = \alpha\right\} \\
&\qquad = P_{\hat{\xi},\phi_0}^{h,\alpha}\left\{\xi_1^{h,\delta} = \xi_0 \pm h\right\}\left(1 - P_{\hat{\xi},\phi_0}^{h,\delta,\alpha}\left\{\phi_1^{h,\delta} = \phi_0 + \delta\right\}\right), \\
&P_{\hat{\xi},\phi_0}^{h,\delta,\alpha}\left\{\phi_1^{h,\delta} = \phi_0 + \delta\right\} = \\
&\quad P\left\{\phi_1^{h,\delta} = \phi_0 + \delta \middle| \bar{\xi}_0^{h,\delta} = \hat{\xi}, \phi_0^{h,\delta} = \phi_0, u_0^{h,\delta} = \alpha\right\} = \frac{\Delta t^h(\hat{\xi},\alpha)}{\Delta t^h(\hat{\xi},\alpha) + \delta},
\end{aligned} \tag{6.1}$$

$$\Delta t^{h,\delta}(\hat{\xi},\alpha) = \frac{\delta \Delta t^h(\hat{\xi},\alpha)}{\Delta t^h(\hat{\xi},\alpha) + \delta}. \tag{6.2}$$

Define

$$\Delta t_n^{h,\delta} = \Delta t^{h,\delta}(\bar{\xi}_n^h, u_n^h), \quad t_n^{h,\delta} = \sum_{i=0}^{n-1} \Delta t_i^{h,\delta}.$$

Interpolations. One could base the continuous-time interpolation used to get $\bar{\xi}_n^{h,\delta}$ on the intervals $\Delta t_n^{h,\delta}$. But then the issues concerning the number of required values of the memory variable would be similar to those of the last section. Consider the alternative where the time variables $\phi_n^{h,\delta}$ determine interpolated time, in that real (i.e., interpolated) time advances (by an amount δ) only when the time variable is incremented and it does not advance otherwise. To make this precise, consider $\xi_n^{h,\delta}$ at only the times that $\phi_n^{h,\delta}$ changes. Define $\mu_0^{h,\delta} = 0$, and, for $n > 0$, set

$$\mu_n^{h,\delta} = \inf\{i > \mu_{n-1}^{h,\delta} : \phi_i^{h,\delta} - \phi_{i-1}^{h,\delta} = \delta\}.$$

Define the "memory" path segment $\bar{\xi}_n^{h,\delta}(\theta), \theta \in [-\tau, 0]$, as follows. For any l and n satisfying $\mu_l^{h,\delta} \le n < \mu_{l+1}^{h,\delta}$, set

$$\begin{aligned}
&\bar{\xi}_n^{h,\delta}(0) = \xi_n^{h,\delta}, \\
&\bar{\xi}_n^{h,\delta}(\theta) = \xi_{\mu_l^{h,\delta}}^{h,\delta}, \qquad \theta \in [-\delta, 0), \\
&\vdots \\
&\bar{\xi}_n^{h,\delta}(\theta) = \xi_{\mu_{l-Q+1}^{h,\delta}}^{h,\delta}, \quad \theta \in [-\tau, -\tau + \delta).
\end{aligned} \tag{6.3}$$

Let $\hat{\xi}$ denote the canonical value of $\bar{\xi}_0^{h,\delta}$. It can be represented as the piecewise constant right continuous interpolation with interval δ of its values

$$(\hat{\xi}(-\tau), \ldots, \hat{\xi}(-\delta), \hat{\xi}(0))$$

with a discontinuity at $t = 0$, as usual. The possible transitions are as follows. Let $\hat{\xi}(0) \in G_h$. Then $\hat{\xi}$ transits as either (if the time variable does not advance)

$$\hat{\xi} = (\hat{\xi}(-\tau), \ldots, \hat{\xi}(-\delta), \hat{\xi}(0)) \to (\hat{\xi}(-\tau), \ldots, \hat{\xi}(-\delta), \hat{\xi}(0) \pm h)$$

or (if the time variable advances)

$$(\hat{\xi}(-\tau), \ldots, \hat{\xi}(-\delta), \hat{\xi}(0)) \to (\hat{\xi}(-\tau + \delta), \ldots, \hat{\xi}(-\delta), \hat{\xi}(0), \hat{\xi}(0)).$$

For the reflecting points, we have the immediate transition

$$(\hat{\xi}(-\tau), \ldots, \hat{\xi}(-\delta), -h) \to (\hat{\xi}(-\tau), \ldots, \hat{\xi}(-\delta), 0),$$

$$(\hat{\xi}(-\tau), \ldots, \hat{\xi}(-\delta), B + h) \to (\hat{\xi}(-\tau), \ldots, \hat{\xi}(-\delta), B).$$

Thus $\bar{\xi}_n^{h,\delta}(\theta) \in G_h$ for $\theta \neq 0$.

Local properties and the dynamical equations. Define $\delta z_n^{h,\delta}, \beta_n^{h,\delta}$ analogously to the definitions used in Sections 3 and 4. By (6.1) and (6.2),

$$E\left[\phi_{n+1}^{h,\delta} - \phi_n^{h,\delta} \middle| \zeta_i^{h,\delta}, u_i^{h,\delta}, i \leq n, \xi_n^{h,\delta} \in G_h\right] = \Delta t_n^{h,\delta}.$$

Define the martingale difference $\beta_{0,n}^{h,\delta} = \left(\phi_{n+1}^{h,\delta} - \phi_n^{h,\delta}\right) - \Delta t_n^{h,\delta}$. With the definitions (6.1), (6.2), and $\hat{\xi} = \bar{\xi}_0^{h,\delta}, \hat{\xi}(0) \in G_h$, we have $E_{\hat{\xi},\phi_0}^{h,\delta,\alpha} \Delta \xi_0^{h,\delta} = b^h(\hat{\xi}, \alpha) \Delta t^{h,\delta}(\hat{\xi}, \alpha)$, and the conditional covariance of the martingale difference term $\beta_0^{h,\delta}$ is

$$\sigma^h(\hat{\xi}) \sigma^h(\hat{\xi}) \Delta t^{h,\delta}(\hat{\xi}, \alpha).$$

Analogously to the expression below (3.6), define the stopping time

$$d^{h,\delta}(t) = \max\left\{ n : \sum_{i=0}^{n-1} \Delta t_i^{h,\delta} = t_n^{h,\delta} \leq t \right\}.$$

As in Theorem 4.1, approximate the initial condition $\bar{x}(0)$ by $\bar{\xi}_0^{h,\delta}$ (in the sense of uniform convergence as $h \to 0, \delta \to 0$), and let it be constant on the intervals $[-\tau + k\delta, -\tau + k\delta + \delta), k = 0, \ldots, Q - 1$, with the values at $-k\delta, k = 0, \ldots, Q$, being in G_h. Let $\xi^{h,\delta}(\cdot)$ and $\phi^{h,\delta}(\cdot)$ denote the continuous-time interpolations of the $\{\xi_n^{h,\delta}\}$ and $\{\phi_n^{h,\delta}\}$ with the intervals $\{\Delta t_n^{h,\delta}\}$. With $\xi_0^{h,\delta} = \bar{\xi}_0^{h,\delta}(0)$, we can write

$$\xi^{h,\delta}(t) = \xi_0^{h,\delta} + \sum_{i=0}^{d^{h,\delta}(t)-1} b^h(\bar{\xi}_i^{h,\delta}, u_i^{h,\delta}) \Delta t_i^{h,\delta} + \sum_{i=0}^{d^{h,\delta}(t)-1} \beta_i^{h,\delta} + \sum_{i=0}^{d^{h,\delta}(t)-1} \delta z_i^{h,\delta}, \quad (6.4)$$

$$\phi^{h,\delta}_{n+1} = \phi^{h,\delta}_n + \Delta t^{h,\delta}_n + \beta^{h,\delta}_{0,n}. \tag{6.5}$$

Next let us get the analog of the process (4.4). Define the random variables $\nu^{h,\delta}_n, n = 0, 1 \ldots$, i.i.d., exponentially distributed, and independent of the $\{\zeta^{h,\delta}_n, u^h_n\}$. Then set $\Delta\tau^{h,\delta}_n = \nu^{h,\delta}_n \Delta t^{h,\delta}_n$, $\tau^{h,\delta}_n = \sum_{i=0}^{n-1} \Delta\tau^{h,\delta}_n$. Define $d^{h,\delta}_\tau(s) = \max\{n : \tau^{h,\delta}_n \le s\}$. Define $\bar\xi^{h,\delta}(s) = \bar\xi^{h,\delta}_{d^{h,\delta}_\tau(s)}$ and $q^{h,\delta}_\tau(s) = t^{h,\delta}_{d^{h,\delta}_\tau(s)}$. With these definitions, $\bar\xi^{h,\delta}_{d^{h,\delta}_\tau(s)} = \bar\xi^{h,\delta}(q^{h,\delta}_\tau(s))$. Let $\psi^{h,\delta}(\cdot)$ denote the interpolation with the random intervals $\Delta\tau^{h,\delta}_n$. Then, analogously to (4.4),

$$\psi^{h,\delta}(t) = \xi^{h,\delta}_0 + \int_0^t \int_{\mathcal{U}^h} b^h(\bar\xi^{h,\delta}_{d^{h,\delta}_\tau(s)}, \alpha) m^{h,\delta}_\psi(d\alpha ds) + M^{h,\delta}(t) + z^{h,\delta}_\psi(t), \tag{6.6}$$

where the quadratic variation of the martingale $M^{h,\delta}(\cdot)$ is

$$\int_0^t \sigma^h(\bar\xi^{h,\delta}(q^{h,\delta}_\tau(s)))\sigma^h(\bar\xi^{h,\delta}(q^{h,\delta}_\tau(s)))ds.$$

It follows from the proof of Theorem 3.1 that the time scales used in the $\xi^{h,\delta}(\cdot)$ and the $\psi^{h,\delta}(\cdot)$ processes coincide asymptotically. I.e., $q^{h,\delta}(s) - s \to 0$, $\tau^{h,\delta}(s) - t^{h,\delta}(s) \to 0$, $q^{h,\delta}_\tau(s) - s \to 0$. The following theorem asserts this result and the fact that $\phi^{h,\delta}(t)$ converges to t.

Theorem 6.1. *Let $\phi^{h,\delta}(\cdot)$ denote the interpolation of the $\phi^{h,\delta}_n$ with the intervals $\Delta t^{h,\delta}_n$. Then $\phi^{h,\delta}(\cdot)$ converges weakly to the process with value t at time t. Also,*

$$\lim_{h\to 0,\delta\to 0} \sup_{u^{h,\delta}} E \sup_{s\le t} \left[\sum_{i=0}^{d^{h,\delta}(s)} (\Delta\tau^{h,\delta}_i - \Delta t^{h,\delta}_i)\right]^2 = 0.$$

The last assertion holds with $d^{h,\delta}_\tau(\cdot)$ replacing $d^{h,\delta}(\cdot)$.

Proof. By (6.5),

$$\phi^{h,\delta}(t) = \sum_{i=0}^{d^{h,\delta}(t)-1} \Delta t^{h,\delta}_i + \sum_{i=0}^{d^{h,\delta}(t)-1} \beta^{h,\delta}_{0,i}.$$

The first sum equals t, modulo $\sup_n \Delta t^{h,\delta}_n$. The variance of the martingale term is δt, modulo $\delta + \sup_n \Delta t^{h,\delta}_n$, and the term converges weakly to the zero process. The proof of the second assertion of the theorem is just that of Theorem 3.1. The last assertion of the theorem follows from the second. ■

It follows from Theorem 6.1 that

$$\sup_{-\tau\le\theta\le 0, t\le T} \left|\xi^{h,\delta}(t+\theta) - \bar\xi^{h,\delta}_{d^{h,\delta}_\tau(t)}(\theta)\right| \to 0. \tag{6.7}$$

The cost function. Consider the cost function

$$W^{h,\delta}(\hat{\xi},u^h)=E^{h,\delta,u^h}_{\hat{\xi},\phi_0}\sum_{i=0}^{\infty}e^{-\beta\phi^{h,\delta}_n}\left[k(\bar{\xi}^{h,\delta}_n,u^h_n)\delta I_{\{\phi^{h,\delta}_{n+1}\neq\phi^{h,\delta}_n\}}+c'\delta y^{h,\delta}_n\right]. \quad (6.8)$$

By using (6.1) and (6.2) and a conditional expectation, the term $\delta I_{\{\phi^{h,\delta}_{n+1}\neq\phi^{h,\delta}_n\}}$ can be replaced by $\Delta t^{h,\delta}_n$. We next show that (6.8) is well defined.

Theorem 6.2. *For small h,δ,* (6.8) *is asymptotically equal to* (*uniformly in such h,δ*)

$$E^{h,\delta,u^h}_{\hat{\xi},\phi_0}\sum_{i=0}^{\infty}e^{-\beta t^{h,\delta}_n}\left[k(\bar{\xi}^{h,\delta}_n,u^{h,\delta}_n)\Delta t^{h,\delta}_n+c'\delta y^{h,\delta}_n\right]. \quad (6.9)$$

Proof. To show that the term involving $k(\cdot)$ in (6.8) is well defined first note that it can be bounded by a constant times the expectation of $\int_0^\infty e^{-\beta\phi^{h,\delta}(s)}ds$. By the computations in Theorem 6.1, for each $K>0$ there are $\epsilon_i>0$, that do not depend on the controls, such that for small enough h,δ

$$P\left\{\phi^{h,\delta}(T+K)-\phi^{h,\delta}(T)\geq\epsilon_1\middle|\text{data to } T\right\}>\epsilon_2$$

w.p.1 for each T. Hence, there is $\epsilon_3>0$, not depending on the controls, such that for small enough h,δ

$$E\left[e^{-\beta(\phi^{h,\delta}(T+K)-\phi^{h,\delta}(T))}\middle|\text{data to } T\right\}\leq e^{-\epsilon_3}$$

w.p.1 for each T. This implies that the "tail" of the sum (6.8) can be neglected and we need only consider the sum $\sum_{i=0}^{N^{h,\delta}(t)}$ where $N^{h,\delta}(t)=\min\{n:t^{h,\delta}_n\geq t\}$ for arbitrary t. But, by Theorem 6.1, for such a sum the asymptotic values are the same if $\phi^{h,\delta}_i, i\leq N^{h,\delta}(t)$, is replaced by $t^{h,\delta}_i, i\leq N^{h,\delta}(t)$. Hence the terms involving $k(\cdot)$ in (6.8) and (6.9) are asymptotically equal. The above estimates and the inequality (4.3) yield the same result for the terms involving $\delta y^{h,\delta}_n$. ■

The Bellman equation. With the form (6.9), the effective canonical cost rate is just $k(\hat{\xi},\alpha)$ times δ times the probability that the time variable advances, namely $k(\hat{\xi},\alpha)\Delta t^{h,\delta}(\hat{\xi},\alpha)$. This can be seen from (6.9), or from (6.8) with the replacement noted below it.

The Bellman equation can be based on either (6.8) or (6.9). They will yield different results, but will be asymptotically equal by Theorem 6.2. For (6.8) and $\hat{\xi}(0)=\xi^{h,\delta}_0\in G_h$, the Bellman equation is (the time variable ϕ does not

appear in the state since the dynamical terms are time independent)

$$V^{h,\delta}(\hat{\xi}) = \inf_{\alpha\in\mathcal{U}^h}\Bigg[\sum_{\pm} P^{h,\delta,\alpha}_{\hat{\xi},\phi_0}\Big\{\xi_1^{h,\delta} = \hat{\xi}(0)\pm h\Big\}V^{h,\delta}\Big(\hat{\xi}(-\tau),\ldots,\hat{\xi}(-\delta),\hat{\xi}(0)\pm h\Big)$$
$$+e^{-\beta\delta}P^{h,\delta,\alpha}_{\hat{\xi},\phi_0}\Big\{\phi_1^{h,\delta} = \phi_0+\delta\Big\}V^{h,\delta}\Big(\hat{\xi}(-\tau+\delta),\ldots,\hat{\xi}(-\delta),\hat{\xi}(0),\hat{\xi}(0)\Big)$$
$$+k(\hat{\xi},\alpha)\Delta t^{h,\delta}(\hat{\xi},\alpha)\Bigg]. \tag{6.10}$$

If $\hat{\xi}(0)$ is a reflecting point $-h$ or $B+h$, then

$$V^{h,\delta}(\hat{\xi}) = V^{h,\delta}\Big(\hat{\xi}(-\tau),\ldots,\hat{\xi}(-\delta),0\Big) + c_1 h \text{ for } \hat{\xi}(0) = -h,$$

$$V^{h,\delta}(\hat{\xi}) = V^{h,\delta}\Big(\hat{\xi}(-\tau),\ldots,\hat{\xi}(-\delta),B\Big) + c_2 h \text{ for } \hat{\xi}(0) = B+h.$$

These equations make it clear that the full state at iterate n is $\bar{\xi}_n^{h,\delta}$, namely, the current value of the spatial variable $\xi_n^{h,\delta}$, together with its value at the last $Q = \tau/\delta$ times that the time variable advances.

Since we can use (6.9) for the cost function when proving convergence, the proof of the next theorem is nearly identical to that of Theorem 4.1 which is given in Section 8.

Theorem 6.3. *Let $\xi_n^{h,\delta}, \Delta t_n^h$ be locally consistent with the model* (1.1)*, whose initial condition is $\hat{x} = \bar{x}(0)$, a continuous function. Let $\bar{\xi}_0^{h,\delta}$ approximate $\hat{x}$ as in Theorem* 4.1*. Assume* (A1.1), (A1.3)*, and the analog of* (A4.1) *for the implicit procedure. With either* (6.10) *or the Bellman equation for* (6.9) *used, $V^h(\bar{\xi}_0^{h,\delta}) \to V(\hat{x})$.*

7 The Number of Points for the Implicit Method-State Delay Only

Comment on the value of δ. Consider solving a parabolic PDE on a finite time interval, and with the classical estimates of rate of convergence holding. Typically $\delta = O(h)$ and the rate of convergence is $O(h^2)+O(\delta^2)$ for the implicit procedure, vs. $O(h^2) + O(\text{max time increment})$ for the explicit procedure [14, Chapter 6]. But, for the explicit procedure the value of the time increment is $O(h^2)$. Thus, for $\delta = O(h)$, the rates of convergence would be of the same order. There is no proof that such estimates hold for the control problem of concern here. But numerical data for the no-delay problems suggests that one should use $\delta = O(h)$.

The implicit procedure actually updates the path values using the time increment $\Delta t_n^{h,\delta}$ (see (6.1) and (6.2)), which is close to Δt_n^h when $\delta = O(h)$ and $\Delta t_n^h = O(h^2)$. After some random number of updates of the path component, the time component advances. Thus the interpolation defined by (6.3) is essentially a sampling of the process $\xi_n^{h,\delta}$ at random intervals. Since the average sum of the $\Delta t_n^{h,\delta}$ between advances of time is approximately δ, the sampling is approximately at time intervals δ. This would give a more accurate construction than would a process ξ_n^δ constructed with the much larger discretization interval δ. Additionally, the implicit procedure allows us to use the original time intervals $\Delta t_n^{h,\delta} \approx \Delta t_n^h$, and not the minimal value $\overline{\Delta}^h$. This is computationally advantageous when the values $\Delta t^h(\hat{\xi}, \alpha)$ vary a great deal, for example when the upper bound on the control is large. It will next be argued that when $\delta = O(h)$, the implicit procedure can be approximated such that it has a much smaller memory requirement than the explicit procedure.

Reduced memory requirements. The vector $\hat{\xi}$ can be represented in terms of the vector

$$\hat{d} = \Big(\hat{\xi}(-\tau) - \hat{\xi}(-\tau+\delta), \ldots, \hat{\xi}(-\delta) - \hat{\xi}(0), \hat{\xi}(0)\Big) = (d(Q), \ldots, d(0)).$$

If $\hat{\xi}(0)$ is a reflection point, then it moves immediately to the closest point in G_h. Otherwise, with this representation, the transitions are to (if the time variable does not advance)

$$(d(Q), \ldots, d(2), d(1) \mp h, d(0) \pm h)$$

or to (if the time variable advances)

$$(d(Q-1), \ldots, d(2), 0, d(0)).$$

The variable $d(0)$ takes $B/h + 3$ possible values. Since there are a potentially unbounded number of steps before the time variable increases, the differences $d(i), i \geq 2$, can take values in the set $G_h - G_h$, which is the set of points $\{B, B-h, \ldots, -B\}$. Hence there are $2B/h + 1$ possible values. The $d(1)$ can take values in $G_h - G_h^+ = \{B+h, B, B-h, \ldots, -B-h\}$, since $\hat{\xi}(0)$ takes values in G_h^+. But over the number of steps that are required for the time variable to advance, with a high probability the sample number of values taken by the $d(i)$ will be much less due to cancellations of positive and negative steps. This idea can be exploited by truncating the possible values of the $d(i), i \geq 1$, by some $N_1 < 2B/h+1$ such that the probability that $d(i)$ takes more than N_1 values is smaller than some predetermined number.[8] The maximum required number of points is $(B/h+3)(2B/h+3)(2B/h+1)^{\tau/\delta-1}$. Comparing this with the number

$$(B/h+3)2^{\tau/\overline{\Delta}^h}, \quad \text{or } (B/h+3)3^{\tau/\overline{\Delta}^h},$$

[8] More generally, one can approximate the range of the $d(i), i \geq 1$, by allowing them to take some prespecified values that might not be integral multiples of h.

for the explicit procedure, we see that there might be considerable saving even without truncation, since typically $\overline{\Delta}^h = O(h^2)$ and $\delta = O(h)$. This saving is due to the fact that, for the implicit procedure the memory consists of the differences in the values attained over many steps, and not set of differences in the values for each of those steps.

We next present some computations concerning the value of N_1.

An example. For illustrative purposes, first consider the simplest example. Let $\sigma(\cdot)$ be constant and $b(\cdot)$ zero. There is no delay, but it will be seen that the estimates are typical even with a drift and delay. Then ξ_n^h is a simple symmetric random walk with reflection and $\Delta t^h = h^2/\sigma^2$. Then use $\delta = h/\sigma$, which is very close to the probability that the time variable advances at each step if h is small. Let v be the number of steps required for the time variable to advance, and define $M_n^{h,\delta} = \sum_{i=0}^n \beta_i^{h,\delta}$. Then $Ev^2 \approx \sigma^2/h^2$ and

$$P\left\{\sup_{n\le v}\left|M_n^{h,\delta}\right| \ge \epsilon\right\} \le \frac{E|M_v^{h,\delta}|^4}{\epsilon^4} \approx \frac{h^4 Ev^2}{\epsilon^4}. \tag{7.1}$$

Let $N_1 = 2N+1, \epsilon = hN$. Then the probability above is bounded by σ^2/N^4h^2. Thus we need N to increase slightly faster than $1/\sqrt{h}$ to have an asymptotically negligible error. The number of memory points needs is approximately $(B/h+3)[2/f(h)+1]^{\tau/\delta} = (B/h+3)[2/f(h)+1]^{\sigma\tau/h}$, where $f(h)/\sqrt{h} \to 0$, as compared to the much larger number $(B/h+3)2^{\sigma^2\tau/h^2}$ for the explicit procedure. The best way of getting N or of efficiently approximating the range of the $d(i), i \ge 1$, is not clear and much further work and numerical experience is required. But the idea is very appealing.

Now, extend the above case by letting $b(\cdot)$ be non-zero, with delayed arguments and satisfying $|b(\hat{\xi},\alpha)|h \le \sigma^2/2$. We still have $\Delta t^h = h^2/\sigma^2$. The number of points needed for the explicit method was noted above. Write $b_n^h = b(\bar{\xi}_n^h, u_n^h)$ and $b_0 = \sup|b(\hat{\xi},\alpha)|$. Now estimate $\sup_{n\le v}\sum_{i=0}^n [b_n^h \Delta t^{h,\delta} + \beta_i^{h,\delta}]$ by splitting the terms. For the drift term, we have the estimate

$$P\{\sup_{n\le v} b_0 \sum_{i=0}^n \Delta t_i^{h,\delta} \ge Nh/2\} \le 4\frac{b_0^2[h^4/\sigma^4]Ev^2}{N^2h^2} \approx \frac{4b_0^2}{N^2\sigma^2}.$$

The estimate (7.1) continues to hold for the martingale term. Thus the martingale term dominates and conclusions of the simpler case continue to hold.

8 Comments on the Proof of Theorems 4.1 and 4.2

Proof. For notational simplicity, let us start with the case where the control is not delayed. The proof is close to that for the no-delay case and the structure

will be outlined. Let $u^h = \{u_n^h, n < \infty\}$ be the optimal control sequence for the chain.

The main new issue over the no-delay case is that the process $\bar{\xi}^h(t)$ appears in the dynamical terms $b(\cdot), \sigma(\cdot), k(\cdot)$. Recall the definitions, for $s \geq 0$, $q_\tau^h(s) = \max\{n : \tau_n^h \leq s\}$ and $q_\tau^h(s) = t_{q_\tau^h(s)}^h$. Recall (4.4):

$$\psi^h(t) = \xi_0^h + \int_0^t \int_{\mathcal{U}^h} b^h(\bar{\xi}^h(q_\tau^h(s)), \alpha) m_\psi^h(d\alpha\, ds) + M^h(t) + z_\psi^h(t), \tag{8.1}$$

where $\psi^h(\cdot)$ is the continuous-time interpolation with intervals $\nu_n^h \Delta t_n^h$, $M^h(\cdot)$ is a martingale with quadratic variation process

$$\int_0^t \left[\sigma^h(\bar{\xi}^h(q_\tau^h(s)))\right]^2 ds,$$

and can be written as [10, Section 10.4.1]

$$M^h(t) = \int_0^t \sigma^h(\bar{\xi}^h(q_\tau^h(s)))dw^h(s), \tag{8.2}$$

where $w^h(\cdot)$ is a martingale with quadratic variation process It. The discontinuities of $w^h(t)$ go to zero as $h \to 0$, and it converges to a standard Wiener process. The proofs of these assertions are the same as for the no-delay case in [10, Section 10.4]. Keep in mind that in all cases, $\xi^h(\cdot)$ and $\psi^h(\cdot)$ are constructed from the basic ξ_n^h, via the appropriate interpolation. Theorem 3.1 applies and shows that the time scales for $\xi^h(\cdot)$ and $\psi^h(\cdot)$ are asymptotically equal.

The sequence $\{\psi^h(\cdot), \xi^h(\cdot), m_\psi^h(\cdot), w^h(\cdot), y_\psi^h(\cdot), q_\tau^h(\cdot)\}$, where $z_\psi^h(\cdot) = y_{1,\psi}^h(\cdot) - y_{2,\psi}^h(\cdot)$, is tight and all weak sense limits are continuous. Take a weakly convergent subsequence, also indexed by h and with limit denoted by $(x(\cdot), \xi(\cdot), m(\cdot), w(\cdot), y(\cdot), q(\cdot))$, with $z(\cdot) = y_1(\cdot) - y_2(\cdot)$. The asymptotic continuity of $q_\tau^h(\cdot)$ is implied by Theorem 6.1, and that of $y_{i,\psi}^h(\cdot)$ is similar to the proof of a similar result in Theorem 2.1. See also the reference (Chapter 10). Assume the Skorohod representation so that the limits can be assumed to be w.p.1. By Theorem 6.1, $x(\cdot) = \xi(\cdot)$ and $q(t) = t$. Thus the process $\bar{\xi}^h(\cdot)$ converges to $\bar{x}(\cdot)$. Now, by the continuity conditions on $b(\hat{x}, \alpha)$ and $\sigma(\hat{x})$ in (A1.1),

$$\int_0^t \int_{\mathcal{U}^h} b^h(\bar{\xi}^h(q_\tau^h(s)), \alpha) m_\psi^h(d\alpha\ ds) \to \int_0^t \int_{\mathcal{U}} b(\bar{x}(s), \alpha) m(d\alpha\ ds), \tag{8.3}$$

$$\int_0^t \sigma^h(\bar{\xi}^h(q_\tau^h(s)))dw^h(s) \to \int_0^t \sigma(\bar{x}(s))dw(s). \tag{8.4}$$

Equation (8.3) follows directly due to the weak convergence and the continuity properties of $b(\cdot)$. To prove (8.4), due to the "stochastic integral," we need to discretize as in the proof of Theorem 2.1. For any function of a real variable $g(\cdot)$ and $\kappa > 0$, let $g_\kappa(s) = g(n\kappa), n\kappa \leq s < n\kappa + \kappa$. By the martingale

and quadratic variation properties of $w^h(\cdot)$, the mean square value of

$$\int_0^t \sigma^h(\bar{\xi}^h(q_\tau^h(s)))dw^h(s) - \int_0^t \sigma^h(\bar{\xi}_\kappa^h(q_\tau^h(s)))dw^h(s)$$
$$= \int_0^t \left[\sigma^h(\bar{\xi}^h(q_\tau^h(s))) - \sigma^h(\bar{\xi}_\kappa^h(q_\tau^h(s)))\right] dw^h(s)$$

is just the mean value of the square of the term in brackets in the right hand integrand, and it goes to zero as $\kappa \to 0$, uniformly in h. The integral in the right side of the first line can be written as a sum and the weak convergence implies that its limit as $h \to 0$ is $\int_0^t \sigma(x_\kappa(s))dw(s)$. The nonanticipativity of $(x(\cdot), y(\cdot), m(\cdot))$ with respect to the Wiener process $w(\cdot)$ is proved by using the analog of (2.2), namely,

$$Eh\left(\psi^h(s_i), \xi^h(s_i), w^h(s_i), y^h(s_i), \left\langle m_\psi^h, g_j \right\rangle (s_i), i \le I, j \le J\right) \times$$
$$\left(w^h(t+\tau) - w^h(t)\right) = 0,$$

and proceeding as below (2.2). Now, with the nonanticipativity proved, let $\kappa \to 0$ in $\int_0^t \sigma(x_\kappa(s))dw(s)$ to get (8.4). Thus we have proved that the limit satisfies (1.2) for some relaxed control $m(\cdot)$.

Since $m_\psi^h(\cdot)$ is the relaxed control representation of the interpolation of the optimal control sequence u_n^h with intervals $\Delta\tau_n^h$, by the definitions of $W^h(\cdot)$ and $V^h(\cdot)$ we have $V^h(\bar{\xi}_0^h) = W^h(\bar{\xi}_0^h, u^h)$. By the weak convergence and the continuity properties of $k(\cdot)$ in (A1.1), $W^h(\bar{\xi}_0^h, u^h) \to W(\hat{x}, m)$. By the minimality of $V(\hat{x})$, we must have $\liminf_h V^h(\bar{\xi}_0^h) \ge V(\hat{x})$. We need only prove that

$$\limsup_h V^h(\bar{\xi}_0^h) \le V(\hat{x}). \tag{8.5}$$

The proof of (8.5) for the no-delay case in [10, Chapter 10] can be readily adapted to the delay case. The proof depends on getting a piecewise constant approximation to the optimal control for (1.2) in terms of past samples of the driving Wiener process and control. The details of this approximation are a little more complicated for the present case, but the method in [10, Theorem 3.1, Chapter 10] carries over with some notational changes. The construction depended only on the continuity of the dynamical terms and weak-sense uniqueness as (A1.3). One needs to add the full initial condition of interest, including the initial control segment, where applicable. With this in hand, The converse inequality (8.5) is obtained as in the book. Owing to lack of space, the details will not be given here.

The proof of Theorem 6.3 is the same. Just use the interpolations $\xi^{h,\delta}(\cdot)$ and $\psi^{h,\delta}(\cdot)$ defined by (6.4) and (6.6), resp.

Delay in the control. Now consider Theorem 4.2, where the control is also delayed. The bracketed term in (4.9) converges to

$$\int_0^t \int_{\mathcal{U}} b(\bar{x}(s), \alpha, v) m(d\alpha, ds + v)$$

for all $v \in [-\tau, 0]$. The rest of the details of the convergence proof are as for the case where only the state is delayed. ■

References

[1] P. Billingsley. *Convergence of Probability Measures; Second edition.* Wiley, New York, 1999.

[2] R. F. Curtain and H. J. Zwart. *An Introduction to Infinite-Dimensional Linear Systems Theory.* Springer-Verlag, Berlin and New York, 1995.

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

[4] A. Ichikawa. Quadratic control of evolution equations with delays in control. *SIAM J. Control and Optim.*, 20:645–668, 1982.

[5] D. Jarvis and H.J. Kushner. Codes for optimal stochastic control: documentation and users guide. Technical report, Brown University, Lefschetz Center for Dynamical Systems Report 96-3, 1996. Documentation and codes are available on the internet: www.dam.brown.edu, then open Lefschetz Center.

[6] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus.* Springer-Verlag, New York, 1988.

[7] H.J. Kushner. *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations.* Academic Press, New York, 1977.

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

[9] H.J. Kushner and D. Barnea. On the control of linear functional-differential equations with quadratic cost. *SIAM J. Control Optim.*, 8:257–272, 1970.

[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.

[11] H.J. Kushner, D. Jarvis, and J. Yang. Controlled and optimally controlled multiplexing systems: A numerical exploration. *Queueing Systems*, 20:255–291, 1995.

[12] A. Manitius and T. Tran. Numerical approximations for hereditary systems with input and output delays: Convergence results abd convergence rates. *SIAM J. Control and Optimiz.*, 32:1332–1363, 1994.

[13] X. Mao. *Stability of Stochastic Differential Equations With Respect to Semimartingales.* Longman Scientific and Technical, Harlow, Essex, UK, 1991.

[14] J. C. Strikwerda. *Finite Difference Schemes and Partial Differential Equations*. Wadsworth and Brooks/Cole, Pacific Grove, CA, 1989.

[15] R. B. Vinter and R. H. Kwong. The infinite time quadratic control problem for linear systems with state and control delays: An evolution equation approach. *SIAM J. Control and Optim.*, 19:139–153, 1981.