Inaugural - Dissertation

zur
Erlangung der Doktorwürde
der
Naturwissenschaftlich-Mathematischen Gesamtfakultät
der
Ruprecht - Karls - Universität
Heidelberg

vorgelegt von

Diplom-Mathematiker
Markus Fischer
aus Berlin

Datum (Tag der mündlichen Prüfung): 20. November 2007

Discretisation of continuous-time stochastic optimal control problems with delay

Gutachter:

Prof. Dr. Markus Reiß
Universität Heidelberg

Prof. Salah-Eldin A. Mohammed
Southern Illinois University, Carbondale

Abstract

In the present work, we study discretisation schemes for continuous-time stochastic optimal control problems with time delay. The dynamics of the control problems to be approximated are described by controlled stochastic delay (or functional) differential equations. The value functions associated with such control problems are defined on an infinite-dimensional function space.

The discretisation schemes studied are obtained by replacing the original control problem by a sequence of approximating discrete-time Markovian control problems with finite or finite-dimensional state space. Such a scheme is convergent if the value functions associated with the approximating control problems converge to the value function of the original problem.

Following a general method for the discretisation of continuous-time control problems, sufficient conditions for the convergence of discretisation schemes for a class of stochastic optimal control problems with delay are derived. The general method itself is cast in a formal framework.

A semi-discretisation scheme for a second class of stochastic optimal control problems with delay is proposed. Under standard assumptions, convergence of the scheme as well as uniform upper bounds on the discretisation error are obtained. The question of how to numerically solve the resulting discrete-time finite-dimensional control problems is also addressed.

Zusammenfassung

In der vorliegenden Arbeit untersuchen wir Schemata zur Diskretisierung von zeitstetigen stochastischen Kontrollproblemen mit Zeitverzögerung. Die Dynamik solcher Probleme wird von gesteuerten stochastischen Differentialgleichungen mit Gedächtnis beschrieben. Die zugehörigen Wertfunktionen sind auf einem unendlich-dimensionenalen Funktionenraum definiert.

Man erhält die Diskretisierungsschemata, die wir betrachten, indem man das Ausgangsproblem durch eine Folge approximierender zeitdiskreter Markovscher Kontrollprobleme ersetzt, deren Zustandsraum endlich-dimensional oder endlich ist. Ein solches Schema ist konvergent, wenn die Wertfunktionen der approximierenden Steuerungsprobleme gegen die Wertfunktion des ursprünglichen Problems streben.

Indem wir eine allgemeine Methode zur Diskretisierung zeitstetiger Kontrollprobleme anwenden, erhalten wir hinreichende Bedingungen für die Konvergenz von Diskretisierungsschemata für eine Klasse von stochastischen Steuerungsproblemen mit Zeitverzögerung. Die Methode zur Konvergenzanalyse selbst wird in einen formalen Rahmen gefasst.

Wir führen dann ein Semidiskretisierungsschema für eine zweite Klasse von stochastischen Steuerungsproblemen mit Zeitverzögerung ein. Unter üblichen Annahmen werden die Konvergenz des Schemas, aber auch gleichmäßige obere Schranken für den Diskretisierungsfehler hergeleitet. Schließlich widmen wir uns der Frage, wie die resultierenden endlich-dimensionalen Steuerungsprobleme numerisch gelöst werden können.

Danksagung

Für den Vorschlag des Gebietes, in das die vorliegende Arbeit fällt, und die Betreuung und ununterbrochene Unterstützung in allen fachlichen und auch außerfachlichen Fragen danke ich Prof. Markus Reiß. Mein Dank gebührt Prof.ssa Giovanna Nappo von der Universität "La Sapienza" für die fruchtbare Zusammenarbeit und ihre Gastfeundschaft während meines Aufenthalts in Rom von April bis September 2006. Für die Unterstützung bei der Organisation dieses Aufenthalts danke ich Prof. Peter Imkeller.

Den Mitgliedern der Forschungsgruppe "Stochastische Algorithmen und Nichtparametrische Statistik" am Weierstraß Institut (WIAS) in Berlin und den Mitgliedern der Statistikgruppe an der Universität Heidelberg danke ich für die angenehme gemeinsam verbrachte Zeit.

Für wertvolle Hinweise zur Theorie deterministischer Kontrollprobleme danke ich Prof. Maurizio Falcone. Für fachliche Gespräche, Anregungen und Unterstützung danke ich Stefan Ankirchner, Christian Bender, Christine Grün, Jan Johannes, Alexander Linke, Jan Neddermeyer, Eva Saskia Rohrbach und Karsten Tabelow.

Finanzielle Unterstützung von Seiten der Deutschen Forschungsgemeinschaft (DFG) und des ESF-Programms "Advanced Methods in Mathematical Finance" erkenne ich dankend an.

Schließlich und endlich danke ich Stella für ihre Geduld und ihr Verständnis.

Contents

Notation and abbreviations

$a \wedge b$	the smaller of the two numbers a, b
$a \vee b$	the bigger of the two numbers a, b
$\mathbf{1}_A$	indicator function of the set A
$\lfloor x \rfloor$	Gauß bracket of the real number x, that is, the largest integer not greater than x
$\lceil x \rceil$	the least integer not smaller than the real number x
$\mathbb{N}$	the set of natural numbers starting from one
$\mathbb{N}_0$	the set of all non-negative integers
$\mathbb{Z}$	the set of all integers
$\mathbf{B}(X)$	the space of all bounded real-valued functions on the set X
$\mathbf{C}(X, Y)$	the space of all continuous functions from the topological space X to the topological space Y
$\mathbf{C}(X)$	the space of all continuous real-valued functions on the topological space X
$\mathbf{D}(I)$	the Skorohod space of all real-valued càdlàg functions on the interval I
$\mathcal{C}$	in Chapter 3: the space $\mathbf{C}([-r, 0], \mathbb{R}^d)$
$\mathcal{C}_N$	in Chapter 3: the space $\mathbf{C}([-r - \frac{r}{N}, 0], \mathbb{R}^d)$
$\hat{\mathcal{C}}(N)$	in Chapter 3: the space of all $\varphi \in \mathcal{C}$ which are piecewise linear w. r. t. the grid $\{k\frac{r}{N} \mid k \in \mathbb{Z}\} \cap [-r, 0]$
A^{T}	transpose of the matrix A
càdlàg	right-continuous with left-hand limits (French acronym)
iff	if and only if
w. r. t.	with respect to

Chapter 1

Introduction

In this thesis, discretisation schemes for the approximation of continuous-time stochastic optimal control problems with time delay in the state dynamics are studied. Optimal control problems of this kind are infinite-dimensional control problems in a sense to be made precise below; they arise in engineering, economics and finance, among others.

We will derive results about the convergence of discretisation schemes. For a more specific semi-discretisation scheme, a priori bounds on the discretisation error will also be obtained. Such results are useful in the numerical solution of the original control problems.

Section 1.1 presents the class of optimal control problems we will be concerned with. In Section 1.2, some examples of optimal control problems with delay are given. Section 1.3 provides an overview over approaches and some results from the literature related to the discretisation of continuous-time optimal control problems – with or without delay. The organisation of the main part of the present work, its aim and scope are specified in Section 1.4

1.1 Stochastic optimal control problems with delay

Here, we introduce the type of optimal control problems we will be concerned with in this thesis. An optimal control problem is composed of two parts: a controlled system and a performance criterion. Given an initial condition of the system and a strategy, the system produces a unique output. A numerical value is assigned to each output according to the performance criterion. In this way, the "performance" of any strategy for any given initial condition is measured. The objective is to find strategies which perform as good as possible, and to calculate optimal performance values.

A controlled system is usually modelled as a discrete- or continuous-time (parametrised) dynamical system. In continuous time, controlled systems are often described by some kind of differential equation. The continuous-time controlled systems we are interested in, here, are modelled as stochastic (or deterministic) delay differential equations. We describe this class of equations in Subsection 1.1.1; a standard reference is Mohammed (1984). In Subsection 1.1.2, the class of stochastic optimal control problems with delay we study in this work is introduced. If the time delay is zero, then those problems reduce to ordinary stochastic optimal control problems. For this latter class of problems a well-developed theory exists; see, for instance, Yong and Zhou (1999) or Fleming and Soner

(2006). Basic optimality criteria, in particular the Principle of Dynamic Programming, are also mentioned in Subsection 1.1.2.

1.1.1 Stochastic delay differential equations

An ordinary Itô stochastic differential equation (SDE) is an equation of the form

$$dX(t) \;=\; b\big(t, X(t)\big)dt \;+\; \sigma\big(t, X(t)\big)dW(t), \quad t \geq 0, \tag{1.1}$$

where b is the *drift coefficient*, σ the *diffusion coefficient* and $W(.)$ a Wiener process. When the diffusion coefficient σ is zero, then Equation (1.1) takes on the form of an ordinary differential equation (ODE).

Let the state space be $\mathbb{R}^d$. The unknown function $X(.)$ in Equation (1.1) is then an $\mathbb{R}^d$-valued stochastic process with continuous or càdlàg[1] trajectories. The drift coefficient b is a function $[0,\infty) \times \mathbb{R}^d \to \mathbb{R}^d$, the diffusion coefficient σ a matrix-valued function $[0,\infty) \times \mathbb{R}^d \to \mathbb{R}^{d\times d_1}$, and W is a d_1-dimensional Wiener process defined on a filtered probability space $(\Omega, \mathcal{F}, \mathbf{P})$ adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$. In the notation, we often omit the dependence on $\omega \in \Omega$.

Equation (1.1) is to be understood as an integral equation. Standard assumptions on the coefficients b, σ guarantee that the initial value problem

$$X(t) \;=\; \begin{cases} X(0) \;+\; \int_0^t b\big(s, X(s)\big)ds \;+\; \int_0^t \sigma\big(s, X(s)\big)dW(s), & t > 0, \\ x, \quad t = 0, \end{cases} \tag{1.2}$$

possesses, for each $x \in \mathbb{R}^d$, a unique *strong solution*, that is, there is a unique (up to indistinguishability) $\mathbb{R}^d$-valued stochastic process $X = (X(t))_{t\geq 0}$ with continuous (or càdlàg) trajectories which is defined on $(\Omega, \mathcal{F}, \mathbf{P})$ and adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ such that Equation (1.2) is satisfied. The initial condition may also be stochastic, namely an $\mathcal{F}_0$-measurable $\mathbb{R}^d$-valued random variable.

Standard assumptions guaranteeing (strong) existence and uniqueness of solutions to Equation (1.2) are that b, σ are jointly measurable, Lipschitz continuous in the second variable (uniformly in the first) and that they satisfy a condition of sublinear growth in the second variable uniformly in the first; see paragraph 5.2.9 in Karatzas and Shreve (1991: p. 289), for example.

An important property of solutions of SDEs is that they are *Markov processes* w. r. t. the given filtration. Another equally important property is that they are continuous *semi-martingales* with semi-martingale decomposition given by the SDE itself.

In addition to the notion of *strong solution*, there is the notion of *weak solution* to an SDE. While strong solutions must live on the given probability space and must be adapted to the given filtration, weak solutions are only required to exist on some suitable stochastic basis; for example, the given filtration may be the one induced by the driving Wiener process, but solutions exist only when they are adapted to some larger filtration. Thus, there are two notions of existence and also two notions of uniqueness for an SDE, cf. Karatzas and Shreve (1991: Sects. 5.2 & 5.3).

[1] A function defined on an interval is *càdlàg* iff it is right-continuous with limits from the left.

The basic existence and uniqueness results carry over to the case of random coefficients, that is, b, σ are defined on $[0,\infty)\times\mathbb{R}^d\times\Omega$, provided b, σ are $(\mathcal{F}_t)$-adapted.[2] A *controlled SDE* can be represented in the form

$$dX(t) \;=\; b\big(t, X(t), u(t)\big)dt \;+\; \sigma(t, X(t), u(t)\big)dW(t), \quad t \geq 0, \tag{1.3}$$

where $u(.)$ is a *control process*, that is, an $(\mathcal{F}_t)$-adapted function $[0,\infty)\times\Omega\to\Gamma$. Here, Γ is a separable metric space, called the *space of control actions*. The coefficients in Equation (1.3) are deterministic functions $[0,\infty)\times\mathbb{R}^d\times\Gamma\to\mathbb{R}^d$ and $[0,\infty)\times\mathbb{R}^d\times\Gamma\to\mathbb{R}^{d\times d_1}$, respectively. For any given control process $u(.)$, however, $b(.,.,u(.))$, $\sigma(.,.,u(.))$ are adapted random coefficients.

A control process $u(.)$ such that the initial value problem corresponding to the controlled equation, here Equation (1.3), has a unique solution for each initial condition of interest will be called an *admissible strategy* or, simply, a *strategy*.

Throughout this thesis, we will represent control processes and strategies as (Γ-valued) functions defined on $[0,\infty)\times\Omega$, that is, defined on the product of time and scenario space. In the deterministic case, control processes reduce to functions $[0,\infty)\to\Gamma$, so-called *open-loop controls*. In the literature, control processes are often represented as *feedback controls*, that is, as deterministic functions defined on the product of time and state space. This representation, though being "natural" for the control of Markov processes, leads to technical difficulties already for discrete-time control problems, see Bertsekas and Shreve (1996). Feedback controls give rise to control processes in the form considered here.

Systems with delay are characterised by the property that their future evolution, as seen from any instant t, depends not only on t and the current state at t (and possibly the control), but also on states of the system a certain amount of time into the past. We will assume throughout that the system has *bounded memory*; thus, there is some finite $r>0$ such that the future evolution of the system as seen from time t depends only on t and system states over the period $[t-r,t]$. The parameter r is the maximal length of the memory or *delay*.

Stochastic delay differential equations (SDDEs) model systems with delay. The drift and diffusion coefficient of an SDDE are functions of time and trajectory *segments* (and, possibly, the control action). For an $\mathbb{R}^d$-valued function $\psi=\psi(.)$ living on the time interval $[-r,\infty)$, the *segment* of length r at time $t\in[0,\infty)$ is the function

$$\psi_t : [-r,0]\to\mathbb{R}^d, \qquad\qquad \psi_t(s) :=\; \psi(t+s), \quad s\in[-r,0].$$

If ψ is a continuous function, then the segment ψ_t at time t is a continuous function defined on $[-r,0]$. Likewise, if ψ is a càdlàg function, then the segment ψ_t at time t is a càdlàg function defined on $[-r,0]$.

Accordingly, if $(X(t))_{t\geq -r}$ is an $\mathbb{R}^d$-valued stochastic process with continuous trajectories, then the associated segment process $(X_t)_{t\geq 0}$ is a stochastic process taking its values in $\mathcal{C} := \mathbf{C}([-r,0],\mathbb{R}^d)$, the space of all $\mathbb{R}^d$-valued continuous functions on the interval $[-r,0]$. In this work, the space $\mathcal{C}$ will always be equipped with the supremum norm induced by the standard norm on $\mathbb{R}^d$.

[2]Strictly speaking, the statement about SDEs with random coefficients is true only if existence and uniqueness are understood in the strong sense. The notions of weak existence and weak uniqueness make sense also for solutions to controlled SDEs (with or without delay), cf. Section 3.1.

The segment process associated with an $\mathbb{R}^d$-valued stochastic process with càdlàg trajectories takes its values in the space $D_0 := D([-r,0],\mathbb{R}^d)$ of all $\mathbb{R}^d$-valued càdlàg functions on $[-r,0]$. We will refer to the space of trajectory segments as the *segment space.* As segment space we will choose either D_0 or $\mathcal{C}$. Notice that both spaces depend on the dimension d and the maximal length of the delay r; both d and r may vary. In the notation just introduced, an SDDE is of the form

$$dX(t) \;=\; b\big(t, X_t\big)dt \;+\; \sigma\big(t, X_t\big)dW(t), \quad t \geq 0. \tag{1.4}$$

The coefficients b, σ are now functions defined on $[0,\infty) \times \mathcal{D}$ or, in the case of random coefficients, on $[0,\infty) \times \mathcal{D} \times \Omega$, where $\mathcal{D}$ is the segment space. In order to obtain unique solutions, as initial condition we have to prescribe not a point $x \in \mathbb{R}^d$, but an initial segment $\varphi \in \mathcal{D}$. The initial segment might also be stochastic, namely a $\mathcal{D}$-valued $\mathcal{F}_0$-measurable random variable.

Let the segment space be the space $\mathcal{C}$ of continuous functions. Theorem II.2.1 in Mohammed (1984: p. 36) gives sufficient conditions such that, for each initial segment $\varphi \in \mathcal{C}$, the initial value problem

$$X(t) \;=\; \begin{cases} X(0) \;+\; \int_0^t b\big(s, X_s\big)ds \;+\; \int_0^t \sigma\big(s, X_s\big)dW(s), & t > 0, \\ \varphi(t), \quad t \in [-r,0], & \end{cases} \tag{1.5}$$

possesses a unique strong solution. Sufficient conditions are that the coefficients b, σ are measurable, are Lipschitz continuous in their segment variable under the supremum norm on $\mathcal{C}$ uniformly in the time variable, satisfy a linear growth condition and, in case they are random, are $(\mathcal{F}_t)$-progressively measurable.

Existence and uniqueness results for SDDEs can also be derived from the existence and uniqueness results for general functional SDEs as given, for instance, in Protter (2003: Ch. 5). There, the coefficients of the SDE are allowed to be random and to depend on the entire trajectory of the solution from time zero up to the current time. Initial conditions, however, are not trajectory segments, but points in $\mathbb{R}^d$ (or $\mathbb{R}^d$-valued random variables). Hence, to transfer the results, the drift and diffusion coefficient of the SDDE have to be redefined according to the given initial condition.

A *controlled SDDE* can be represented in the form

$$dX(t) \;=\; b\big(t, X_t, u(t)\big)dt \;+\; \sigma\big(t, X_t, u(t)\big)dW(t), \quad t \geq 0, \tag{1.6}$$

where $u(.)$ is a Γ-valued control process as above and b, σ are deterministic functions defined on $[0,\infty)\times\mathcal{D}\times\Gamma$. Existence and uniqueness are again a consequence of the general results applied to the random coefficients $b(.,.,u(.))$, $\sigma(.,.,u(.))$.

Observe that, in Equation (1.6), there is no delay in the control. At time $t > 0$, the coefficients b, σ depend on $u(t)$, and $u(t)$ is $\mathcal{F}_t$-measurable. Systems with delay in the control are outside the scope of the present work. Some kind of *implementation delay*, however, can be captured. Let w be some measurable function $\Gamma \to \mathbb{R}^l$. We can now add l additional dimensions to the state space $\mathbb{R}^d$ and consider an SDDE of the form

$$\begin{aligned} dX(t) \;&=\; \tilde{b}\big(t, X_t, Y_t, u(t)\big)dt \;+\; \tilde{\sigma}\big(t, X_t, Y_t, u(t)\big)dW(t), \\ dY(t) \;&=\; w\big(u(t)\big)dt, \end{aligned}$$

where $X(.)$ represents the first d components and $Y(.)$ the remaining l components. The coefficients $\tilde{b}$, $\tilde{\sigma}$ do not directly depend on the trajectory of $u(.)$, but, through Y_t, on segments of the process $(\int_0^t w(u(s))ds)_{t\geq 0}$ (and the initial segment Y_0); $\tilde{b}$, $\tilde{\sigma}$ may, for example, be functions of the difference $Y(t-\delta) - Y(t-r)$, where $\delta \in [0, r)$. In this way, *distributed* implementation delay can be modelled.

The solution of an SDDE like Equation (1.4) is, in general, not a Markov process. Suppose the coefficients of the SDDE are deterministic and uncontrolled (or else a constant control is applied), and let $X(.)$ be a solution process. Then the segment process $(X_t)_{t\geq 0}$ associated with $X(.)$ enjoys the Markov property, cf. Theorem III.1.1 in Mohammed (1984: p. 51). The Markov semigroup of linear operators induced by the transition probabilities of the segment process is weakly, but not strongly continuous. In particular, only the weak infinitesimal generator exists. A representation of the weak infinitesimal generator on a subset of its domain as a differential operator can be derived, cf. Theorem III.4.3 in Mohammed (1984: pp. 109-110).

The solution of an SDDE like Equation (1.4), although generally not a Markov process, is an Itô diffusion and a continuous semi-martingale (after time zero), and the Itô formula is applicable as usual. However, the usual Itô formula does not apply to the segment process. It is possible to develop an Itô-like calculus also for the segment processes associated with solutions of SDDEs, see Hu et al. (2004) and Yan and Mohammed (2005).

In this thesis, the driving noise process of the continuous-time systems will always be a Wiener process. Extensions of some of the results of this thesis, in particular the convergence analysis of Section 2.3, to systems driven by more general Lévy processes are possible.

In Chapters 2 and 3, we will be concerned with the discretisation of controlled systems with delay; here, we give some references to works concerned with the discretisation of uncontrolled systems with delay. An overview of numerical methods for uncontrolled SDDEs is given in Buckwar (2000). The simplest discretisation procedure is the Euler-Maruyama scheme. The work by Mao (2003) gives the rate of convergence for this scheme provided the SDDE has globally Lipschitz continuous coefficients and the dependence on the segments is in the form of generalised distributed delays; Proposition 3.3 in Section 3.2 of the present work provides a partial generalisation of Mao's results and uses arguments similar to those in Calzolari et al. (2007). The most common first order scheme is due to Milstein; see Hu et al. (2004) for the rate of convergence of this scheme applied to SDDEs with point delay.

1.1.2 Optimal control problems with delay

Recall that an optimal control problem is composed of a controlled system and a performance criterion. In what follows, the controlled system will always be described by a controlled SDDE like Equation (1.6) in Subsection 1.1.1. As initial condition, an element of the segment space $\mathcal{D}$ has to be prescribed; the segment space $\mathcal{D}$ will be either $\mathcal{C} := \mathbf{C}([-r, 0], \mathbb{R}^d)$ or $D_0 := D([-r, 0], \mathbb{R}^d)$. When, in addition to the initial segment $\varphi \in \mathcal{D}$, also the initial time $t_0 \in [0, \infty)$ is allowed to vary, then the system output for initial condition (t_0, φ) under control process $u(.)$ is determined by

(1.7)
$$X(t) \;=\; \begin{cases} \varphi(0) \;+\; \int_0^t b\big(t_0{+}s, X_s, u(s)\big)ds \;+\; \int_0^t \sigma(t_0{+}s, X_s, u(s)\big)dW(s), & t > 0, \\ \varphi(t), \quad t \in [-r, 0], \end{cases}$$

provided a unique solution $X = X^{t_0,\varphi,u}$ exists. Notice that the solution process $X(.)$ is defined over time $[-r, \infty)$, and the evolution of the system starts at time zero. The initial time t_0 only appears in the time argument of the coefficients.

The performance criterion is usually given in terms of a *cost functional*. The cost functionals we will consider are of the form

$$\big((t_0, \varphi), u(.)\big) \;\mapsto\; \mathbf{E}\left(\int_0^\tau f\big(t_0{+}s, X_s, u(s)\big)ds \;+\; g\big(X_\tau\big)\right), \tag{1.8}$$

where $X = X^{t_0,\varphi,u}$ is the solution to Equation (1.7) with initial condition (t_0, φ) under strategy $u(.)$ and τ is the *remaining time*, which may depend on t_0 and $X^{t_0,\varphi,u}$. The functions f, g are called the *cost rate* and *terminal cost*, respectively; they may depend on segments of the solution process; in general, f is a function $[0, \infty) \times \mathcal{D} \times \Gamma \to \mathbb{R}$, while g is a function $\mathcal{D} \to \mathbb{R}$.

Two versions of (1.8) will play a role. The first version gives rise to optimal control problems with *finite time horizon*. Choose $T > 0$, the deterministic *time horizon*, and set $\tau := T - t_0$. For the second version, choose a bounded open set $O \subset \mathbb{R}^d$, let $\hat{\tau}_O$ be the time of first exit of $X^{t_0,\varphi,u}$ from O and set $\tau := \hat{\tau}_O \wedge (T - t_0)$, where $T \in (0, \infty]$. This leads to optimal control problems with *random time horizon*.

Let $T > 0$ be finite, and let τ in (1.8) be $T - t_0$. Denote by $\mathcal{U}$ the set of admissible strategies, that is, the set of all those control processes $u(.)$ such that the initial value problem (1.7) yields a unique solution and the expectation in (1.8) a finite value for each initial condition $(t_0, \varphi) \in [0, T] \times \mathcal{D}$. Let the function $J\colon [0, T] \times \mathcal{D} \times \mathcal{U} \to \mathbb{R}$ be defined according to (1.8). Then J is the cost functional of an optimal control problem with finite time horizon.

Given an optimal control problem, there is a twofold objective: determine the minimal costs and find an *optimal strategy* for any initial condition. A strategy u^* is optimal for a given initial condition (t_0, φ) iff

$$J(t_0, \varphi, u^*) \;=\; \inf_{u \in \mathcal{U}} J(t_0, \varphi, u). \tag{1.9}$$

Existence of optimal strategies is not always guaranteed. Let us assume that the right hand side of Equation (1.9) is finite for all initial conditions (which is not necessarily the case). A direct minimisation of $J(t_0, \varphi, .)$ over the set $\mathcal{U}$ is usually not possible. Observe that initial conditions are time-state pairs; here, "states" are segments, that is, continuous or càdlàg functions on $[-r, 0]$.

A simple, yet fundamental approach, associated with the work of R. Bellman, to solving the dynamic optimisation problem is as follows. Introduce the function which assigns the minimal costs to each time-state pair. This function is called the *value function*. The values of the value function are, of course, unknown at this stage. If the system, the set of strategies and the cost functional have a certain additive structure in time, then the value function obeys *Bellman's Principle of Optimality* or, as it is also called, the *Principle of*

Dynamic Programming (PDP). Let V denote the value function of some optimal control problem; thus, V is a function $I \times \mathcal{S} \to \mathbb{R}$, where I is a time interval and $\mathcal{S}$ the "state space". Bellman's Principle then states that V satisfies

$$V(t,x) \;=\; \mathcal{T}_{t,r}\big(V(r,.)\big)(t,x) \quad \text{for all } x \in \mathcal{S},\ t,r \in I,\ t \leq r, \tag{1.10}$$

where $(\mathcal{T}_{t,r})$ is a two-parameter semigroup of monotone operators, called *Bellman operators*; see Fleming and Soner (2006: Sect. II.3) for this abstract formulation of the PDP. In the case at hand, the value function is defined by

$$V:\ [0,T] \times \mathcal{D} \to \mathbb{R}, \qquad V(t_0,\varphi) := \inf_{u \in \mathcal{U}} J(t_0,\varphi,u). \tag{1.11}$$

The Principle of Dynamic Programming takes on the form

$$V(t_0,\varphi) \;=\; \inf_{u\in\mathcal{U}} \mathbf{E}\left(\int_0^t f\big(t_0{+}s, X^u_s, u(s)\big)ds \;+\; V(t_0{+}t, X^u_t)\right), \quad 0 \leq t \leq T{-}t_0, \tag{1.12}$$

where X^u is the solution to Equation (1.7) under control process u with initial condition (t_0,φ). The minimisation on the right hand side of Equation (1.12) could be restricted to strategies defined on the time interval $[0,t]$.

Observe that the validity of the PDP has to be verified for each class of optimal control problems. For finite horizon stochastic (and deterministic) optimal control problems with delay, the PDP is indeed valid, see Larssen (2002) and also Appendix A.1 for the precise statement. The Markov property of the segment processes associated with solutions to Equation (1.7) under certain strategies is essential for the validity of the PDP in the form of Equation (1.12).

Notice that an optimal control problem with delay is, generally, infinite-dimensional in the sense that the corresponding value function lives on an infinite-dimensional function space, namely the segment space.

When the controlled processes are controlled Markov processes with finite-dimensional state space and the value function is sufficiently smooth, then the PDP in conjunction with Dynkin's formula allows to derive a partial differential equation (PDE) which is solved by the value function. Such a PDE, which involves the family of infinitesimal generators associated with the controlled Markov processes and characterises the value function, is called Hamilton-Jacobi-Bellman equation (HJB equation). In general, the value function need not be sufficiently smooth; consequently, the HJB equation does not necessarily possess classical solutions. Viscosity solutions provide the "right" generalisation of the concept of solution for HJB equations, see Fleming and Soner (2006).

In principle, it is possible to derive an HJB equation and define viscosity solutions also for controlled Markov processes with infinite-dimensional state space. See Chang et al. (2006) for results in this direction in connection with controlled SDDEs; also cf. Subsection 1.2.1. While, in Chapter 3, we will make extensive use of the PDP, we will not need any kind of HJB equation.

Let us also mention the fact that knowledge of the value function of an optimal control problem enables us to construct optimal or "nearly" optimal strategies. When time is discrete and the space of control actions Γ is finite or compact, then optimal strategies can be constructed in feedback form (and for each initial condition). We will return to this point in Section 3.4.

A second fundamental approach to optimal control problems is via *Pontryagin's Maximum Principle*, see Yong and Zhou (1999: Chs. 3 & 7) for the case of finite-dimensional controlled SDEs. Pontryagin's Principle provides *necessary* conditions which an optimal strategy and the associated optimal process (if such exist) have to satisfy in terms of the so-called adjoint equations, which evolve "backwards" in time. Under certain additional assumptions, the necessary conditions become sufficient. Versions of this principle for the control of deterministic systems with delay exist, cf. the example in Subsection 1.2.2. For stochastic control problems with delay of a special form, a version of the Pontryagin Maximum Principle is derived in Øksendal and Sulem (2001). For the results of this thesis, we will not rely on the Maximum Principle.

We have not made precise any assumptions on the coefficients of the control problems introduced above. This will be done in Subsection 2.3.1 and Section 3.1, respectively, where we specify the classes of continuous-time control problems to be approximated.

1.2 Examples of optimal control problems with delay

Some examples of continuous-time optimal control problems with delay, mostly from the literature, are given in this section. Control problems with linear dynamics and a "quadratic" cost criterion are well-studied in many settings. In Subsection 1.2.1, we cite results concerning the representation of optimal strategies for a class of linear quadratic regulators with point as well as distributed delay. In Subsection 1.2.2, a simple deterministic problem with point delay modelling the optimal allocation of production resources is presented. Subsection 1.2.3 describes a stochastic optimal control problem with delay which may arise in finance when pricing derivatives that depend on market external processes. Special cases of optimal control problems with delay are really equivalent to finite-dimensional control problems without delay. Subsection 1.2.4 contains results from the literature about those reducible problems.

A further example is the deterministic infinite horizon model of optimal economic growth studied in Boucekkine et al. (2005). Optimal control problems also arise in finance when the asset prices in a financial market are modelled as SDDEs, see Øksendal and Sulem (2001), for instance.

1.2.1 Linear quadratic control problems

When the system dynamics are linear in the state as well as in the control variable, the noise is additive and the cost functional has a quadratic form over a finite or infinite time horizon, then it is possible to derive a representation of the optimal strategies of the control problem. Such control problems are referred to as *linear quadratic problems* or *linear quadratic regulators*. Optimal strategies are given in feedback form; the representation involves the solution of an associated system of deterministic differential equations, the so-called *Riccati equations*. This is the case not only for finite-dimensional stochastic and deterministic systems, but also for systems described by abstract evolution equations (cf. Bensoussan et al., 2007).

Here, we just cite a result for finite horizon linear quadratic systems with one point and one distributed delay and additive noise, see Kolmanovskiĭ and Shaĭkhet (1996: Ch. 5).

We consider the time-homogeneous case. The dynamics of the control problem are given by the affine-linear equation

$$
\begin{aligned}
dX(t) &= A\,X(t)dt \;+\; A_1X(t-r)dt \;+\; \left(\int_{-r}^{0} G(s)X(t+s)ds\right)dt \\
&\quad +\; B\,u(t)dt \;+\; \sigma dW(t), \quad t>0,
\end{aligned}
\tag{1.13}
$$

where $r > 0$ is the delay length, $W(.)$ a d_1-dimensional standard Wiener process adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$, $u(.)$ a strategy, σ is a $d \times d_1$-matrix, A, A_1 are $d \times d$-matrices, G is a bounded continuous function $[-r, 0] \to \mathbb{R}^{d\times d}$, and B is a $d \times l$-matrix.

The strategy $u(.)$ in Equation (1.13) is any $\mathbb{R}^l$-valued $(\mathcal{F}_t)$-adapted square integrable process. Let $\mathcal{U}$ denote the set of all such processes. Let $D_0 := D([-r, 0], \mathbb{R}^d)$ denote the space of all $\mathbb{R}^d$-valued càdlàg functions on $[-r, 0]$. Given $\varphi \in D_0$ and a strategy $u(.) \in \mathcal{U}$, there is a unique (up to indistinguishability) d-dimensional càdlàg process $X(.) = X^{\varphi,u}(.)$ such that Equation (1.13) is satisfied and $X(t) = \varphi(t)$ for all $t \in [-r, 0]$.

Let $T > 0$ be the deterministic time horizon. The quadratic cost functional (for fixed initial time zero) is the function $J\colon D_0 \times \mathcal{U} \to \mathbb{R}$ given by

$$
J(\varphi,u) := \mathbf{E}\left(X^{\mathsf{T}}(T)C\,X(T) \;+\; \int_0^T \big(X^{\mathsf{T}}(t)\tilde{C}\,X(t) + u^{\mathsf{T}}(t)M\,u(t)\big)dt\right),
\tag{1.14}
$$

where C, $\tilde{C}$ are positive semi-definite $d\times d$-matrices and M is a positive definite $l\times l$-matrix. The associated value function (at initial time zero) is defined by

$$
V(\varphi) := \inf_{u\in\mathcal{U}} J(\varphi,u), \quad \varphi\in D_0.
$$

For the control problem determined by (1.13) and (1.14), a version of the Hamilton-Jacobi-Bellman equation[3] allows to derive a representation in feedback form of the optimal strategies. Define the function $u_0\colon [0, T] \times D_0 \to \mathbb{R}^l$ by

$$
u_0(t,\varphi) := -M^{-1}B^{\mathsf{T}}\left(P(t)\varphi(0) \;+\; \int_{-r}^{0} Q(t,s)\varphi(s)ds\right),
$$

where P, Q are matrix-valued functions $[0, T] \to \mathbb{R}^{d\times d}$ and $[0, T] \times [-r, 0] \to \mathbb{R}^{d\times d}$, respectively. The functions P, Q are determined by the following system of differential equations, which involves, in addition, the unknown functions $R\colon [0, T] \times [-r, 0] \times [-r, 0] \to \mathbb{R}^{d\times d}$ and $g\colon [0, T] \to \mathbb{R}$:

(1.15)

$$
\begin{aligned}
&\tfrac{d}{dt}P(t) + A^{\mathsf{T}}P(t) + P(t)A(t) + Q(t,0) + Q^{\mathsf{T}}(t,0) + \tilde{C} &&= P(t)B\,M^{-1}B^{\mathsf{T}}P(t),\\
&\big(\tfrac{\partial}{\partial t} - \tfrac{\partial}{\partial s}\big)Q(t,s) + P(t)G(t,s) + A^{\mathsf{T}}Q(t,s) + R(t,0,\tau) &&= P(t)B\,M^{-1}B^{\mathsf{T}}Q(t,s),\\
&\big(\tfrac{\partial}{\partial t} - \tfrac{\partial}{\partial s} - \tfrac{\partial}{\partial \tau}\big)R(t,s,\tau) + G^{\mathsf{T}}(t,s)Q(t,\tau) + Q^{\mathsf{T}}(t,s)G(t,\tau) &&= Q^{\mathsf{T}}(t,s)B\,M^{-1}B^{\mathsf{T}}Q(t,\tau),\\
&\tfrac{d}{dt}g(t) + \mathrm{trace}\big(\sigma^{\mathsf{T}}P(t)\sigma\big) \;=\; 0, && t\in[0,T],\; s,\tau\in[-r,0],
\end{aligned}
$$

[3] The derivation of the HJB equation in Kolmanovskiĭ and Shaĭkhet (1996: Ch. 5) is not completely rigorous; see Chang et al. (2006) and the references therein for a more careful treatment. The development there starts from the expression for the weak infinitesimal generator of the segment process as derived in Mohammed (1984).

with boundary conditions

$$(1.16)\qquad \begin{array}{llll} P(T) = C, & R(T,s,\tau) = 0, & Q(T,s) = 0, \\ g(T) = 0, & P(t)A_1 = Q(t,-r), & A_1^{\mathsf{T}} Q(t,s) = R(t,-r,s), \\ & & t \in [0,T],\ s,\tau \in [-r,0]. \end{array}$$

Equations (1.15) can be shown to possess a unique continuously differentiable solution (P,Q,R,g) under boundary conditions (1.16), see Theorem 5.2.1 in Kolmanovskiĭ and Shaĭkhet (1996: p. 124). It is also shown that u_0 is indeed an optimal feedback control. This means the following. Let $\varphi \in D_0$, and let $X^* = X^{*,\varphi}$ be the unique solution to

$$(1.17)\qquad X^*(t) = \begin{cases} \varphi(0) + \int_0^t \Big(A\,X^*(\tau) + A_1 X^*(\tau-r) + \Big(\int_{-r}^0 G(s) X^*(\tau+s) ds\Big)\Big) d\tau \\ \quad + \int_0^t B\,u_0\big(\tau, X^*_\tau)\big) d\tau + \sigma\,W(t) \quad \text{if } t > 0, \\ \varphi(t) \quad \text{if } t \in [-r,0]. \end{cases}$$

Recall the notation X^*_τ for the segment of $X^*(.)$ at time τ. Observe that u_0 is Lipschitz continuous (in supremum norm) in its segment variable, whence strong existence and uniqueness of the solution X^* are guaranteed. Indeed, due to the form of u_0, Equation (1.17) is an affine-linear uncontrolled SDDE, and X^* can be expressed by a variation-of-constants formula. Set $u^*(t) := u_0(t, X^*)$, $t \geq 0$. Then it holds that $J(\varphi, u^*) = V(\varphi)$, that is, u^* is an optimal strategy and X^* is the optimal process for the given initial condition φ.

In special cases, Equations (1.15) can be solved explicitly. For general linear quadratic problems, they may serve as a starting point for the numerical computation of optimal strategies and minimal costs.

1.2.2 A simple model of resource allocation

The following finite-horizon deterministic problem can be interpreted as a simplified model of optimal resource allocation; see Bertsekas (2005: Ex. 3.1.2, 3.3.2) for the non-delay case. Let $T > 0$ be the time horizon, let $r \in [0,T)$ be the length of the time delay, and $c > 0$ a parameter. The dynamics of the model are given by

$$(1.18)\qquad \begin{cases} \dot{x}(t) = c\,u(t)\,x(t-r), & \text{if } t > 0, \\ x(t) = \varphi(t), & \text{if } t \in [-r,0], \end{cases}$$

where the initial path φ is in $\mathcal{C}_+ := \mathbf{C}([-r,0],(0,\infty))$; if $r = 0$, then φ is just a positive real number. An admissible strategy $u(.)$ is any element of the set $\mathcal{U}$ of all Borel measurable functions $[0,\infty) \to [0,1]$.

The initial time will be fixed and equal to zero. The objective is to maximise, for each initial segment $\varphi \in \mathcal{C}_+$, the cost functional

$$\tilde{J}(\varphi,u) := \int_0^T \big(1 - u(t)\big) x(t-r) dt$$

over $u \in \mathcal{U}$. Clearly, this is equivalent to minimising

$$(1.19)\qquad J(\varphi,u) := \int_0^T \big(u(t) - 1\big) x(t-r) dt$$

over $u \in \mathcal{U}$, since $\sup_{u\in\mathcal{U}} \tilde{J}(.,.,u) = -\inf_{u\in\mathcal{U}} J(.,.,u)$.

A possible interpretation of the control problem determined by (1.18) and (1.19) is the following (cf. Bertsekas, 2005: Ex. 3.1.2). The state trajectory $x(.) = x^u(.)$ describes the production rate of certain commodities (e. g. wheat). Consequently, the total amount of goods produced in any time period $[0,\tau]$ is $\int_0^\tau x(t)dt$ (in suitable units). During the entire production period (from time zero to time T) the producer has the choice between producing for reinvestment and the production of storable goods. This means that, at any time $t \in [0,T]$, a portion $u(t) \in [0,1]$ of the production rate is allocated to reinvestment, while the remaining portion $1 - u(t)$ goes into the production of storable goods. The production rate changes in proportion to the level of reinvestment. If reinvestment is zero, then the production rate will remain constant.

In order to justify Equation (1.18), it is instructive to consider small time steps. Denote by $y(t)$ the total amount of goods produced up to time t, that is, $y(t) = y(0) + \int_0^t x(s)ds$, where $x(.)$ is the production rate. Let $h > 0$ be the length of a small time step. Clearly, $y(t+h) = y(t) + \int_t^{t+h} x(s)ds$. On the other hand,

$$x(t+h) \;\approx\; x(t) \;+\; c \cdot u(t)\,(y(t) - y(t-h))\,,$$

where the parameter $c > 0$ regulates the effectiveness of reinvestment. Letting h tend to zero and taking into account the initial condition, we obtain (1.18).

The objective is to maximise the total amount of stored goods, that is, to maximise $\tilde{J}(\varphi,u)$ over all strategies $u \in \mathcal{U}$ for each initial condition $\varphi \in \mathcal{C}_+$ on the production rate. Equivalently, we can minimise $J(\varphi,u)$ over $u \in \mathcal{U}$ for each $\varphi \in \mathcal{C}_+$.

The parameter r – when positive – introduces a time delay. At time $t \geq 0$, instead of allocating a portion $u(t)$ of the *current* production rate $x(t)$, the producer may allocate a portion of the *past* production rate $x(t-r)$. The total amount of stored goods is measured accordingly, namely by $\int_0^T (1 - u(t))x(t-r)dt$. We may think of r as the time it takes to transform or sell the goods produced.

The control problem described above can be solved explicitly, and optimal strategies can be found. In the non-delay case, this is possible by relying on the Pontryagin Maximum Principle, see Theorem 3.2.1 in Yong and Zhou (1999: p. 103), for example. Pontryagin's Maximum Principle gives a set of *necessary* conditions an optimal strategy must satisfy (if it exists) in terms of the so-called adjoint variable. Under additional assumptions, those conditions are also sufficient for a strategy to be optimal, cf. Theorem 3.2.5 in Yong and Zhou (1999: p. 112).

In case $r = 0$, the solution of the above simple control problem by means of the Maximum Principle is given in Bertsekas (2005: pp. 121-122). For $r \geq 0$, we may rely on a version of Pontryagin's Principle for deterministic systems with delay, cf. Gabasov and Kirillova (1977: p. 840).

Given any initial segment $\varphi \in \mathcal{C}_+$, it can be shown that a corresponding optimal strategy satisfies

$$u^*(t) \;=\; \begin{cases} 1 & \text{if } p(t) \geq \frac{1}{c}, \\ 0 & \text{if } p(t) < \frac{1}{c}, \end{cases} \qquad t \in [0,T],$$

where $p(.)$ is the solution to the adjoint terminal-value problem given, in the case at hand, by

$$(1.20)\qquad \begin{cases} p(t) \;=\; 0, & t \in [T-r, T], \\ \dot{p}(t) \;=\; -1, & t \in [T-r-\frac{1}{c}, T-r], \\ \dot{p}(t) \;=\; -c\,p(t+r), & t \in [0, T-r-\frac{1}{c}]. \end{cases}$$

Equations (1.20) describe a deterministic "backward" delay differential equation with terminal condition. It follows that an optimal strategy is given by

$$(1.21)\qquad u^*(t) \;=\; \begin{cases} 1 & \text{if } t \in [0, T-r-\frac{1}{c}], \\ 0 & \text{if } t \in [T-r-\frac{1}{c}, T]. \end{cases}$$

Observe that u^* depends on the delay length r and the "effectiveness" parameter $c > 0$, but not on the initial condition. The minimal costs $J(\varphi, u^*)$, however, depend on $\varphi \in \mathcal{C}_+$. If $r = 0$, we have an explicit solution, if $r > 0$, we can integrate in steps of length r.

The optimal strategy as given by Equation (1.21) is of *bang-bang* type. It consists in reinvesting as much as possible before a critical switching time $T-r-\frac{1}{c}$, and then not to reinvest any more, but to produce and store until the final time is reached.

1.2.3 Pricing of weather derivatives

The example problem of this subsection is based on Ankirchner et al. (2007), where pricing and hedging of insurance derivatives that depend on external physical processes is studied.

Let $X(.)$ be a continuous-time stochastic process (one-dimensional, for simplicity) describing some physical quantity, e. g. surface temperature at a given place or averaged over a certain region. Suppose X can be modelled as an SDDE of the form

$$(1.22)\qquad dX(t) \;=\; b\big(t, X_t\big)dt \;+\; \sigma\big(t, X_t\big)dW(t), \quad t > 0,$$

where X_t is the segment of length $r > 0$ of $X(.)$ at time t, $W(.)$ a standard Wiener process and b, σ are appropriate functions; Equation (1.22) should possess a unique solution for each initial condition $\varphi \in \mathcal{D}$, where $\mathcal{D} = \mathbf{C}([-r, 0])$, for example.

Suppose further that an economic agent A (e. g. an insurance company) intends to sell a financial derivative on the process $X(.)$. At maturity $T > 0$, the derivative yields – from the perspective of A – an income $F(X_T)$, where F is some deterministic function $\mathcal{D} \to \mathbb{R}$. The income thus may depend on the evolution of $X(.)$ over the period $[T-r, T]$. Notice that the length r of the time delay may be artificially increased.

The question is which price A should ask for the derivative corresponding to F. It is assumed that A has the possibility to invest in a financial market. In this market, there is a risky asset with price process $S(.)$ such that $S(.)$ and $X(.)$ are correlated. We assume that $S(.)$ is given by the modified Black and Scholes model

$$(1.23)\qquad dS(t) \;=\; \mu\big(t, S(t)\big)S(t)dt \;+\; \beta_1 S(t)dW(t) \;+\; \beta_2 S(t)d\tilde{W}(t),$$

where $\tilde{W}$ is a second standard Wiener process independent of the first. The processes $S(.)$ and $X(.)$ are correlated through $\beta_1 \neq 0$.

The financial market is incomplete, as the physical quantity described by X is not traded. The price p of the derivative that A should ask can be determined as the *utility* *indifference price*, provided a *utility function* describing A's attitude towards risk is given. Let $\Psi\colon \mathbb{R} \to \mathbb{R}$ denote such a function. We assume that Ψ is an exponential utility function. Then the price p is determined by the equation

$$\sup_{u\in\mathcal{U}} \mathbf{E}\left(\Psi\big(V^u(T) + F(X_T) - p\big)\right) \;=\; \sup_{u\in\mathcal{U}} \mathbf{E}\left(\Psi\big(V^u(T)\big)\right), \tag{1.24}$$

where $V^u(.)$ is the value of A's portfolio under investment strategy $u \in \mathcal{U}$; see Ankirchner et al. (2007) for the details. For an exponential utility function Ψ, the unknown p in Equation (1.24) factors out, and, on the left hand side of (1.24), we have a stochastic optimal control problem with delay of the type studied in Chapter 3.

1.2.4 Delay problems reducible to finite dimension

In this subsection, we follow Bauer and Rieder (2005); but also cf. Elsanosi et al. (2000) and Larssen and Risebro (2003), where a similar approach is taken.

The value function of an optimal control problem with delay lives, for fixed initial time, on the segment space associated with the system dynamics. The segment space is, apart from the case when the delay length r is equal to zero, an infinite-dimensional space of functions, say $\mathcal{D}$; for example, $\mathcal{D} = \mathbf{C}([-r,0],\mathbb{R}^d)$. In general, it is not possible to reduce the value function to a finite-dimensional object, that is, it is not generally possible to find a number $n \in \mathbb{N}$ and continuous functions $\Theta\colon \mathcal{D} \to \mathbb{R}^n$, $\Psi\colon \mathbb{R}^n \to \mathbb{R}$ such that $V = \Psi \circ \Theta$.

If the controlled SDDE as well as the cost functional have a special form and certain additional assumptions are fulfilled, then the optimal control problem with delay is reducible to a control problem without delay, that is, the problem is effectively finite-dimensional.

Let Γ be a closed subset of Euclidean space (of any dimension). Let W be a one-dimensional standard Wiener process adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$. Denote by $\mathcal{U}$ the set of all $(\mathcal{F}_t)$-progressively measurable Γ-valued processes. Let $r > 0$, and let the dynamics of the control problem with delay be given by the one-dimensional controlled SDDE

$$\begin{aligned} dX(t) \;&= \mu_1\big(t, X(t), Y(t), u(t)\big)dt \;+\; \mu_2\big(X(t), Y(t)\big)\xi(t)dt \\ &\quad +\; \sigma\big(t, X(t), Y(t), u(t)\big)dW(t), \quad t > 0, \end{aligned} \tag{1.25}$$

where $u \in \mathcal{U}$ is a strategy, $\xi(t) := w(X(t-r))$ and $Y(t) := \int_{-r}^{0} e^{\lambda\cdot s} w(X(t+s))ds$ for some continuously differentiable function $w\colon \mathbb{R} \to \mathbb{R}$ and a constant $\lambda \in \mathbb{R}$. Here, we only give the one-dimensional result with initial time set to zero; see Bauer and Rieder (2005) for a full account. The coefficients of Equation (1.25) are measurable functions

$$\mu_1 :\; [0,\infty) \times \mathbb{R} \times \mathbb{R} \times \Gamma \to \mathbb{R}, \qquad \mu_2 :\; \mathbb{R} \times \mathbb{R} \to \mathbb{R},$$
$$\sigma :\; [0,\infty) \times \mathbb{R} \times \mathbb{R} \times \Gamma \to \mathbb{R}.$$

Equation (1.25) describes a system whose evolution depends not only on the current state $X(.)$, but also on a certain weighted average over the past, namely $Y(.)$, as well as a point delay, namely $\xi(.)$. Let us assume, for example, that

- $|\mu_2|$ is bounded and Lipschitz continuous,
- there is a constant $K > 0$ such that for all $t \geq 0$, $\gamma \in \Gamma$, $x, y \in \mathbb{R}$,

$$|\mu_1(t,x,y,\gamma)| \;+\; |\sigma(t,x,y,\gamma)| \;\leq\; K\big(1 + |x| \vee |y|\big),$$

- μ_1, σ are Lipschitz continuous in their respective second and third variable uniformly in the other variables.

Then, for every initial segment $\varphi \in \mathcal{C} := \mathbf{C}([-r,0])$ and every strategy $u \in \mathcal{U}$, there is a unique continuous process $X = X^{\varphi,u}$ such that Equation (1.25) is satisfied and $X(t) = \varphi(t)$ for all $t \in [-r, 0]$.

Let $T > 0$ be the deterministic time horizon. The cost functional of the optimal control problem is the function $J\colon \mathcal{C} \times \mathcal{U} \to \mathbb{R}$ given by

$$J(\varphi, u) := \mathbf{E}\left(\int_0^T f\big(t, X(t), Y(t), u(t)\big)dt \;+\; g\big(X(T), Y(T)\big)\right). \tag{1.26}$$

The associated value function V is defined by $V(\varphi) := \inf_{u\in\mathcal{U}} J(\varphi, u)$, $\varphi \in \mathcal{C}$.

At this point, an idea could be that $V(\varphi)$ depends on its argument $\varphi \in \mathcal{C}$ only through $\varphi(0)$ (correspondig to $X(t)$) and $\int_{-r}^0 w(\varphi(s))ds$ (correspondig to $Y(t)$). Observe that in Equation (1.25) there is still the point delay $\xi(t) = w(X(t-r))$. Also notice that the process $Y(.)$ is of bounded variation. Let $\Psi \in \mathbf{C}^{2,1}(\mathbb{R}\times\mathbb{R})$. By Itô's formula, for any solution $X(.)$ and the associated process $Y(.)$,

$$\begin{aligned} d\Psi\big(X(t), Y(t)\big) \;&=\; \tfrac{\partial}{\partial x}\Psi\big(X(t), Y(t)\big)dX(t) \;+\; \tfrac{\partial}{\partial y}\Psi\big(X(t), Y(t)\big)dY(t) \\ &\quad+\; \tfrac{\partial^2}{\partial x\partial x}\Psi\big(X(t), Y(t)\big)d\langle X, X\rangle(t). \end{aligned}$$

While expressions for $dX(t)$ and $d\langle X, X\rangle(t)$ now follow from Equation (1.25), we have

$$dY(t) \;=\; w\big(X(t)\big)dt \;-\; e^{-\lambda r}\xi(t)dt \;-\; \lambda Y(t)dt, \tag{1.27}$$

by construction of Y. Introduce the hypothesis that

$$\text{there is } \Psi \in \mathbf{C}^{2,1}(\mathbb{R}\times\mathbb{R}) \text{ such that for all } x, y \in \mathbb{R}, \\ \tfrac{\partial}{\partial x}\Psi(x,y)\mu_2(x,y) \;-\; e^{-\lambda r}\tfrac{\partial}{\partial y}\Psi(x,y) \;=\; 0. \tag{HT}$$

If Hypothesis (HT) holds, then the transformed process $\Psi(X, Y)$ obeys an equation of the form

$$d\Psi\big(X(t), Y(t)\big) \;=\; \tilde{\mu}\big(t, X(t), Y(t), u(t)\big)dt \;+\; \tilde{\sigma}\big(t, X(t), Y(t), u(t)\big)dW(t), \tag{1.28}$$

where the coefficients $\tilde{\mu}$, $\tilde{\sigma}$ can be expressed in terms of the original coefficients. Notice that the point delay $\xi(t)$ has disappeared. Indeed, Hypothesis (HT) has been chosen so that the "$\xi(t)$" term stemming from Equation (1.25) and the "$\xi(t)$" term in Equation (1.27) cancel out. The appearance of the point delay in Equation (1.27), on the other hand, is inevitable in view of the form of Y.

If the coefficients $\tilde{\mu}$, $\tilde{\sigma}$ are such that they depend on their x- and y-variable only through $\Psi(x, y)$, then $\Psi(X, Y)$, the transformed process, obeys an ordinary SDE of the form

$$d\Psi\big(X(t), Y(t)\big) \;=\; \bar{\mu}\big(t, \Psi(X(t), Y(t)), u(t)\big)dt \;+\; \bar{\sigma}\big(t, \Psi(X(t), Y(t)), u(t)\big)dW(t),$$

where $\bar{\mu}$, $\bar{\sigma}$ are the new coefficients which can be found by hypothesis.

Under Hypothesis (HT) and the reducibility hypothesis, the transformed dynamics can be written in terms of $Z(t) := \Psi(X(t), Y(t))$. If also the coefficients f, g in (1.26) are reducible, that is, if the coefficients of the cost functional depend on their x- and y-variable only through $\Psi(x, y)$, then a finite-dimensional control problem without delay arises which is related to the original control problem through the transformation Ψ and the corresponding reduction of the coefficients. Notice that the reducibility of the coefficients is a second hypothesis.

If the Hamilton-Jacobi-Bellman equation associated with the finite-dimensional control problem without delay admits a classical solution and if optimal strategies exist, then the finite-dimensional and the delay problem are equivalent in that their value functions are equivalent, see Theorem 1 in Bauer and Rieder (2005). That all hypotheses can be satisfied at once is shown in Bauer and Rieder (2005: Sects. 4-6) by way of specific examples: a linear quadratic regulator, a model of optimal consumption, and a deterministic model for congestion control.

1.3 Approximation of continuous-time control problems

There are various possible approaches to approximating continuous-time optimal control problems. We focus on those approaches which yield an approximation to the value function of the original problem. Recall that knowledge of the value function allows to choose optimal or nearly optimal strategies so that an optimal control problem is essentially solved once its value function is known. The methods we mention were mostly developed for finite-dimensional systems – stochastic as well as deterministic.

A basic idea is to replace the original control problem by a sequence of control problems which are numerically solvable in such a way that the associated value functions converge to the value function of the original problem. It is often possible to reinterpret a given scheme in terms of approximating control problems even though the scheme itself need not be defined in these terms.

A natural ansatz for constructing a suitable sequence of control problems is to derive their dynamics and cost functionals from a discretisation of the dynamics and cost functional of the original problem. This method, known as the "Markov chain method", was introduced by H. J. Kushner and is well-established in the case of finite-dimensional stochastic and deterministic optimal control problems, see Kushner and Dupuis (2001) and the references therein. The method allows to prove convergence of the approximating value functions to the value function of the original problem under very broad conditions. The most important condition to be satisfied is that of "local consistency" of the discretised dynamics with the original dynamics.

Due to its general nature, the Markov chain method can also be applied to control problems with delay. In Chapter 2, we will study this method in detail and develop an abstract framework for the proof of convergence. The framework may serve as a guide for using the Markov chain method in the convergence analysis of approximation schemes for various classes of optimal control problems. In Section 2.3, the convergence analysis is carried out for the discretisation of stochastic optimal control problems with delay and a random time horizon. We note, however, that while the method is well-suited for establishing convergence of a scheme, it usually provides no information about the speed of convergence.

The value function of a continuous-time finite-dimensional optimal control problem can often be characterised as the unique viscosity solution of an associated partial differential equation. For classical control problems, that equation is the Hamilton-Jacobi-Bellman equation (HJB equation) associated with the control problem, which is a first order PDE in the case of a deterministic system and a second order PDE in the case of a stochastic system driven by a Wiener process. Examples show that the value function of a deterministic or degenerate stochastic control problem is not necessarily continuously differentiable (e. g. Fleming and Soner, 2006: II.2), whence classical solutions to the HJB equation do not always exist.[4]

An approximation to the value function of a continuous-time optimal control problem can be obtained by discretising the associated HJB equation. In particular, finite difference schemes can be used for the discretisation. In the case of finite-dimensional deterministic optimal control problems, convergence as well as rates of convergence for such schemes were obtained in the 1980s, see, for instance, Capuzzo Dolcetta and Ishii (1984) or Capuzzo Dolcetta and Falcone (1989). Mere convergence of a discretisation scheme for finite-dimensional deterministic and stochastic equations – without error bounds – can be checked by relying on a theorem due to Barles and Souganidis (1991). Their result is not limited to the analysis of HJB equations arising in control theory in that it applies to a wide class of equations possessing a viscosity solution.

About ten years ago, N. V. Krylov was the first to obtain rates of convergence for finite difference schemes approximating finite-dimensional stochastic control problems with controlled and possibly degenerate diffusion matrix, see Krylov (1999, 2000) and the references therein. The error bound obtained there in the special case of a time discretisation scheme with coefficients that are Lipschitz continuous in space and $\frac{1}{2}$-Hölder continuous in time is of order $h^{1/6}$ with h the length of the time step. Notice that in Krylov (1999) the order of convergence is given as $h^{1/3}$, where the time step has length h^2. When the space too is discretised, the ratio between time and space step is like h^2 against h or, equivalently, h vs. $\sqrt{h}$, which explains why the order of convergence is expressed in two different ways.

In Krylov (2005), sharp error bounds are obtained for fully discrete finite difference schemes in a special form; the bounds are of order $h^{1/2}$ in the mesh size h of the space discretisation and of order $\tau^{1/4}$ in the length τ of the time step.

Using purely analytic techniques from the theory of viscosity solutions, Barles and Jakobsen (2005, 2007) obtain error bounds for a broad class of finite difference schemes for the approximation of PDEs of Hamilton-Jacobi-Bellman type. In the case of a simple time discretisation scheme, the estimate for the speed of convergence they find is of order $h^{1/10}$ in the length h of the time step.

A possible ansatz for extending those results to the approximation of control problems with delay is to try to derive a HJB equation for the value function. Recall that a version of the Principle of Dynamic Programming still holds for delay systems, cf. Appendix A.1. As in the finite-dimensional setting, such an HJB equation is not guaranteed to admit classical (i. e. Fréchet-differentiable) solutions, and viscosity solutions have to be defined. The HJB equation can then be used as a starting point for constructing finite difference

[4]Generalised solutions for the HJB equation can be shown to exist also in the case when there are no classical solutions, but uniqueness of generalised solutions does not always hold. For viscosity solutions, on the other hand, existence and uniqueness can be guaranteed; moreover, viscosity solutions are the right solutions in the sense that they coincide with the value function of the underlying control problem.

schemes; see Chang et al. (2006) for first results in this direction.

A different approach to the approximation of control problems with delay is to start from a representation of the system dynamics as an evolution equation in Hilbert space. A suitable Hilbert space for this purpose is the space $M_2 := L^2([-r,0],\mathbb{R}^d) \times \mathbb{R}^d$, the *Delfour-Mitter space*, where $r > 0$ is the maximal length of the delay. Notice that the segment space $\mathbf{C}([-r,0],\mathbb{R}^d)$ introduced in Section 1.1 can be continuously embedded into M_2. Projection methods could be used to obtain an approximation scheme. For the representation of controlled deterministic systems with delay, especially linear systems, see Bensoussan et al. (2007: II.4); for how to represent stochastic systems with delay in Hilbert space, see Da Prato and Zabczyk (1992).

A further approach to the discretisation of optimal control problems is based on the Markov property. For a suitable choice of the state space, the controlled processes enjoy the Markov property provided only feedback controls are used as strategies. In the case of problems with delay, the Markov property holds for the segment processes. The dynamics of the original problem are represented by the family of controlled Markov semigroups. Discretisation schemes, especially for time discretisation, can then be studied in terms of convergence of the infinitesimal generators associated with the Markov semigroups; see van Dijk (1984) for an early work. Observe, however, that in order to obtain rates of convergence strong regularity hypotheses may be necessary already in the finite-dimensional case; this amounts to assuming that an optimal strategy in feedback form with sufficiently regular (e. g. Lipschitz continuous) feedback function exists or that the value function is two or three times continuously differentiable.

In this work, we will not use any infinite-dimensional representation of the system dynamics; instead, we will stick to the semi-martingale setting. The Markov property of the (infinite-dimensional) segment processes will nevertheless be exploited. In Section 2.3, we construct approximating discrete-time processes as "extended Markov chains". In Chapter 3, we will make extensive use of a version of the Principle of Dynamic Programming, which is based on the Markov property of the segment processes, cf. Appendix A.1.

Working in the semi-martingale setting has several advantages. Existence and uniqueness results for controlled SDDEs are well-established. There is an elaborate theory characterising weak convergence of $\mathbb{R}^d$-valued semi-martingales (e. g. Jacod and Shiryaev, 1987). This theory will be essential for the convergence analysis of Section 2.3. When the noise process of the dynamics of the original system is a Wiener process – as will be the case in this work –, then the solution processes are Itô diffusions. Strong results on their path regularity, in particular on the moments of their moduli of continuity, are available, cf. Appendix A.2 and Section 3.2. In Section 3.3, we will make use of a finite-dimensional "stochastic mean value theorem" due to N. V. Krylov. The main ingredients in the proof of that result are a mollification trick, the usual PDP and the Itô formula, cf. Theorem A.2 in Appendix A.3.

1.4 Aim and scope

The aim of this thesis is to study discretisation schemes for continuous-time stochastic optimal control problems with time delay in the state dynamics. The noise process driving the system of the original control problem will always be a Wiener process – one-dimensional in Section 2.3 and multi-dimensional in Chapter 3. The object to be approximated is the value function associated with the original control problem. We are concerned with questions of convergence as well as rates of convergence or bounds on the discretisation error. Error bounds tell how much cannot be lost (or gained) in passing from the original model to a discretised model. This is also the first step in the approximate numerical solution of continuous-time models. For a continuous-time control problem, an approximate numerical solution is usually the only kind of explicit solution available.

The general idea we follow is to replace the original continuous-time control problem by a sequence of approximating discrete-time control problems which are easier to solve numerically. Observe that the value function associated with a continuous-time control problem with delay of the type studied here lives, in general, on a function space, namely the segment space, whence the problem may be considered to be infinite-dimensional.

We will take two approaches. In Chapter 2, we follow the Markov chain method mentioned above, which is a recipe for constructing discretisation schemes and proving convergence in the sense of convergence of associated value functions. In Section 2.1, we present the method as it is found in the work of H. J. Kushner and others. In Section 2.2, we develop an abstract framework in which to state sufficient conditions guaranteeing convergence of approximation schemes. We then apply the method to the discretisation of a class of stochastic optimal control problems with delay and a random time horizon (the time of first exit from a compact set), cf. Section 2.3.

In Chapter 3, we study a more specific scheme, which applies to finite-horizon stochastic control problems with delay, controlled and possibly degenerate diffusion coefficient and multi-dimensional state as well as noise process, cf. Section 3.1. According to the scheme, time and segment space are discretised in two steps, see Sections 3.2 and 3.3. Under quite natural assumptions, we obtain not only convergence, but also bounds on the error of the discretisation scheme, see Section 3.4. The worst-case bound on the discretisation error in the general case is of order nearly $h^{1/12}$ in the length of the (inner) time step h.

The two-step scheme produces a sequence of approximating finite-dimensional control problems in discrete time. In Section 3.5, we address the question of how to solve these problems numerically. Instead of further discretising the state space – as in Section 2.3 –, we propose to use a variant of "approximate Dynamic Programming", exploiting the two-step structure of the scheme. Memory requirements, in particular, can be kept at a realistic level.[5] Notwithstanding the special structure of the discretisation scheme, its use is not confined to the approximation of finite horizon control problems. It should also apply to systems with a reflecting boundary or systems controlled up to the time of first exit from a compact set.

In this thesis, we are interested in discretisation schemes which yield an approximation to the value function of the original problem. The value function gives the *globally* minimal costs, and knowing it allows to construct *globally* optimal or nearly optimal strategies (for each initial condition). There are efficient procedures for finding *locally* optimal strategies and calculating *locally* minimal costs, but we will not be concerned with any of them. Moreover, we will not use any hypotheses on the regularity of optimal strategies (not

[5] The amount of computer memory required for the two-step scheme depends on the mesh size of the outer time grid. In terms of the length $\tilde{h}$ of this outer time step, a worst-case error bound of order $\tilde{h}^{1/2}\ln(1/\tilde{h})$ holds.

even existence) nor any regularity assumptions on the value function which are not a consequence of properties of the system coefficients. If such hypotheses were assumed, it would be possible to derive much better rates of convergence. The reason why we refrain from making such assumptions is that they are, usually, difficult or impossible to verify based on the information available about the system to be controlled.

Chapter 2

The Markov chain method

There is a general procedure, known as the “Markov chain method” and developed by Harold J. Kushner, for rendering optimal control problems in continuous time accessible to numerical computation. The basic idea is to construct a family of discrete optimal control problems by discretising the original dynamics and the original cost functional in time and space. The important point to establish then is whether the value functions associated with the discrete problems converge to the original value function as the mesh size of the discretisation tends to zero.

If the value functions converge, then the discrete control problems are a valid approximation to the original problem and standard algorithms, notably those based on Dynamic Programming (e. g. Bertsekas, 2005, 2007), can be applied – at least in principle – to calculate the minimal costs and to find optimal strategies for each of the discrete control problems.

When the dynamics of the original problem are given by ordinary deterministic or stochastic differential equations, suitable discrete control problems are obtained by replacing the original controlled differential equations with controlled Markov chains whose transition probabilities are consistent with the original dynamics. Under compactness and continuity assumptions on the original problem, a condition of *local consistency* for the transition probabilities of the controlled Markov chains suffices to guarantee convergence of the corresponding value functions.

In Section 2.1 we describe the Markov chain method following Kushner and Dupuis (2001) by means of a deterministic example problem. Section 2.2 sets up an abstract framework for approximating a given optimal control problem by a sequence of discrete problems. There the continuity and compactness assumptions underlying Kushner's method are made explicit. In Section 2.3, we apply the method to a class of stochastic control problems with delay and a stopping condition as time horizon. Most of the material of that section has been published in Fischer and Reiß (2007). In Kushner (2005), discretisation schemes for a class of stochastic control problems with delay and reflection are studied; however, the proofs for the delay case do not seem to be as closely analogous to the non-delay case as is suggested there. Section 2.4 contains a brief discussion of the scope of the Markov chain method.

2.1 Kushner's approximation method

As an illustration of how Kushner's method works, let us consider a deterministic optimal control problem with finite time horizon. The system dynamics are described by a controlled ordinary differential equation:

$$\dot{x}(t) \;=\; b\big(t_0+t, x(t), u(t)\big), \quad t > 0, \tag{2.1}$$

where b is a measurable function $[0,\infty) \times \mathbb{R}^d \times \Gamma \to \mathbb{R}^d$ and $u(.)$ a measurable function $[0,\infty) \to \Gamma$. The space Γ is called the space of *control actions* and it is assumed that Γ is a compact metric space. This hypothesis will be crucial later.

The initial state is $x(0) = y$ for some $y \in \mathbb{R}^d$. In the formulation adopted here, solutions $x(.)$ to Equation (2.1) – if there are any – always start at time zero, while the initial time $t_0 \geq 0$ enters the equation through the coefficient b. Let $\mathcal{U}_{ad}$ be the set of all Borel measurable functions $u : [0,\infty) \to \Gamma$ such that Equation (2.1) possesses a unique absolutely continuous solution $x(.) = x^{t_0,y,u}(.)$ for each $(t_0, y) \in [0,\infty) \times \mathbb{R}^d$. The elements of $\mathcal{U}_{ad}$ are called *admissible strategies* or, simply, *strategies*. Let $T > 0$ be the deterministic time horizon. Associated with strategy $u \in \mathcal{U}_{ad}$ and initial condition $(t_0, y) \in [0,T] \times \mathbb{R}^d$ are the costs

$$J_{det}\big(t_0, y, u(.)\big) := \int_0^{T-t_0} f\big(t_0+t, x^{t_0,y,u}(t), u(t)\big)dt \;+\; g\big(x^{t_0,y,u}(T-t_0)\big), \tag{2.2}$$

where f and g are suitable measurable functions $[0,\infty) \times \mathbb{R}^d \times \Gamma \to \mathbb{R}$ and $\mathbb{R}^d \to \mathbb{R}$, respectively, such that the above integral makes sense as an element of $[-\infty,\infty]$. The value function of the control problem determined by (2.1) and (2.2) is given by

$$V_{det}(t_0, y) := \inf_{u\in\mathcal{U}_{ad}} J_{det}\big(t_0, y, u(.)\big), \quad (t_0, y) \in [0,T] \times \mathbb{R}^d.$$

The idea is now to construct a suitable family $(\mathcal{P}_M)_{M\in\mathbb{N}}$ of optimal control problems in discrete time and with discrete state space so that the corresponding value functions converge pointwise to V_{det}. The problem $\mathcal{P}_M$ of degree M may be obtained as follows. Let $S_M \subset \mathbb{R}^d$ be a regular triangulation of the state space $\mathbb{R}^d$. Hence, any state $y \in \mathbb{R}^d$ can be represented as the convex combination of at most $d+1$ elements of S_M.

The dynamics of the control problem $\mathcal{P}_M$ are determined by the choice of a time-inhomogeneous controlled transition function $p^M : \mathbb{N}_0 \times S_M \times \Gamma \times S_M \to [0,1]$, that is, a function p^M which is jointly measurable and such that $p^M(n, y, \gamma, .)$ defines a probability distribution on S_M for all $n \in \mathbb{N}_0$, $y \in S_M$, $\gamma \in \Gamma$. Observe that the set S_M is at most countable. The number $p^M(n, y, \gamma, z)$ should be interpreted as the probability that, between time step n and $n+1$, the system switches from state $y \in S_M$ to state $z \in S_M$ under the action of control $\gamma \in \Gamma$.

Admissible strategies for the problem $\mathcal{P}_M$ are adapted sequences $(u(n))_{n\in\mathbb{N}_0}$ of Γ-valued random variables such that, for each initial condition $(n_0, y) \in \mathbb{N}_0 \times S_M$, there is an adapted S_M-valued sequence $(\xi(n))_{n\in\mathbb{N}_0}$ whose transition probabilities are given by the function p^M. Strictly speaking, an admissible strategy consists in a (complete) probability space $(\Omega, \mathcal{F}, \mathbf{P})$ equipped with a filtration $(\mathcal{F}_n)$ and an $(\mathcal{F}_n)$-adapted sequence $(u(n))$ of Γ-valued random variables; thus, the underlying filtered probability space is part of the strategy. For simplicity, we usually omit the stochastic basis from the notation.

If u is admissible, then, for each $(n_0, y) \in \mathbb{N}_0 \times S_M$, there is a discrete-time process $(\xi(n))$ such that for all $z \in S_M$, all $n \in \mathbb{N}_0$,

$$\mathbf{P}\big(\xi(n{+}1) = z \mid \mathcal{F}_n\big) \;=\; p^M\big(n_0{+}n, \xi(n), u(n), z\big), \qquad \xi(0) \;=\; y \quad \mathbf{P}\text{-a.s.},$$

where $\mathbf{P}$ is the probability measure which is part of the strategy. The distribution of ξ is uniquely determined by u, the transition function p^M and the initial condition (n_0, y). denote by $\mathcal{U}_{ad}^M$ the set of all admissible strategies of degree M.

An alternative to the above definition of the set of admissible strategies is to take feedback controls as strategies, that is, measurable functions of the time step, the current and past states of the system and the past control actions; see, for example, Hernández-Lerma and Lasserre (1996: Ch. 2). The advantage of the seemingly more complicated formulation here is that the admissible strategies are defined directly on the underlying probability space. The admissibility requirement above means that the underlying stochastic basis is rich enough so that the controlled process $(\xi(n))$ corresponding to a Γ-valued adapted sequence $(u(n))$ can be constructed. In Section 2.3, we will restrict the set of admissible stochastic bases to those carrying a Wiener process.

To conclude the construction of the problem $\mathcal{P}_M$ we need an analogue of the cost functional (2.2). Let $T_M \in \mathbb{N}_0$ be the discrete time horizon of degree M; for example, T_M could be equal to $\lfloor M \cdot T \rfloor$. We replace the integral by a sum and take expectations since $\mathcal{P}_M$ is a stochastic control problem. For $n \in \{0, \ldots, T_M\}$, $y \in S_M$, $(u(n)) \in \mathcal{U}_{ad}^M$ set

$$J_{det}^M\big(n_0, y, u\big) := \mathbf{E}\left(\sum_{n=0}^{T_M - n_0 - 1} f_M\big(n_0{+}n, \xi(n), u(n)\big)dt \;+\; g_M\big(\xi(T_M{-}n_0)\big)\right), \tag{2.3}$$

where $\xi = (\xi(n))$ is the discrete-time process associated with strategy u and initial condition (n_0, y). The functions f_M, g_M should be appropriate discretisations of f and g, respectively.

Suppose that one time step for the discrete problem of degree M corresponds to a step of length $h_M := \frac{1}{M}$ in continuous time. Then the requirement that the family $(p^M)_{M\in\mathbb{N}}$ of transition functions be locally consistent with the original dynamics means that for all $n, n_0 \in \mathbb{N}_0$, $y, z \in S_M$, $\gamma \in \Gamma$,

$$\sum_{z\in S_M} p^M(n_0{+}n, y, \gamma, z)\, z \quad = \quad y \;+\; h_M \cdot b\big(\tfrac{n_0+n}{M}, y, \gamma\big) \;+\; o(h_M), \tag{2.4}$$

where $o(h_M)$ is the M-th element of a sequence that tends to zero faster than $(h_M)_{M\in\mathbb{N}}$. In addition, one only has to require that the maximal jump size of the associated controlled Markov chains tend to zero as the discretisation degree M goes to infinity. Condition (2.4) can also be expressed in terms of the controlled Markov chains, cf. Section 2.3.3.

It is straightforward to construct a sequence of transition functions such that the jump height and the local consistency conditions can be fulfilled. We may define the function p^M by, for example,

$$p^M(n, y, \gamma, z) := \begin{cases} \lambda_i & \text{if } z = x_i \\ 0 & \text{else,} \end{cases}$$

where $x_1, \ldots, x_{d+1} \in S_M$ are the vertices of the simplex containing $y + h_M b(n, y, \gamma)$ and $\lambda_1, \ldots, \lambda_{d+1} \in [0, 1]$ are such that

$$y + h_M\, b(n, y, \gamma) \;=\; \sum_{i=1}^{d+1} \lambda_i\, x_i.$$

This choice yields a family of locally consistent transition functions provided the mesh size of the triangulations S_M, $M \in \mathbb{N}$, tends to zero like $h_M = \frac{1}{M}$ as M goes to infinity.

Besides local consistency of the family of transition probabilities there is a second important hypothesis in Kushner's method, namely the (semi-)continuity of the cost functionals with respect to a suitable notion of convergence. The cost functional J_{det} of the original problem, for instance, can be extended to a mapping which takes an initial condition $(t_0, y) \in [0, T] \times \mathbb{R}^d$, a strategy $u(.) \in \mathcal{U}_{ad}$ and an absolutely continuous function $x(.)$ and which yields a real number (or $\pm\infty$). In (2.2), the definition of J_{det}, on the other hand, the connection between $(t_0, y, u(.))$ and $x(.)$ is determined by the system dynamics as given by Equation (2.1). For the discrete stochastic problem of degree M, the cost functional J^M_{det} is a mapping which assigns a cost to any initial condition $(n_0, y) \in \{0, \ldots, T_M\} \times S_M$, strategy $(u(n)) \in \mathcal{U}^M_{ad}$ and S_M-valued adapted sequence $(\xi(n))$.

Consequently, we may interpret the cost functionals as defined on product spaces whose components are the set of initial conditions, the space of strategies and a suitable space of functions or random processes encompassing all possible trajectories of the system. We can even find a common product space for all the cost functionals involved. This can often be achieved by replacing the strategies and state sequences of the discrete-time problems by their piecewise constant continuous-time interpolations. As for the example problem, the deterministic strategies and solutions of the system equation are re-interpreted as particular "degenerate" random processes. The product space forming the domain of the extended cost functionals is endowed with a notion of convergence, namely that of weak convergence of random processes (or weak convergence of the associated probability distributions). The induced topology renders the cost functionals J_{det}, J^M_{det} continuous provided the coefficients f, g and f_M, g_M in (2.2) and (2.3), respectively, are continuous. We will see more details of this construction in Section 2.2.

There is a last important point in the set-up of Kushner's method: the compactification of the space of admissible strategies. Remember that the space of control actions Γ is assumed to be a compact metric space. Nonetheless, the space $\mathcal{U}_{ad}$ of admissible strategies, equipped with the topology of weak convergence, need not be compact. The reason why compactness of the strategy space is desirable, here, is that it guarantees the existence of optimal strategies for the original problem (for discrete-time control problems the compactness of Γ itself is sufficent). More generally, any sequence of strategies will possess limit points that are themselves strategies. A similar compactness property will be implicit in the assumptions of Theorem 2.1 in Section 2.2, the convergence result for the abstract framework. There, however, topological properties will regard the system space only.

We should stress that Kushner's method is not confined to such simple schemes as we have sketched for the example problem. In particular, for the discretisation of time, the grid need not be uniformly spaced. It is possible to analyse non-deterministic, adaptive schemes. Also a wide variety of different deterministic and stochastic optimal control problems can be handled. The system dynamics, for instance, might be described as a controlled jump-diffusion and the performance criterion might involve a random time horizon.

In the remainder of this Section, we introduce the concept of relaxed controls necessary for the compactification of the space of admissible strategies of a continuous-time control problem, cf. Kushner (1990: Ch. 3) and the references therein.

Definition 2.1. A *deterministic relaxed control* over a compact metric space Γ is a positive measure ρ on $\mathcal{B}(\Gamma \times [0,\infty))$, the Borel σ-algebra on $\Gamma \times [0,\infty)$, such that

$$\rho(\Gamma \times [0,t]) \;=\; t \quad \text{for all } t \geq 0. \tag{2.5}$$

Denote by $\mathcal{R}(\Gamma)$ the set of all deterministic relaxed controls over Γ.

For each $G \in \mathcal{B}(\Gamma)$, the function $t \mapsto \rho(G \times [0,t])$ is absolutely continuous with respect to the Lebesgue measure on $[0,\infty)$ by virtue of property (2.5). Denote by $\dot\rho(.,G)$ any density of $\rho(G \times [0,.])$. The family of densities $\dot\rho(.,G)$, $G \in \mathcal{B}(\Gamma)$, can be chosen in a Borel measurable way such that $\dot\rho(t,.)$ is a probability measure on $\mathcal{B}(\Gamma)$ for each $t \geq 0$, and

$$\rho(B) \;=\; \int_0^\infty \int_\Gamma \mathbf{1}_{\{(\gamma,t)\in B\}}\, \dot\rho(t,d\gamma)\, dt \quad \text{for all } B \in \mathcal{B}(\Gamma \times [0,\infty)).$$

The space $\mathcal{R}(\Gamma)$ of all deterministic relaxed controls over Γ is equipped with the *weak-compact topology* induced by the following notion of convergence: a sequence $(\rho_n)_{n\in\mathbb{N}}$ of relaxed controls converges to $\rho \in \mathcal{R}(\Gamma)$ if and only if

$$\int\limits_{\Gamma\times[0,\infty)} g(\gamma,t)\, d\rho_n(\gamma,t) \;\stackrel{n\to\infty}{\longrightarrow}\; \int\limits_{\Gamma\times[0,\infty)} g(\gamma,t)\, d\rho(\gamma,t) \quad \text{for all } g \in \mathbf{C}_c(\Gamma \times [0,\infty)),$$

where $\mathbf{C}_c(\Gamma \times [0,\infty))$ is the space of all real-valued continuous functions on $\Gamma \times [0,\infty)$ having compact support. By the compactness of Γ, $\mathcal{R}(\Gamma)$ is (sequentially) compact under the weak-compact topology.

Suppose $(\rho_n)_{n\in\mathbb{N}}$ is a convergent sequence in $\mathcal{R}(\Gamma)$ with limit point ρ. Given $T > 0$, let $\rho_{n|T}$ denote the restriction of ρ_n to the Borel σ-algebra on $\Gamma \times [0,T]$, and denote by $\rho_{|T}$ the restriction of ρ to $\mathcal{B}(\Gamma \times [0,T])$. Then $\rho_{n|T}$, $n \in \mathbb{N}$, $\rho_{|T}$ are all finite measures and $(\rho_{n|T})$ converges weakly to $\rho_{|T}$.

Any ordinary deterministic strategy $u(.)$ gives rise to a deterministic relaxed control, namely to

$$\rho(B) := \int_0^\infty \int_\Gamma \mathbf{1}_{\{(\gamma,t)\in B\}}\, \delta_{u(t)}(d\gamma)\, dt, \quad B \in \mathcal{B}(\Gamma \times [0,\infty)), \tag{2.6}$$

where δ_γ is the Dirac measure at $\gamma \in \Gamma$. Moreover, any deterministic relaxed control can be approximated – in the weak-compact topology – by a sequence of ordinary deterministic strategies.

The dynamics of a control problem described by controlled ordinary differential equations can be rewritten using relaxed controls. The relaxed version in integral form of Equation (2.1), for instance, is

$$\tilde{x}(t) \;=\; y \;+\; \int_{\Gamma\times[0,t]} b\big(t_0{+}s, \tilde{x}(s), \gamma\big)\, d\rho(\gamma,s), \quad t \geq 0, \tag{2.7}$$

where $(t_0, y) \in [0,\infty) \times \mathbb{R}^d$ is the initial condition.

In the stochastic case, the analogue of deterministic relaxed controls are relaxed control processes.

Definition 2.2. A *relaxed control process* over a compact metric space Γ is an $\mathcal{R}(\Gamma)$-valued random variable R defined on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbf{P})$ such that the mapping

$$\Omega \ni \omega \mapsto R(G \times [0,t])(\omega) \in [0,t]$$

is $\mathcal{F}_t$-measurable for all $t \geq 0$, $G \in \mathcal{B}(\Gamma)$.

Since, by definition, Condition (2.5) holds scenario-wise for a relaxed control process R, there is a family $(\dot{R}(t,.))$ of derivative measures such that, $\mathbf{P}$-almost surely,

$$R(B)(\omega) \;=\; \int_0^\infty \int_\Gamma \mathbf{1}_{\{(\gamma,t)\in B\}}\, \dot{R}(t,d\gamma)(\omega)\, dt \quad \text{for all } B \in \mathcal{B}(\Gamma \times [0,\infty)).$$

The family $(\dot{R}(t,.))$ can be constructed in a measurable way (cf. Kushner, 1990: p. 52). Any ordinary control process, that is, any Γ-valued $(\mathcal{F})_t$-adapted process, can be represented as a relaxed control process: for u an ordinary control process, set

$$R(B)(\omega) := \int_0^\infty \int_\Gamma \mathbf{1}_{\{(\gamma,t)\in B\}}\, \delta_{u(t,\omega)}(d\gamma)\, dt, \quad B \in \mathcal{B}(\Gamma \times [0,\infty)),\ \omega \in \Omega, \tag{2.8}$$

where δ_γ is the Dirac measure at $\gamma \in \Gamma$. Then R is the relaxed control representation of u.

2.2 An abstract framework

In this section we provide an abstract framework for the convergence analysis of discretisation schemes constructed according to the Markov chain method. The framework not only formalises the ideas outlined in Section 2.1, it also extends their scope of applicability. This is possible because Kushner's method does not require the system dynamics or cost functional to have any special structure. In particular, no additivity properties like the Principle of Dynamic Programming are exploited, not even the Markov property of the system is needed.

The definitions to be given below are illustrated by means of the deterministic control problem from Section 2.1. In Section 2.3, the convergence analysis for a class of stochastic optimal control problems with delay is carried out in detail. The work to be done there consists mainly in verifying that the hypotheses of Theorem 2.1 below are satisfied.

2.2.1 Optimisation and control problems

Optimal control problems are parametrised optimisation problems; the parameters correspond to the (initial) data for the system dynamics. Since the parameter set may be a singleton, we omit the modifier "parametrised" in the following definition.

Definition 2.3. An *optimisation problem* is a triple $(\mathcal{D}, \mathcal{A}, F)$, where $\mathcal{D}$, $\mathcal{A}$ are non-empty sets and F is a mapping $\mathcal{D} \times \mathcal{A} \to [-\infty, \infty]$.

The function F of an optimisation problem $(\mathcal{D}, \mathcal{A}, F)$ is called the *objective function* or *target function.* The set $\mathcal{D}$ is the *data set* of the problem, the set $\mathcal{A}$ may be called the *restrictor set.* Given a datum $\varphi \in \mathcal{D}$, the aim is to minimise (or maximise) $F(\varphi, .)$ over $\mathcal{A}$.

Definition 2.4. Let $\mathcal{P} = (\mathcal{D}, \mathcal{A}, F)$ be an optimisation problem. The function $V : \mathcal{D} \to [-\infty, \infty]$ defined by $V(\varphi) := \inf\{F(\varphi, \alpha) \mid \alpha \in \mathcal{A}\}$ is called the *value function* associated with $\mathcal{P}$. The problem $\mathcal{P}$ is *finite* iff its value function is finite, that is, iff V is $\mathbb{R}$-valued.

We restrict attention to minimisation problems. Maximisation problems can be rewritten as minimisation problems in the obvious way. More general optimisation problems could be formulated by letting the target function attain values in an arbitrary partially ordered set. Clearly, there is more structure to an optimal control problem than to an optimisation problem.

Definition 2.5. An *optimal control problem* is a quintuple $(\mathcal{D}, \mathcal{A}, \mathcal{H}, \Psi, J)$, where $\mathcal{D}$, $\mathcal{A}$, $\mathcal{H}$ are non-empty sets, Ψ is a mapping $\mathcal{D} \times \mathcal{A} \to \mathcal{H}$ and J is a mapping $\mathcal{H} \to [-\infty, \infty]$.

The components of an optimal control problem $(\mathcal{D}, \mathcal{A}, \mathcal{H}, \Psi, J)$ are denominated as follows: $\mathcal{D}$ is the *data set*, $\mathcal{A}$ is the *set of admissible strategies* or, simply, the set of strategies, $\mathcal{H}$ is called the *system space*, the mapping Ψ is the *system functional*, and J is called the *cost functional.* An optimal control problem $(\mathcal{D}, \mathcal{A}, \mathcal{H}, \Psi, J)$ gives rise to an optimisation problem, namely the triple $(\mathcal{D}, \mathcal{A}, F)$ with $F := J \circ \Psi$.

Definition 2.6. The *value function* associated with an optimal control problem is defined to be the value function associated with the induced optimisation problem. An optimal control problem is *finite* iff its value function is finite.

For simplicity, we will use the expression "control problem" without the modifying "optimal" also in the sense of "optimal control problem". Let us illustrate the definitions by applying them to the deterministic example problem of Section 2.1. In order to identify the components of that control problem according to Definition 2.5, set $\mathcal{D}_{det} := [0, T] \times \mathbb{R}^d$, let $\mathcal{A}_{det}$ be the set of strategies $\mathcal{U}_{ad}$, and let $\mathcal{H}_{det}$ be the set $\mathcal{D}_{det} \times \mathcal{A}_{det} \times \mathbf{C}([0, \infty), \mathbb{R}^d)$. Define the system functional Ψ_{det} as the mapping

$$\mathcal{D}_{det} \times \mathcal{A}_{det} \to \mathcal{H}_{det}, \qquad ((t_0, y), u(.)) \mapsto ((t_0, y), u(.), x^{t_0,y,u}(.)),$$

where $x^{t_0,y,u}(.)$ is the unique solution to Equation (2.1) under strategy $u(.)$ and initial condition (t_0, y). Lastly, define the cost functional J_{det} to be the mapping $\mathcal{H}_{det} \to [-\infty, \infty]$ given by

$$((t_0, y), u(.), x(.)) \mapsto \int_0^{T-t_0} f\big(t_0+t, x(t), u(t)\big)dt \; + \; g\big(x(T-t_0)\big).$$

Notice that in the above definition of J_{det} the function $x(.)$ is not necessarily a solution to Equation (2.1), but may be any continuous function $[0, \infty) \to \mathbb{R}^d$.

The quintuple $(\mathcal{D}_{det}, \mathcal{A}_{det}, \mathcal{H}_{det}, \Psi_{det}, J_{det})$ thus defined is an optimal control problem in the sense of Definition 2.5. Let V_{det} be the associated value function according to Definition 2.6. Then V_{det} coincides with the value function induced by the cost functional (2.2) in Section 2.1.

The representation of our example problem as an optimal control problem in the sense of Definition 2.5 is not unique. For example, in the definition of the system space $\mathcal{H}_{det}$ we could replace the component $\mathbf{C}([0,\infty),\mathbb{R}^d)$ by $\mathbf{C}_{abs}([0,\infty),\mathbb{R}^d)$, the space of all absolutely continuous functions $[0,\infty) \to \mathbb{R}^d$. The definitions given so far are about sets without any additional structure. For the discretisation and convergence analysis, we will require the system space to carry a suitable topology and the system functional and cost functional to have certain continuity properties. Nevertheless, we will neither obtain nor need unique representations of control problems. Before coming to this, let us introduce some basic relations between control problems.

Definition 2.7. Let $\mathcal{P} = (\mathcal{D},\mathcal{A},\mathcal{H},\Psi,J)$, $\tilde{\mathcal{P}} = (\tilde{\mathcal{D}},\tilde{\mathcal{A}},\tilde{\mathcal{H}},\tilde{\Psi},\tilde{J})$ be optimal control problems. Then $\mathcal{P}$ is a *subproblem* of $\tilde{\mathcal{P}}$ iff there are injective mappings $\iota_D\colon \mathcal{D} \hookrightarrow \tilde{\mathcal{D}}$, $\iota_A\colon \mathcal{A} \hookrightarrow \tilde{\mathcal{A}}$, $\iota_H\colon \mathcal{H} \hookrightarrow \tilde{\mathcal{H}}$ such that $\tilde{\Psi} \circ (\iota_D \times \iota_A) = \Psi$ and $\tilde{J} \circ \iota_H = J$.

The mappings ι_D, ι_A, ι_H which make $\mathcal{P}$ a subproblem of $\tilde{\mathcal{P}}$ are called *data embedding*, *strategy embedding* and *system embedding*, respectively. Notice that these embeddings need not be unique. We may say that $\tilde{\mathcal{P}}$ is a *superproblem* of $\mathcal{P}$ to indicate that $\mathcal{P}$ is a subproblem of $\tilde{\mathcal{P}}$.

Let V, $\tilde{V}$ be the value functions associated with the control problems $\mathcal{P}$ and $\tilde{\mathcal{P}}$, respectively. Suppose $\mathcal{P}$ is a subproblem of $\tilde{\mathcal{P}}$ with data embedding ι_D. Then, by definition, $\tilde{V}(\iota_D(\varphi)) \leq V(\varphi)$ for all $\varphi \in \mathcal{D}$. However, Definition 2.7 does not guarantee that $\tilde{V} \circ \iota_D = V$. The relation defined next is to ensure this property.

Definition 2.8. Let $\mathcal{P}$, $\tilde{\mathcal{P}}$ be optimal control problems with associated value functions V and $\tilde{V}$, respectively. Then $\tilde{\mathcal{P}}$ is a *relaxation* of $\mathcal{P}$ iff $\mathcal{P}$ is a subproblem of $\tilde{\mathcal{P}}$ for some data embedding ι_D such that $\tilde{V} \circ \iota_D = V$. The control problem $\tilde{\mathcal{P}}$ is a *restriction* of $\mathcal{P}$ iff $\tilde{\mathcal{P}}$ is a subproblem of $\mathcal{P}$ for some data embedding ι_D such that $\tilde{V} = V \circ \iota_D$.

Definition 2.9. Two optimal control problems $\mathcal{P}$ and $\tilde{\mathcal{P}}$ are said to be *compatible* iff $\tilde{\mathcal{P}}$ is a relaxation or restriction of $\mathcal{P}$ such that the data embedding involved is onto. In this case, we also say that $\mathcal{P}$ is *compatible with* $\tilde{\mathcal{P}}$ or vice versa.

Passing to a relaxation or restriction of a given control problem allows us to vary the set of strategies as well as the system. Hence, when two control problems are compatible, we can replace one with the other, at least as far as the value functions are concerned.

Definition 2.10. Let $(\mathcal{D},\mathcal{A},\mathcal{H},\Psi,J)$ be a control problem and $\varphi \in \mathcal{D}$. A strategy $\alpha \in \mathcal{A}$ is called an *optimal strategy* for the datum φ iff $J(\varphi,\alpha) = V(\varphi)$. A strategy $\alpha \in \mathcal{A}$ is called an *ε-optimal strategy* iff $\varepsilon > 0$ and $J(\varphi,\alpha) \leq V(\varphi) + \varepsilon$.

Thus, if α is an optimal strategy for a given datum $\varphi \in \mathcal{D}$, then $J(\varphi,.)$ attains its minimum at α. The existence of an optimal strategy cannot always be guaranteed, see Kushner and Dupuis (2001: p. 86) for a deterministic example. The passage to a relaxation of the control problem may allow us to work with a larger set of strategies where optimal strategies are guaranteed to exist. Recall that, at least in discrete time, the value function can be used for the synthesis of optimal or ε-optimal strategies. Observe that, while compatible control problems have value functions that can be identified with each other, the corresponding optimal or nearly optimal strategies do not, in general, coincide since the system functionals may be different.

Let us again apply these notions to the deterministic example problem. Recall the definition of the quintuple $\mathcal{P}_{det} = (\mathcal{D}_{det}, \mathcal{A}_{det}, \mathcal{H}_{det}, \Psi_{det}, J_{det})$. We construct a control problem $\tilde{\mathcal{P}}_{det} = (\tilde{\mathcal{D}}_{det}, \tilde{\mathcal{A}}_{det}, \tilde{\mathcal{H}}_{det}, \tilde{\Psi}_{det}, \tilde{J}_{det})$ such that $\tilde{\mathcal{P}}_{det}$ is a relaxation of $\mathcal{P}_{det}$. To this end, set $\tilde{\mathcal{D}}_{det} := \mathcal{D}_{det} = [0,T] \times \mathbb{R}^d$. Recall from Definition 2.1 how the set $\mathcal{R}(\Gamma)$ of deterministic relaxed controls with values in Γ was defined. Let $\tilde{\mathcal{A}}_{det}$ be the set of all $\rho \in \mathcal{R}(\Gamma)$ such that Equation (2.7) under ρ has a unique absolutely continuous solution for each initial condition $(t_0, \varphi) \in [0,\infty) \times \mathbb{R}^d$.

Define the system space $\tilde{\mathcal{H}}_{det}$ in analogy to $\mathcal{H}_{det}$ as the set $\tilde{\mathcal{D}}_{det} \times \tilde{\mathcal{A}}_{det} \times \mathbf{C}([0,\infty), \mathbb{R}^d)$. Define the system functional $\tilde{\Psi}_{det}$ by

$$\tilde{\mathcal{D}}_{det} \times \tilde{\mathcal{A}}_{det} \to \tilde{\mathcal{H}}_{det}, \qquad ((t_0,y), u(.)) \mapsto ((t_0,y), \rho, \tilde{x}^{t_0,y,\rho}(.)),$$

where $\tilde{x}^{t_0,y,\rho}(.)$ is the unique solution to Equation (2.7), the relaxed version of the system equation (2.1), under the deterministic relaxed control ρ and initial condition (t_0, y). Finally, define $\tilde{J}_{det}$ to be the mapping $\tilde{\mathcal{H}}_{det} \to [-\infty, \infty]$ given by

$$((t_0,y), \rho, x(.)) \mapsto \int_{\Gamma \times [0,T-t_0]} f\big(t_0+t, x(t), \gamma\big) d\rho(\gamma, t) \ + \ g\big(x(T-t_0)\big).$$

In order to verify that $\tilde{\mathcal{P}}_{det} = (\tilde{\mathcal{D}}_{det}, \tilde{\mathcal{A}}_{det}, \tilde{\mathcal{H}}_{det}, \tilde{\Psi}_{det}, \tilde{J}_{det})$ thus constructed is indeed a relaxation of $\mathcal{P}_{det}$, we recall from Section 2.1 that any ordinary control is associated with a deterministic relaxed control according to (2.6). This defines the strategy embedding. The data embedding is just the identity on $[0,T] \times \mathbb{R}^d$. The system embedding again uses the interpretation of ordinary control strategies as relaxed controls. The value functions of $\mathcal{P}_{det}$ and $\tilde{\mathcal{P}}_{det}$ are identical, because any deterministic relaxed control can be approximated by ordinary deterministic strategies, the cost functionals J_{det} and $\tilde{J}_{det}$ coincide for ordinary strategies, and $\tilde{J}_{det}$ is continuous with respect to the weak topology on $\tilde{\mathcal{H}}_{det}$.

The problem $\tilde{\mathcal{P}}_{det}$ can be further relaxed by allowing for relaxed control processes as strategies. In place of Equation (2.7), we then have the random ordinary differential equation

$$\tilde{x}(t,\omega) \ = \ y \ + \int_{\Gamma \times [0,t]} b\big(t_0+s, \tilde{x}(s,\omega), \gamma\big)\, dR(\gamma, s, \omega), \quad t \geq 0,\ \omega \in \Omega, \tag{2.9}$$

where R is a relaxed control process on the stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbf{P})$ in the sense of Definition 2.2. In the cost functional we must now take expectations, that is, instead of $\tilde{J}_{det}$ we have

$$((t_0,y), R, X) \mapsto \mathbf{E}\left(\int_{\Gamma \times [0,T-t_0]} f\big(t_0+t, X(t), \gamma\big) dR(\gamma, t) \ + \ g\big(X(T-t_0)\big)\right),$$

where $X(.)$ is any $\mathbb{R}^d$-valued continuous stochastic process adapted to the filtration coming with the relaxed control process R. The value functions of $\mathcal{P}_{det}$, $\tilde{\mathcal{P}}_{det}$ and the new problem will still be identical, because an ε-optimal strategy of the deterministic problem is also ε-optimal for almost all trajectories of the randomized problem.

Let us denote by $\hat{\mathcal{P}}_{det} = (\hat{\mathcal{D}}_{det}, \hat{\mathcal{A}}_{det}, \hat{\mathcal{H}}_{det}, \hat{\Psi}_{det}, \hat{J}_{det})$ the stochastic relaxation of $\tilde{\mathcal{P}}_{det}$ and $\mathcal{P}_{det}$. We choose $\hat{\mathcal{D}}_{det} := \tilde{\mathcal{D}}_{det}$ as the data set. The set of strategies $\hat{\mathcal{A}}_{det}$ is the set of pairs of stochastic bases and adapted relaxed control processes over Γ. Observe that the new cost functional as defined above actually depends only on the joint distribution of the processes X and R, and on the initial data. Therefore, as system space we could choose

$$\hat{\mathcal{D}}_{det} \times \{\text{probability measures on } \mathcal{B}(\mathcal{R}(\Gamma) \times \mathbf{C}([0,\infty),\mathbb{R}^d))\}.$$

In place of $\mathbf{C}([0,\infty),\mathbb{R}^d)$ we will take $D([0,\infty),\mathbb{R}^d)$, the *Skorohod space* of all functions $[0,\infty) \to \mathbb{R}^d$ which are continuous from the right and have limits from the left. The space $D([0,\infty),\mathbb{R}^d)$ is equipped with the *Skorohod topology*, cf. Billingsley (1999: Ch. 3) and also Section 2.3.1. The Skorohod space allows for an easier approximation of functions, even when they are continuous, in particular, for the approximation by piecewise constant functions. Hence, we define the system space $\hat{\mathcal{H}}_{det}$ to be the product space

$$\hat{\mathcal{D}}_{det} \times \{\text{probability measures on } \mathcal{B}(\mathcal{R}(\Gamma) \times D([0,\infty),\mathbb{R}^d))\}.$$

The system functional $\hat{\Psi}_{det}$ is the mapping

$$\hat{\mathcal{D}}_{det} \times \hat{\mathcal{A}}_{det} \ni ((t_0,\varphi),((\Omega,\mathcal{F},(\mathcal{F}_t),\mathbf{P}),R)) \mapsto ((t_0,\varphi),\mathbf{P}_{(R,X)}) \in \hat{\mathcal{H}}_{det}, \tag{2.10}$$

where $X = \tilde{x}^{t_0,y,R}(.)$ is the solution to Equation (2.9) and $\mathbf{P}_{(R,X)}$ denotes the joint distribution of R and X under $\mathbf{P}$, the probability measure which is part of the admissible strategy. For $\hat{\Psi}_{det}$ to be well-defined, we need that solutions to Equation (2.9) be unique in distribution. Lastly, we rewrite the cost functional and define $\hat{J}_{det}$ to be the mapping $\hat{\mathcal{H}}_{det} \to [-\infty,\infty]$ given by

$$((t_0,y),\mathbf{Q}) \mapsto \int \left(\int_{\Gamma\times[0,T-t_0]} f\big(t_0+t,\tilde{x}(t),\gamma\big)d\rho(\gamma,t) \;+\; g\big(\tilde{x}(T-t_0)\big)\right) d\mathbf{Q}(\rho,\tilde{x}(.)), \tag{2.11}$$

where the integral with respect to the probability measure $\mathbf{Q}$ is over $\mathcal{R}(\Gamma) \times D([0,\infty),\mathbb{R}^d)$.

For the approximation of a given control problem we need the following notions of discretisation.

Definition 2.11. Let $\mathcal{P} = (\mathcal{D},\mathcal{A},\mathcal{H},\Psi,J)$, $\bar{\mathcal{P}} = (\bar{\mathcal{D}},\bar{\mathcal{A}},\bar{\mathcal{H}},\bar{\Psi},\bar{J})$ be control problems. Then $\bar{\mathcal{P}}$ is a *direct discretisation* of $\mathcal{P}$ iff $\bar{\mathcal{H}} = \mathcal{H}$ and there is a surjective mapping $\pi^D\colon \mathcal{D} \to \bar{\mathcal{D}}$ and an injective mapping $\iota_A\colon \bar{\mathcal{A}} \hookrightarrow \mathcal{A}$.

The mappings π^D, ι_A which make $\bar{\mathcal{P}}$ a *direct discretisation* of $\mathcal{P}$ are called *data projection* and *strategy embedding*, respectively.

Definition 2.12. A control problem $\bar{\mathcal{P}}$ is a *discretisation* of $\mathcal{P}$ iff there are control problems $\mathcal{P}_0$, $\bar{\mathcal{P}}_0$ such that $\mathcal{P}_0$ is compatible with $\mathcal{P}$, $\bar{\mathcal{P}}_0$ is compatible with $\bar{\mathcal{P}}$, and $\bar{\mathcal{P}}_0$ is a direct discretisation of $\mathcal{P}_0$.

Clearly, a control problem $\mathcal{P}$ which is a direct discretisation of some other control problem $\tilde{\mathcal{P}}$ is also a discretisation of $\tilde{\mathcal{P}}$. When we have two control problems where one is a direct discretisation of the other, then both problems must have the same system space. This is not necessarily the case when one control problem is just a discretisation of the other.

Let us see how the discrete control problems described in Section 2.1 fit into our framework. We define them in such a way that they are direct discretisations of the control problem $\hat{\mathcal{P}}_{det}$.

For $M \in \mathbb{N}$, a control problem $\hat{\mathcal{P}}_M = (\hat{\mathcal{D}}_M, \hat{\mathcal{A}}_M, \hat{\mathcal{H}}_M, \hat{\Psi}_M, \hat{J}_M)$ compatible with the control problem of degree M can be defined in the following way. Set $\hat{\mathcal{D}}_M := \{0, \dots, T_M\} \times S_M$, where S_M is a regular triangulation of $\mathbb{R}^d$ as in Section 2.1. Set $\hat{\mathcal{A}}_{det} := \mathcal{U}_{ad}^M$. Thus, admissible strategies are pairs of stochastic bases $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\in\mathbb{N}_0}, \mathbf{P})$ and $(\mathcal{F}_n)$-adapted Γ-valued sequences $(\bar{u}(n))_{n\in\mathbb{N}_0}$. The system space $\hat{\mathcal{H}}_M$ is the same as $\hat{\mathcal{H}}_{det}$. Denote by ι_{rel} the mapping induced by the representation of ordinary Γ-valued control processes as relaxed control processes according to (2.8). Define the system functional $\hat{\Psi}_M$ as the mapping

$$\hat{\mathcal{D}}_M \times \hat{\mathcal{A}}_M \ni ((n_0, y), ((\Omega, \mathcal{F}, (\mathcal{F}_n), \mathbf{P}), \bar{u}(n))) \mapsto ((\tfrac{n_0}{M}, y), \mathbf{Q}) \in \hat{\mathcal{H}}_{det},$$

where the probability measure $\mathbf{Q}$ is the distribution under $\mathbf{P}$ of the random variable

$$\Omega \ni \omega \mapsto \left(\iota_{rel}\left([0,\infty) \ni t \mapsto \bar{u}(\lfloor \tfrac{t}{M} \rfloor, \omega)\right), [0,\infty) \ni t \mapsto \xi(\lfloor \tfrac{t}{M} \rfloor, \omega)\right) \in \mathcal{R}(\Gamma) \times D([0,\infty), \mathbb{R}^d),$$

and $(\xi(n))$ is the S_M-valued $(\mathcal{F}_n)$-adapted sequence determined by the strategy $\bar{u}$, the initial condition (n_0, y) and the transition function p_M from Section 2.1. The cost functional $\hat{\mathcal{J}}_M$ is defined to be the mapping $\hat{\mathcal{H}}_{det} \to [-\infty, \infty]$ given by

$$((t_0, y), \mathbf{Q}) \mapsto \int \left(\sum_{n=0}^{T_M - n_0 - 1} \int_\Gamma f_M\big(n_0 + n, \tilde{x}(\tfrac{n}{M}), \gamma\big) d\rho(\gamma, \tfrac{n}{M}) \;+\; g_M\big(\tilde{x}(\tfrac{T_M - n_0}{M})\big) \right) d\mathbf{Q}(\rho, \tilde{x}(.)),$$

where the integral with respect to $\mathbf{Q}$ is again over $\mathcal{R}(\Gamma) \times D([0,\infty), \mathbb{R}^d)$.

In order to check whether the control problem $\hat{\mathcal{P}}_M$ just constructed is indeed a direct discretisation of $\hat{\mathcal{P}}_{det}$, it is enough to find a strategy embedding ι_A^M and a data projection π_M^D. Define ι_A^M to be the mapping

$$\hat{\mathcal{A}}_M \ni (\bar{u}_n)_{n\in\mathbb{N}_0} \mapsto \iota_{rel}\big([0,\infty) \ni t \mapsto \bar{u}(\lfloor \tfrac{t}{M} \rfloor)\big) \in \hat{\mathcal{A}}_{det},$$

that is, the sequence $(\bar{u}_n)$ is associated with the relaxed control representation of its piecewise constant interpolation relative to a grid of mesh size $\frac{1}{M}$. This is the same operation as in the definition of the system functional $\hat{\mathcal{J}}_M$. As data projection π_M^D we may choose the mapping

$$\hat{\mathcal{D}}_{det} \ni (t_0, \varphi) \mapsto (\lfloor M \cdot t_0 \rfloor, \Lambda_M(\varphi)) \in \hat{\mathcal{D}}_M,$$

where Λ_M maps $\varphi \in \mathbb{R}^d$ to its nearest neighbour in $S_M \subset \mathbb{R}^d$.

2.2.2 Approximation and convergence

Recall that our objective is to provide sufficient conditions for the convergence of the value functions associated with a family of "discrete" control problems to the value function of some given "continuous" control problem. The conditions are stated for approximating control problems which are direct discretisations of the original problem.

Definition 2.13. A sequence $(\mathcal{P}_M)_{M\in\mathbb{N}}$ of optimal control problems *approximates* an optimal control problem $\mathcal{P}$ iff $\mathcal{P}_M$ is a direct discretisation of $\mathcal{P}$ with data projection π_M^D, each $M \in \mathbb{N}$, and $(V_M \circ \pi_M^D)_{M\in\mathbb{N}}$ converges to V pointwise over $\mathcal{D}$, where V_M, V are the value functions associated with $\mathcal{P}_M$, $\mathcal{P}$, respectively.

Let $\mathcal{P} = (\mathcal{D}, \mathcal{A}, \mathcal{H}, \Psi, J)$, $\mathcal{P}_M = (\mathcal{D}_M, \mathcal{A}_M, \mathcal{H}, \Psi_M, J_M)$, $M \in \mathbb{N}$, be control problems such that, for each $M \in \mathbb{N}$, the problem $\mathcal{P}_M$ is a direct discretisation of $\mathcal{P}$ with data projection π_M^D. Note that the system space $\mathcal{H}$ is the same for all control problems involved, while the cost functional may vary depending on the discretisation degree $M \in \mathbb{N}$.

For proving convergence we will suppose that there is a topology on $\mathcal{H}$ such that

(H1) the mapping $J : \mathcal{H} \to (-\infty, \infty]$ is sequentially lower semi-continuous,

(H2) J_M tends to J as $M \to \infty$ uniformly on sequentially compact subsets of $\mathcal{H}$,

(H3) for each $\varphi \in \mathcal{D}$, each $\alpha \in \mathcal{A}$, there is a sequence $(\alpha_M)_{M\in\mathbb{N}}$ with $\alpha_M \in \mathcal{A}_M$ such that $\limsup_{M\to\infty} J \circ \Psi_M(\pi_M^D(\varphi), \alpha_M) \leq J \circ \Psi(\varphi, \alpha)$,

(H4) for each $\varphi \in \mathcal{D}$, any sequence $(\alpha_M)_{M\in\mathbb{N}}$ such that $\alpha_M \in \mathcal{A}_M$, the closure of the set $\{\Psi_M(\pi_M^D(\varphi), \alpha_M) : M \in \mathbb{N}\}$ is sequentially compact in $\mathcal{H}$,

(H5) for each $\varphi \in \mathcal{D}$, any sequence $(\alpha_M)_{M\in\mathbb{N}}$ such that $\alpha_M \in \mathcal{A}_M$, the limit points of the sequence $(\Psi_M(\pi_M^D(\varphi), \alpha_M))_{M\in\mathbb{N}}$ are contained in $\Psi(\varphi, .)(\mathcal{A})$.

The conditions just stated guarantee that the sequence of control problems $(\mathcal{P}_M)_{M\in\mathbb{N}}$ approximates $\mathcal{P}$.

Theorem 2.1. *Let $\mathcal{P}$ be a control problem and $(\mathcal{P}_M)_{M\in\mathbb{N}}$ be a sequence of direct discretisations of $\mathcal{P}$ as above. If all the control problems are finite and there is a topology on the system space $\mathcal{H}$ such that Assumptions (H1)-(H5) hold, then $(\mathcal{P}_M)$ approximates $\mathcal{P}$.*

Proof. Let $\varphi \in \mathcal{D}$. For $M \in \mathbb{N}$ set $\varphi_M := \pi_M^D(\varphi)$. We have to show that $V_M(\varphi_M) \to V(\varphi)$ as $M \to \infty$. The first step is to check that $\limsup_{M\to\infty} V_M(\varphi_M) \leq V(\varphi)$. To this end, let $\varepsilon > 0$ and choose $\alpha \in \mathcal{A}$ such that $V(\varphi) \geq J(\Psi(\varphi, \alpha)) - \varepsilon$. For this φ and this α, choose a sequence $(\alpha_M)_{M\in\mathbb{N}}$ with $\alpha_M \in \mathcal{A}_M$ according to Assumption (H3) such that $\limsup_{M\to\infty} J \circ \Psi_M(\pi_M^D(\varphi), \alpha_M) \leq J \circ \Psi(\varphi, \alpha)$. By Assumptions (H2) and (H4) we have that the difference between $J \circ \Psi_M(\varphi_M, \alpha_M)$ and $J_M \circ \Psi_M(\varphi_M, \alpha_M)$ tends to zero as $M \to \infty$. Hence we find $M_0 \in \mathbb{N}$ such that for all $M \geq M_0$

$$V_M(\varphi_M) \;\leq\; J_M\big(\Psi_M(\varphi_M, \alpha_M)\big) \;\leq\; J\big(\Psi(\varphi, \alpha)\big) + \varepsilon \;\leq\; V(\varphi) + 2\varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, it follows that $\limsup_{M\to\infty} V_M(\varphi_M) \leq V(\varphi)$.

The second step is to show that $\liminf_{M\to\infty} V_M(\varphi_M) \geq V(\varphi)$. For each $M \in \mathbb{N}$, choose $\alpha_M \in \mathcal{A}_M$ such that $V_M(\varphi_M) \geq J_M(\Psi_M(\varphi_M, \alpha_M)) - \frac{1}{M}$. Set $x_M := \Psi_M(\varphi_M, \alpha_M)$. Then, by construction,

$$\liminf_{M\to\infty} V_M(\varphi_M) \;=\; \liminf_{M\to\infty} J_M(x_M).$$

Assume we had $\liminf_{M\to\infty} J_M(x_M) < V(\varphi)$. By Assumption (H4), (x_M) would be contained in a sequentially compact set, whence we could choose a convergent subsequence $(x_{M(i)}) \subset (x_M)$ with unique limit point $x := \lim_{i\to\infty} x_{M(i)}$. By Assumption (H5), there would be a strategy $\tilde{\alpha} \in \mathcal{A}$ such that $x = \Psi(\varphi, \tilde{\alpha})$, whence $J(x) = J(\Psi(\varphi, \tilde{\alpha})) \geq V(\varphi)$. By Assumption (H1) we would have $\liminf_{i\to\infty} J(x_{M(i)}) \geq J(x)$, while Assumptions (H2) and (H4) together would imply that the difference between $J(x_{M(i)})$ and $J_{M(i)}(x_{M(i)})$ tends to zero as $i \to \infty$. This would yield

$$\liminf_{i\to\infty} J_{M(i)}(x_{M(i)}) \;=\; \liminf_{i\to\infty} J_{M(i)}(x_{M(i)}) - J(x_{M(i)}) + J(x_{M(i)}) - J(x) + J(x) \;\geq\; V(\varphi),$$

a contradiction to the hypothesis that $\liminf_{M\to\infty} J_M(x_M) < V(\varphi)$. Therefore, we must have $\liminf_{M\to\infty} J_M(x_M) \geq V(\varphi)$. It follows that $\liminf_{M\to\infty} V_M(\varphi_M) \geq V(\varphi)$. □

The conclusion of Theorem 2.1 continues to hold if we replace Assumptions (H1) and (H5) by the following hypotheses:

(H1') the mapping $J : \mathcal{H} \to (-\infty, \infty]$ is sequentially continuous,

(H5') for all $\varphi \in \mathcal{D}$, any sequence $(\alpha_M)_{M\in\mathbb{N}}$ such that $\alpha_M \in \mathcal{A}_M$, the limit points of the sequence $(\Psi_M(\pi_M^D(\varphi), \alpha_M))_{M\in\mathbb{N}}$ are contained in the closure of $\Psi(\varphi, .)(\mathcal{A})$.

Let us briefly comment on Assumptions (H1)-(H5). Hypothesis (H1) is a continuity assumption on the cost functional of the original problem only. Hypothesis (H2) states that the cost functionals of the approximating problems converge locally uniformly to the costs of the original problem. Hypothesis (H3) could be called a "scattering assumption", because it implies that any "continuous" strategy can be approximated by "discrete" strategies in the sense of converging costs. Hypothesis (H4) is about the existence of convergent subsequences of solutions to the dynamics of the approximating problems. It is usually a consequence of the compactification of the space of strategies mentioned in Section 2.1. Hypothesis (H5) says that limits of solutions to the approximating dynamics can be identified as solutions to the original dynamics.

There are two important points when it comes to applying Theorem 2.1. The first is to re-formulate the control problems involved so that the approximating problems are direct discretisations of the original problem. The second point is the choice of a suitable topology on the system space. As far as the example problem is concerned, $\hat{\mathcal{P}}_{det}$, $\hat{\mathcal{P}}_M$, $M \in \mathbb{N}$, are appropriate reformulations of the original control problem and the associated discrete problems, respectively. Also, for each $M \in \mathbb{N}$, $\hat{\mathcal{P}}_M$ is a direct discretisation of $\hat{\mathcal{P}}_{det}$. It remains to choose the topology on the system space $\hat{\mathcal{H}}_{det}$. In view of how the cost functionals are defined, the topology of weak convergence of probability measures on $\mathcal{B}(\mathcal{R}(\Gamma)\times D([0,\infty), \mathbb{R}^d))$ coupled with the standard topology on $[0,T]\times\mathbb{R}^d$ is a good choice.

We do not provide the convergence analysis for the example problem. Notice that we did not make precise any assumptions regarding the coefficients b, f, g of the original problem. In Section 2.3, however, the details of the convergence analysis for a class of stochastic optimal control problems with delay are worked out.

2.3 Application to stochastic control problems with delay

Here, we study the approximation of certain continuous-time stochastic optimal control problems with time delay in the dynamics according to the Markov chain method. The control problems whose value functions are to be approximated are specified in Subsection 2.3.1. In Subsection 2.3.2, the original control problem is reformulated by enlarging and compactifying the set of admissible strategies. For the resulting relaxed control problem, optimal strategies are guaranteed to exist. The dynamics of the approximating control problems are defined in Subsection 2.3.3; time as well as the state space are discretised, and an appropriate condition of local consistency is given. In Subsection 2.3.4, the cost functionals of the approximating problems are specified and convergence of the corresponding value functions is shown. Subsection 2.3.5 contains a technical result which is needed in the proof of Proposition 2.2.

2.3.1 The original control problem

We consider the control of a dynamical system given by a one-dimensional stochastic delay differential equation driven by a Wiener process. Both drift and diffusion coefficient may depend on the solution's history a certain amount of time into the past. Let $r > 0$ denote the *delay length*, i. e. the maximal length of dependence on the past. In order to simplify the analysis, we restrict attention to the case where only the drift term can be directly controlled.

Typically, the solution process of an SDDE does not enjoy the Markov property, while the segment process associated with that solution does, cf. Subsection 1.1.1. The segment process $(X_t)_{t\geq 0}$ associated with a real-valued càdlàg process $(X(t))_{t\geq -r}$ takes its values in $D_0 := D([-r,0])$, the space of all real-valued càdlàg functions on the interval $[-r,0]$. There are two natural topologies on D_0. The first is the one induced by the supremum norm. The second is the *Skorohod topology* of càdlàg convergence (e. g. Billingsley, 1999: Ch. 3). The main difference between the Skorohod and the uniform topology lies in the different evaluation of convergence of functions with jumps, which appear naturally as initial segments and discretised processes. For continuous functions both topologies coincide. Similar statements hold for $D_\infty := D([-r,\infty))$ and $\tilde{D}_\infty := D([0,\infty))$, the spaces of all real-valued càdlàg functions on the intervals $[-r,\infty)$ and $[0,\infty)$, respectively. The spaces D_∞ and $\tilde{D}_\infty$ will always be supposed to carry the Skorohod topology, while D_0 will canonically be equipped with the uniform topology.

Let (Γ, d_Γ) be a compact metric space, the space of *control actions*. Denote by b the drift coefficient of the controlled dynamics, and by σ the diffusion coefficient. Let $(W(t))_{t\geq 0}$ be a one-dimensional standard Wiener process on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbf{P})$ satisfying the usual conditions, and let $(u(t))_{t\geq 0}$ be a *control process*, i. e. an $(\mathcal{F}_t)$-adapted measurable process with values in Γ. Consider the controlled SDDE

$$dX(t) \;=\; b\big(X_t, u(t)\big)\, dt \;+\; \sigma(X_t)\, dW(t), \qquad t \geq 0. \tag{2.12}$$

The control process $u(.)$ together with its stochastic basis including the Wiener process is called an *admissible strategy* if, for every deterministic initial condition $\varphi \in D_0$, Equation (2.12) has a unique solution which is also weakly unique. Write $\mathcal{U}_{ad}$ for the set of admissible strategies of Equation (2.12). The stochastic basis coming with an admissible control will often be omitted in the notation.

A *solution* in the sense used here is an adapted càdlàg process defined on the stochastic basis of the control process such that the integral version of Equation (2.12) is satisfied. Given a control process together with a standard Wiener process, a solution to Equation (2.12) is *unique* if it is indistinguishable from any other solution almost surely satisfying the same initial condition. A solution is *weakly unique* if it has the same law as any other solution with the same initial distribution and satisfying Equation (2.12) for a control process on a possibly different stochastic basis so that the joint distributions of control and driving Wiener process are the same for both solutions. Let us specify the regularity assumptions to be imposed on the coefficients b and σ:

(A1) Càdlàg functionals: the mappings

$$(\psi,\gamma) \mapsto \big[t \mapsto b(\psi_t,\gamma),\ t \geq 0\big], \qquad\qquad \psi \mapsto \big[t \mapsto \sigma(\psi_t),\ t \geq 0\big]$$

define measurable functionals $D_\infty \times \Gamma \to \tilde{D}_\infty$ and $D_\infty \to \tilde{D}_\infty$, respectively, where D_∞, $\tilde{D}_\infty$ are equipped with their Borel σ-algebras.

(A2) Continuity of the drift coefficient: there is an at most countable subset of $[-r,0]$, denoted by I_{ev}, such that for every $t \geq 0$ the function defined by

$$D_\infty \times \Gamma \ni (\psi,\gamma) \mapsto b(\psi_t,\gamma)$$

is continuous on $D_{ev}(t) \times \Gamma$ uniformly in $\gamma \in \Gamma$, where

$$D_{ev}(t) := \ \{\psi \in D_\infty \mid \psi \text{ is continuous at } t+s \text{ for all } s \in I_{ev}\}.$$

(A3) Global boundedness: $|b|$, $|\sigma|$ are bounded by a constant $K > 0$.

(A4) Uniform Lipschitz condition: There is a constant $K_L > 0$ such that for all $\varphi, \tilde{\varphi} \in D_0$, all $\gamma \in \Gamma$

$$|b(\varphi,\gamma) - b(\tilde{\varphi},\gamma)| \ + \ |\sigma(\varphi) - \sigma(\tilde{\varphi})| \ \leq \ K_L \cdot \sup_{s\in[-r,0]} |\varphi(s) - \tilde{\varphi}(s)|.$$

(A5) Ellipticity of the diffusion coefficient: $\sigma(\varphi) \geq \sigma_0$ for all $\varphi \in D_0$, where $\sigma_0 > 0$ is a positive constant.

Assumptions (A1) and (A4) on the coefficients allow us to invoke Theorem V.7 in Protter (2003: p.253), which guarantees the existence of a unique solution to Equation (2.12) for every piecewise constant control attaining only finitely many different values. The boundedness Assumption (A3) poses no limitation except for the initial conditions, because the state evolution will be stopped when the state process leaves a bounded interval. Assumption (A2) allows us to use "segmentwise approximations" of the solution process, see the proof of Proposition 2.1. The assumptions imposed on the drift coefficient b are satisfied, for example, by

$$b(\varphi,\gamma) := \ f\Big(\varphi(r_1),\dots,\varphi(r_n), \int_{-r}^{0} \varphi(s)w_1(s)ds, \dots, \int_{-r}^{0} \varphi(s)w_m(s)ds\Big) \cdot g(\gamma), \tag{2.13}$$

where $r_1,\dots,r_n \in [-r,0]$ are fixed, f, g are bounded continuous functions and f is Lipschitz, and the weight functions $w_1,\dots,w_m$ lie in $L^1([-r,0])$. Apart from the control term, the diffusion coefficient σ may have the same structure as b in (2.13).

We next give an example of a function that could be taken for σ if the càdlàg continuity in Assumption (A1) were missing. In Subsection 2.3.3 it will become clear that the corresponding control problem cannot be approximated by a simple discretisation procedure, because the evaluation of $\sigma(\varphi)$ for any $\varphi \in D_0$ depends on the discretisation grid. Let A_M be the subset of the interval $[-r,0]$ given by

$$A_M := \ \big\{(t-2^{-3M},t] \ \big|\ t = r(\tfrac{n}{2^M}-1) \text{ for some } n \in \{1,\dots,2^M\}\big\}.$$

Let A be the union of the sets A_M, $M \in \mathbb{N}$. With positive constants σ_0, K, we define a functional $\sigma : D_0 \to \mathbb{R}$ by

$$\sigma(\varphi) := \sigma_0 + K \wedge \sup\{|\varphi(t) - \varphi(t-)| \mid t \in A\}, \tag{2.14}$$

where $\varphi(t-)$ is the left hand limit of φ at $t \in [-r, 0]$. Assumptions (A3) and (A4) are clearly satisfied if we choose σ according to (2.14), but σ would not induce a càdlàg functional $D_\infty \to \tilde{D}_\infty$. This can be seen by considering the mapping $[0, \infty) \ni t \mapsto \sigma(\psi_t)$ for a function $\psi \in D_\infty$ which is constant except for a single discontinuity. If we had defined σ with the set A being the union of only finitely many sets A_M, then we would have obtained a càdlàg functional.

We consider control problems in the weak formulation (cf. Yong and Zhou, 1999: p. 64). Given an admissible control $u(.)$ and a deterministic initial segment $\varphi \in D_0$, denote by $X^{\varphi,u}$ the unique solution to Equation (2.12). Let I be a compact interval with non-empty interior. Define the stopping time $\tau^{\bar{T}}_{\varphi,u}$ of first exit from the interior of I before time $\bar{T} > 0$ by

$$\tau^{\bar{T}}_{\varphi,u} := \inf\{t \geq 0 \mid X^{\varphi,u}(t) \notin \mathrm{int}(I)\} \wedge \bar{T}. \tag{2.15}$$

In order to define the costs, we prescribe a *cost rate* $k \colon \mathbb{R} \times \Gamma \to [0, \infty)$ and a *boundary cost* $g \colon \mathbb{R} \to [0, \infty)$ which we take to be (jointly) continuous bounded functions. Let $\beta \geq 0$ denote the exponential *discount rate*. Then define the *cost functional* on $D_0 \times \mathcal{U}_{ad}$ by

$$J(\varphi, u) := \mathbf{E}\left(\int_0^\tau \exp(-\beta s) \cdot k\big(X^{\varphi,u}(s), u(s)\big)\, ds + g\big(X^{\varphi,u}(\tau)\big)\right), \tag{2.16}$$

where $\tau = \tau^{\bar{T}}_{\varphi,u}$. Our aim is to minimize $J(\varphi, .)$. We introduce the *value function*

$$V(\varphi) := \inf\{J(\varphi, u) \mid u \in \mathcal{U}_{ad}\}, \quad \varphi \in D_0. \tag{2.17}$$

The control problem now consists in calculating the function V and finding admissible controls that minimize J. Such control processes are called *optimal controls* or *optimal strategies*.

2.3.2 Existence of optimal strategies

In the class $\mathcal{U}_{ad}$ of admissible strategies it may happen that there is no optimal control. A way out is to enlarge the class of strategies, allowing for so-called relaxed controls, cf. Subsection 2.3.1 and the discussion after Definition 2.8 in Subsection 2.2.1.

Let R be a relaxed control process in the sense of Definition 2.2. Then Equation (2.12) takes on the form

$$dX(t) = \Big(\int_\Gamma b(X_t, \gamma)\, \dot{R}(t, d\gamma)\Big) dt + \sigma(X_t)\, dW(t), \qquad t \geq 0, \tag{2.18}$$

where $(\dot{R}(t, .))_{t \geq 0}$ is the family of derivative measures associated with R. A relaxed control process together with its stochastic basis including the Wiener process is called *admissible relaxed control* or an *admissible strategy* if, for every deterministic initial condition, Equation (2.18) has a unique solution which is also weakly unique. Denote by $\hat{\mathcal{U}}_{ad}$ the set of all admissible relaxed controls. Instead of (2.16) we define a cost functional on $D_0 \times \hat{\mathcal{U}}_{ad}$ by

$$\hat{J}(\varphi, R) := \mathbf{E}\left(\int_0^\tau \int_\Gamma \exp(-\beta s) \cdot k\big(X^{\varphi,R}(s), \gamma\big)\, \dot{R}(s, d\gamma)\, ds + g\big(X^{\varphi,R}(\tau)\big)\right), \tag{2.19}$$

where $X^{\varphi,R}$ is the solution to Equation (2.18) under the relaxed control process R with initial segment φ and τ is defined in analogy to (2.15). Instead of (2.17) as value function we have

$$\hat{V}(\varphi) := \inf\{\hat{J}(\varphi, R) \mid R \in \hat{\mathcal{U}}_{ad}\}, \quad \varphi \in D_0. \tag{2.20}$$

The cost functional $\hat{J}$ depends only on the joint distribution of the solution $X^{\varphi,R}$ and the underlying control process R, since τ, the time horizon, is a deterministic function of the solution. The distribution of $X^{\varphi,R}$, in turn, is determined by the initial condition φ and the joint distribution of the control process and its accompanying Wiener process. Letting the time horizon vary, we may regard $\hat{J}$ as a function of the law of (X, R, W, τ), that is, as being defined on a subset of the set of probability measures on $\mathcal{B}(D_\infty \times \mathcal{R} \times \tilde{D}_\infty \times [0,\infty])$. Notice that the time interval has been compactified. The domain of definition of $\hat{J}$ is determined by the class of admissible relaxed controls for Equation (2.18), the definition of the time horizon and the distributions of the initial segments X_0.

The following proposition gives the analogue of Theorem 10.1.1 in Kushner and Dupuis (2001: pp. 271-275) for our setting. We present the proof in detail, because the identification of the limit process is different from the classical case.

Proposition 2.1. *Assume (A1) – (A4). Let $((R^M, W^M))_{M\in\mathbb{N}}$ be any sequence of admissible relaxed controls for Equation* (2.18)*, where (R^M, W^M) is defined on the filtered probability space $(\Omega_M, \mathcal{F}^M, (\mathcal{F}^M_t), \mathbf{P}_M)$. Let X^M be a solution to Equation* (2.18) *under control (R^M, W^M) with deterministic initial condition $\varphi^M \in D_0$, and assume that (φ^M) tends to φ uniformly for some $\varphi \in D_0$. For each $M \in \mathbb{N}$, let τ^M be an $(\mathcal{F}^M_t)$-stopping time. Then $((X^M, R^M, W^M, \tau^M))_{M\in\mathbb{N}}$ is tight.*

Denote by (X, R, W, τ) a limit point of the sequence $((X^M, R^M, W^M, \tau^M))_{M\in\mathbb{N}}$. Define a filtration by $\mathcal{F}_t := \sigma(X(s), R(s), W(s), \tau\mathbf{1}_{\{\tau\leq t\}},\ s \leq t)$, $t \geq 0$. Then $W(.)$ is an $(\mathcal{F}_t)$-adapted Wiener process, τ is an $(\mathcal{F}_t)$-stopping time, (R, W) is an admissible relaxed control, and X is a solution to Equation (2.18) *under (R, W) with initial condition φ.*

Proof. Tightness of (X^M) follows from the Aldous criterion (cf. Billingsley, 1999: pp. 176-179): given $M \in \mathbb{N}$, any bounded $(\mathcal{F}^M_t)$-stopping time ν and $\delta > 0$ we have

$$\mathbf{E}_M\left(\left|X^M(\nu+\delta) - X^M(\nu)\right|^2 \mid \mathcal{F}^M_\nu\right) \quad \leq \quad 2K^2\delta(\delta+1)$$

as a consequence of Assumption (A3) and the Itô isometry. Notice that $X^M(0)$ tends to $X(0)$ as M goes to infinity by hypothesis. The sequences (R^M) and (τ^M) are tight, because the value spaces $\mathcal{R}$ and $[0,\infty]$, respectively, are compact. The sequence (W^M) is tight, since all W^M induce the same measure. Finally, componentwise tightness implies tightness of the product (cf. Billingsley, 1999: p. 65).

By abuse of notation, we do not distinguish between the convergent subsequence and the original sequence and assume that $((X^M, R^M, W^M, \tau^M))$ converges weakly to (X, R, W, τ). The random time τ is an $(\mathcal{F}_t)$-stopping time by construction of the filtration. Likewise, R is $(\mathcal{F}_t)$-adapted by construction, and it is indeed a relaxed control process, because $R(\Gamma \times [0,t]) = t$, $t \geq 0$, $\mathbf{P}$-almost surely by weak convergence of the relaxed control processes (R^M) to R. The process W has Wiener distribution and continuous paths with probability one, being the limit of standard Wiener processes. To check that W is an

$(\mathcal{F}_t)$-Wiener process, we use the martingale problem characterization of Brownian motion. To this end, for $g \in \mathbf{C}_c(\Gamma \times [0,\infty))$, $\rho \in \mathcal{R}$ define the pairing

$$(g,\rho)(t) := \int_{\Gamma\times[0,t]} g(\gamma,s)\, d\rho(\gamma,s), \quad t \geq 0.$$

Notice that real-valued continuous functions on $\mathcal{R}$ can be approximated by functions of the form

$$\mathcal{R} \ni \rho \mapsto \tilde{H}\big((g_j,\rho)(t_i),\, (i,j) \in \mathbb{N}_p \times \mathbb{N}_q\big) \in \mathbb{R},$$

where p, q are natural numbers, $\{t_i \mid i \in \mathbb{N}_p\} \subset [0,\infty)$, and $\tilde{H}$, g_j, $j \in \mathbb{N}_q$, are suitable continuous functions with compact support and $\mathbb{N}_N := \{1,\dots,N\}$ for any $N \in \mathbb{N}$. Let $t \geq 0$, $t_1,\dots,t_p \in [0,t]$, $h \geq 0$, $g_1,\dots,g_q$ be functions in $\mathbf{C}_c(\Gamma \times [0,\infty))$, and H be a continuous function of $2p + p\cdot q + 1$ arguments with compact support. Since W^M is an $(\mathcal{F}^M_t)$-Wiener process for each $M \in \mathbb{N}$, we have for all $f \in \mathbf{C}^2_c(\mathbb{R})$

$$\begin{aligned}\mathbf{E}_M\Big(&H\big(X^M(t_i),(g_j,R^M)(t_i),W^M(t_i),\tau^M\mathbf{1}_{\{\tau^M\leq t\}},\,(i,j)\in\mathbb{N}_p\times\mathbb{N}_q\big)\\ &\cdot\Big(f\big(W^M(t+h)\big)-f\big(W^M(t)\big)-\frac{1}{2}\int_t^{t+h}\frac{\partial^2 f}{\partial x^2}\big(W^M(s)\big)ds\Big)\Big) \quad = \quad 0.\end{aligned}$$

By the weak convergence of $((X^M,R^M,W^M,\tau^M))_{M\in\mathbb{N}}$ to (X,W,R,τ) we see that

$$\begin{aligned}\mathbf{E}\Big(&H\big(X(t_i),(g_j,R)(t_i),W(t_i),\tau\mathbf{1}_{\{\tau\leq t\}},\,(i,j)\in\mathbb{N}_p\times\mathbb{N}_q\big)\\ &\cdot\Big(f\big(W(t+h)\big)-f\big(W(t)\big)-\frac{1}{2}\int_t^{t+h}\frac{\partial^2 f}{\partial x^2}\big(W(s)\big)ds\Big)\Big) \quad = \quad 0\end{aligned}$$

for all $f \in \mathbf{C}^2_c(\mathbb{R})$. As H, p, q, t_i, g_j vary over all possibilities, the corresponding random variables $H(X(t_i),(g_j,R)(t_i),W(t_i),\tau\mathbf{1}_{\{\tau\leq t\}},\,(i,j) \in \mathbb{N}_p \times \mathbb{N}_q)$ induce the σ-algebra $\mathcal{F}_t$. Since $t \geq 0$, $h \geq 0$ were arbitrary, it follows that

$$f\big(W(t)\big) - f\big(W(0)\big) - \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}\big(W(s)\big)ds, \quad t \geq 0,$$

is an $(\mathcal{F}_t)$-martingale for every $f \in \mathbf{C}^2_c(\mathbb{R})$. Consequently, W is an $(\mathcal{F}_t)$-Wiener process.

It remains to show that X solves Equation (2.18) under control (R,W) with initial condition φ. Notice that X has continuous paths on $[0,\infty)$ $\mathbf{P}$-almost surely, because the process $(X(t))_{t\geq 0}$ is the weak limit in $\tilde{D}_\infty$ of continuous processes. Fix $T > 0$. We have to check that, $\mathbf{P}$-almost surely,

$$X(t) = \varphi(0) \; + \int_0^t \int_\Gamma b(X_s,\gamma)\, \dot{R}(s,d\gamma)\, ds + \int_0^t \sigma(X_s)\, dW(s) \quad \text{for all } t \in [0,T].$$

By virtue of the Skorohod representation theorem (cf. Billingsley, 1999: p. 70) we may assume that the processes (X^M,R^M,W^M), $M \in \mathbb{N}$, are all defined on the same probability space $(\Omega,\mathcal{F},\mathbf{P})$ as (X,R,W) and that convergence of $((X^M,R^M,W^M))$ to (X,R,W) is

$\mathbf{P}$-almost sure. Since X, W have continuous paths on $[0,T]$ and (φ^M) converges to φ in the uniform topology, one finds $\tilde{\Omega} \in \mathcal{F}$ with $\mathbf{P}(\tilde{\Omega}) = 1$ such that for all $\omega \in \tilde{\Omega}$

$$\sup_{t\in[-r,T]} \left|X^M(t)(\omega) - X(t)(\omega)\right| \overset{M\to\infty}{\longrightarrow} 0, \qquad \sup_{t\in[-r,T]} \left|W^M(t)(\omega) - W(t)(\omega)\right| \overset{M\to\infty}{\longrightarrow} 0,$$

and also $R^M(\omega) \to R(\omega)$ in $\mathcal{R}$. Let $\omega \in \tilde{\Omega}$. We first show that

$$\int_0^t \int_\Gamma b\big(X_s^M(\omega),\gamma\big)\, \dot{R}^M(s,d\gamma)(\omega)\, ds \overset{M\to\infty}{\longrightarrow} \int_0^t \int_\Gamma b\big(X_s(\omega),\gamma\big)\, \dot{R}(s,d\gamma)(\omega)\, ds$$

uniformly in $t \in [0,T]$. As a consequence of Assumption (A4), the uniform convergence of the trajectories on $[-r,T]$ and property (2.5) of the relaxed controls, we have

$$\int_{\Gamma\times[0,T]} \left|b\big(X_s^M(\omega),\gamma\big) - b\big(X_s(\omega),\gamma\big)\right| dR^M(\gamma,s)(\omega) \overset{M\to\infty}{\to} 0.$$

By Assumption (A2), we find a countable set $A_\omega \subset [0,T]$ such that the mapping $(\gamma,s) \mapsto b(X_s(\omega),\gamma)$ is continuous in all $(\gamma,s) \in \Gamma \times ([0,T] \setminus A_\omega)$. Since A_ω is countable we have $R(\omega)(\Gamma\times A_\omega) = 0$. Hence, by the generalized mapping theorem (cf. Billingsley, 1999: p. 21), we obtain for each $t \in [0,T]$

$$\int_{\Gamma\times[0,t]} b\big(X_s(\omega),\gamma\big)\, dR^M(\gamma,s)(\omega) \overset{M\to\infty}{\to} \int_{\Gamma\times[0,t]} b\big(X_s(\omega),\gamma\big)\, dR(\gamma,s)(\omega).$$

The convergence is again uniform in $t \in [0,T]$, as b is bounded and R^M, $M \in \mathbb{N}$, R are all positive measures with mass T on $\Gamma \times [0,T]$. Define càdlàg processes C^M, $M \in \mathbb{N}$, on $[0,\infty)$ by

$$C^M(t) := \varphi^M(0) + \int_{\Gamma\times[0,t]} b(X_s^M,\gamma)\, dR^M(\gamma,s), \quad t \geq 0,$$

and define C in analogy to C^M with φ, R, X in place of φ^M, R^M, X^M, respectively. From the above, we know that $C^M(t) \to C(t)$ holds uniformly over $t \in [0,T]$ for any $T > 0$ with probability one. Define operators $F^M : \tilde{D}_\infty \to \tilde{D}_\infty$, $M \in \mathbb{N}$, mapping càdlàg processes to càdlàg processes by

$$F^M(Y)(t)(\omega) := \sigma\left([-r,0] \ni s \mapsto \begin{cases} Y(t+s)(\omega) & \text{if } t+s \geq 0, \\ \varphi^M(t+s) & \text{else} \end{cases}\right), \quad t \geq 0,\ \omega \in \Omega,$$

and define F in the same way as F^M with φ^M replaced by φ. Observe that X^M solves

$$X^M(t) = C^M(t) + \int_0^t F^M(X^M)(s-)\, dW^M(s), \quad t \geq 0.$$

Denote by $(\hat{X}(t))_{t\geq 0}$ the unique solution to

$$\hat{X}(t) = C(t) + \int_0^t F(\hat{X})(s-)\, dW(s), \quad t \geq 0,$$

and set $\hat{X}(t) := \varphi(t)$ for $t \in [-r,0)$. Assumption (A4) and the uniform convergence of (φ^M) to φ imply that $F^M(\hat{X})$ converges to $F(\hat{X})$ uniformly on compacts in probability (convergence *in ucp*). Theorem V.15 in Protter (2003: p. 265) yields that (X^M) converges to $\hat{X}$ in ucp, that is

$$\sup_{t\in[0,T]} \left|X^M(t) - \hat{X}(t)\right| \overset{M\to\infty}{\longrightarrow} 0 \quad \text{in probability } \mathbf{P} \text{ for any } T > 0.$$

Therefore, X is indistinguishable from $\hat{X}$. By definition of C and F, this implies that $\hat{X}$ solves Equation (2.18) under control (R, W) with initial condition φ, and so does X. □

If the time horizon were deterministic, then the existence of optimal strategies in the class of relaxed controls would be clear. Given an initial condition $\varphi \in D_0$, one would select a sequence $((R^M, W^M))_{M\in\mathbb{N}}$ such that $(\hat{J}(\varphi, R^M))$ converges to its infimum. By Proposition 2.1, a suitable subsequence of $((R^M, W^M))$ and the associated solution processes would converge weakly to (R, W) and the associated solution to Equation (2.18). Taking into account (2.19), the definition of the costs, this in turn would imply that $\hat{J}(\varphi, .)$ attains its minimum value at R or, more precisely, at (X, R, W).

A similar argument is still valid if the time horizon depends continuously on the paths with probability one under every possible solution. That is to say, the mapping

$$\hat{\tau} : D_\infty \to [0, \infty], \qquad \hat{\tau}(\psi) := \inf\{t \geq 0 \mid \psi(t) \notin \text{int}(I)\} \wedge \bar{T}, \tag{2.21}$$

is Skorohod continuous with probability one under the measure induced by any solution $X^{\varphi,R}$, R any relaxed control. This is indeed the case if the diffusion coefficient σ is bounded away from zero as required by Assumption (A5), cf. Kushner and Dupuis (2001: pp. 277-281).

By introducing relaxed controls, we have enlarged the class of possible strategies. The infimum of the costs, however, remains the same for the new class. This is a consequence of the fact that stochastic relaxed controls can be arbitrarily well approximated by piecewise constant ordinary stochastic controls which attain only a finite number of different control values. A proof of this assertion is given in Kushner (1990: pp. 59-60) in case the time horizon is finite, and extended to the case of control up to an exit time in Kushner and Dupuis (2001: pp. 282-286). Notice that nothing hinges on the presence or absence of delay in the controlled dynamics. Let us summarize our findings.

Theorem 2.2. *Assume (A1) – (A5). Given any deterministic initial condition $\varphi \in D_0$, the relaxed control problem determined by* (2.18) *and* (2.19) *possesses an optimal strategy, and the minimal costs are the same as for the original control problem.*

When reformulated along the lines of the example problem in Section 2.2, the relaxed control problem determined by (2.18) and (2.19) is indeed a relaxation in the sense of Definition 2.8 of the original control problem from Subsection 2.3.1.

2.3.3 Approximating chains

In order to construct finite-dimensional approximations to our control problem, we discretise time and state space. In the non-delay case a random time grid permits simpler proofs. Since in the delay case the segment process must be well approximated, a deterministic grid is natural and preferable, but calls for proof techniques deviating from the classical way adopted by Kushner and Dupuis (2001) or Kushner (2005).

Denote by $h > 0$ the mesh size of an equidistant time discretisation starting at zero. Let $S_h := \sqrt{h}\mathbb{Z}$ be the corresponding state space, and set $I_h := I \cap S_h$. Notice that S_h is countable and I_h is finite. Let $\Lambda_h : \mathbb{R} \to S_h$ be a round-off function. We will simplify things even further by considering only mesh sizes $h = \frac{r}{M}$ for some $M \in \mathbb{N}$, where r is the delay length. The number M will be referred to as *discretisation degree.*

The admissible strategies for the finite-dimensional control problems correspond to piecewise constant processes in continuous time. A discrete-time process $u = (u(n))_{n\in\mathbb{N}_0}$ on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbf{P})$ with values in Γ is a *discrete admissible control of degree M* if u takes on only finitely many different values in Γ and $u(n)$ is $\mathcal{F}_{nh}$-measurable for all $n \in \mathbb{N}_0$. Denote by $(\bar{u}(t))_{t\geq 0}$ the piecewise constant càdlàg interpolation to u on the time grid. We call a discrete-time process $(\xi(n))_{n\in\{-M,\dots,0\}\cup\mathbb{N}}$ a *discrete chain of degree M* if $(\xi(n))$ takes its values in S_h and $\xi(n)$ is $\mathcal{F}_{nh}$-measurable for all $n \in \mathbb{N}_0$. In analogy to $\bar{u}$, write $(\bar{\xi}(t))_{t\geq -r}$ for the càdlàg interpolation to the discrete chain $(\xi(n))_{n\in\{-M,\dots,0\}\cup\mathbb{N}}$. We denote by $\bar{\xi}_t$ the D_0-valued segment of $\bar{\xi}(.)$ at time $t \geq 0$.

Let $\varphi \in D_0$ be a deterministic initial condition, and suppose we are given a *sequence of discrete admissible controls* $(u^M)_{M\in\mathbb{N}}$, that is u^M is a discrete admissible control of degree M on a stochastic basis $(\Omega_M, \mathcal{F}^M, (\mathcal{F}^M_t), \mathbf{P}_M)$ for each $M \in \mathbb{N}$. In addition, suppose that the sequence $(\bar{u}^M)$ of interpolated discrete controls converges weakly to some relaxed control R. We are then looking for a sequence approximating the solution X of Equation (2.18) under control (R, W) with initial condition φ, where the Wiener process W has to be constructed from the approximating sequence.

Given M-step or *extended Markov transition functions* $p^M : S_h^{M+1} \times \Gamma \times S_h \to [0, 1]$, $M \in \mathbb{N}$, we define a *sequence of approximating chains* associated with φ and (u^M) as a family $(\xi^M)_{M\in\mathbb{N}}$ of processes such that ξ^M is a discrete chain of degree M defined on the same stochastic basis as u^M, provided the following conditions are fulfilled for $h = h_M := \frac{r}{M}$ tending to zero:

(i) Initial condition: $\xi^M(n) = \Lambda_h(\varphi(nh))$ for all $n \in \{-M, \dots, 0\}$.

(ii) Extended Markov property: for all $n \in \mathbb{N}_0$, all $x \in S_h$

$$\mathbf{P}_M\big(\xi^M(n+1) = x \,\big|\, \mathcal{F}^M_{nh}\big) \quad = \quad p^M\big(\xi^M(n-M), \dots, \xi^M(n), u^M(n), x\big).$$

(iii) Local consistency with the drift coefficient:

$$\begin{aligned}\mu_{\xi^M}(n) \;&:=\; \mathbf{E}_M\big(\xi^M(n+1) - \xi^M(n) \,\big|\, \mathcal{F}^M_{nh}\big)\\ &=\; h\cdot b\big(\bar{\xi}^M_{nh}, u^M(n)\big) + o(h) \quad =:\; h\cdot b_h\big(\bar{\xi}^M_{nh}, u^M(n)\big).\end{aligned}$$

(iv) Local consistency with the diffusion coefficient:

$$\mathbf{E}_M\Big(\big(\xi^M(n+1) - \xi^M(n) - \mu_{\xi^M}(n)\big)^2\Big|\, \mathcal{F}^M_{nh}\Big) \;=\; h\cdot\sigma^2(\bar{\xi}^M_{nh}) + o(h) \;=:\, h\cdot\sigma_h^2(\bar{\xi}^M_{nh}).$$

(v) Jump heights: there is a positive number $\tilde{N}$, independent of M, such that

$$\sup_n |\xi^M(n+1) - \xi^M(n)| \leq \tilde{N}\sqrt{h_M}.$$

It is straightforward, under Assumptions (A3) and (A5), to construct a sequence of extended Markov transition functions such that the jump height and the local consistency conditions can be fulfilled. Assuming that the bounding constant K from (A3) is a natural number, we may define the functions p^M for all $M \in \mathbb{N}$ big enough by, for example,

$$p^M(Z(-M),\dots,Z(0),\gamma,x) := \begin{cases} \frac{1}{2K^2}\sigma(\bar{Z}) + \frac{\sqrt{h}}{2K}b(\bar{Z},\gamma), & \text{if } x = Z(0)+K\sqrt{h},\\ \frac{1}{2K^2}\sigma(\bar{Z}) - \frac{\sqrt{h}}{2K}b(\bar{Z},\gamma), & \text{if } x = Z(0)-K\sqrt{h},\\ 1-\frac{1}{K^2}\sigma(\bar{Z}) & \text{if } x = Z(0)\\ 0 & \text{else,} \end{cases}$$

where $h = h_M$, $Z = (Z(-M),\dots,Z(0)) \in S_h^{M+1}$, $\gamma \in \Gamma$, $x \in S_h$, and $\bar{Z} \in D_0$ is the piecewise constant interpolation associated with Z. The family (p^M) as just defined, in turn, is all we need in order to construct a sequence of approximating chains associated with any given φ, (u^M).

We will represent the interpolation $\bar{\xi}^M$ as a solution to an equation corresponding to Equation (2.12) with control process $\bar{u}^M$ and initial condition φ^M, where φ^M is the piecewise constant S_h-valued càdlàg interpolation to φ, that is $\varphi^M = \bar{\xi}_0^M$. Define the discrete process $(L^M(n))_{n\in\mathbb{N}_0}$ by $L^M(0) := 0$ and

$$\xi^M(n) \quad = \quad \varphi^M(0) \;+\; \sum_{i=0}^{n-1} h\cdot b_h\big(\bar{\xi}_{ih}^M, u^M(i)\big) \;+\; L^M(n), \qquad n \in \mathbb{N}.$$

Observe that L^M is a martingale in discrete time with respect to the filtration $(\mathcal{F}_{nh}^M)$. Setting

$$\varepsilon_1^M(t) := \sum_{i=0}^{\lfloor \frac{t}{h}\rfloor -1} h\cdot b_h\big(\bar{\xi}_{ih}^M, \bar{u}^M(ih)\big) \;-\; \int_0^t b\big(\bar{\xi}_s^M, \bar{u}^M(s)\big)\,ds, \qquad t \geq 0,$$

the interpolated process $\bar{\xi}^M$ can be represented as solution to

$$\bar{\xi}^M(t) \quad = \quad \varphi^M(0) \;+\; \int_0^t b\big(\bar{\xi}_s^M, \bar{u}^M(s)\big)\,ds \;+\; L^M(\lfloor \tfrac{t}{h}\rfloor) \;+\; \varepsilon_1^M(t), \qquad t \geq 0.$$

With $T > 0$, we have for the error term

$$\begin{aligned} \mathbf{E}_M\Big(\sup_{t\in[0,T]} \big|\varepsilon_1^M(t)\big|\Big) \quad &\leq \quad \sum_{i=0}^{\lfloor \frac{T}{h}\rfloor -1} h\,\mathbf{E}_M\Big(\big|b_h\big(\bar{\xi}_{ih}^M, u^M(i)\big) - b\big(\bar{\xi}_{ih}^M, u^M(i)\big)\big|\Big) \;+\; K\cdot h \\ &\quad + \int_0^{h\lfloor \frac{T}{h}\rfloor} \mathbf{E}_M\Big(\big|b\big(\bar{\xi}_{h\lfloor \frac{s}{h}\rfloor}^M, \bar{u}^M(s)\big) - b\big(\bar{\xi}_s^M, \bar{u}^M(s)\big)\big|\Big)\,ds, \end{aligned}$$

which tends to zero as M goes to infinity by Assumptions (A2), (A3), dominated convergence and the defining properties of (ξ^M). Moreover, $|\varepsilon_1^M(t)|$ is bounded by $2K\cdot T$ for all $t \in [0,T]$ and all M big enough, whence also

$$\mathbf{E}_M\Big(\sup_{t\in[0,T]}\big|\varepsilon_1^M(t)\big|^2\Big) \quad \overset{M\to\infty}{\longrightarrow} \quad 0.$$

The discrete-time martingale L^M can be rewritten as a discrete stochastic integral. Define $(W^M(n))_{n\in\mathbb{N}_0}$ by setting $W^M(0) := 0$ and

$$W^M(n) := \sum_{i=0}^{n-1} \frac{1}{\sigma(\bar{\xi}^M_{ih})}\big(L^M(i{+}1) - L^M(i)\big), \qquad n \in \mathbb{N}.$$

Using the piecewise constant interpolation $\bar{W}^M$ of W^M, the process $\bar{\xi}^M$ can be expressed as the solution to

(2.22)

$$\bar{\xi}^M(t) \;=\; \varphi^M(0) \;+\; \int_0^t b\big(\bar{\xi}^M_s, \bar{u}^M(s)\big)\,ds \;+\; \int_0^t \sigma\Big(\bar{\xi}^M_{h\lfloor\frac{s-}{h}\rfloor}\Big)\,d\bar{W}^M(s) \;+\; \varepsilon^M_2(t), \quad t \geq 0,$$

where the error terms (ε^M_2) converge to zero as (ε^M_1) before.

We are now prepared for the convergence result, which should be compared to Theorem 10.4.1 in Kushner and Dupuis (2001: p. 290). The proof is similar to that of Proposition 2.1. We merely point out the main differences.

Proposition 2.2. *Assume (A1) – (A5). For each $M \in \mathbb{N}$, let τ^M be a stopping time with respect to the filtration generated by $(\bar{\xi}^M(s), \bar{u}^M(s), \bar{W}^M(s),\ s \leq t)$. Let R^M denote the relaxed control representation of $\bar{u}^M$. Suppose (φ^M) converges to the initial condition φ uniformly on $[-r, 0]$. Then $((\bar{\xi}^M, R^M, \bar{W}^M, \tau^M))_{M\in\mathbb{N}}$ is tight.*

For a limit point (X, R, W, τ) set $\mathcal{F}_t := \sigma\big(X(s), R(s), W(s), \tau\mathbf{1}_{\{\tau\leq t\}},\ s \leq t\big)$, $t \geq 0$. Then W is an $(\mathcal{F}_t)$-adapted Wiener process, τ is an $(\mathcal{F}_t)$-stopping time, (R, W) is an admissible relaxed control, and X is a solution to Equation (2.18) *under* (R, W) *with initial condition φ.*

Proof. The main differences in the proof are establishing the tightness of $(\bar{W}^M)$ and the identification of the limit points. We calculate the order of convergence for the discrete-time previsible quadratic variations of (W^M):

$$\langle W^M\rangle_n \;=\; \sum_{i=0}^{n-1}\mathbf{E}\big((W^M(i{+}1) - W^M(i))^2 \mid \mathcal{F}^M_{ih}\big) \;=\; nh \;+\; o(h)\sum_{i=0}^{n-1}\frac{1}{\sigma^2(\tilde{\xi}^M_{ih})}$$

for all $M \in \mathbb{N}$, $n \in \mathbb{N}_0$. Taking into account Assumption (A5) and the definition of the time-continuous processes $\bar{W}^M$, we see that $\langle\bar{W}^M\rangle$ tends to $\mathrm{Id}_{[0,\infty)}$ in probability uniformly on compact time intervals. By Theorem VIII.3.11 of Jacod and Shiryaev (1987: p. 432) we conclude that $(\bar{W}^M)$ converges weakly in $\tilde{D}_\infty$ to a standard Wiener process W. That W has independent increments with respect to the filtration $(\mathcal{F}_t)$ can be seen by considering the first and second conditional moments of the increments of W^M for each $M \in \mathbb{N}$ and applying the conditions on local consistency and the jump heights of (ξ^M).

Suppose $((\bar{\xi}^M, R^M, \bar{W}^M))$ is weakly convergent with limit point (X, R, W). The remaining different part is the identification of X as a solution to Equation (2.18) under the relaxed control (R, W) with initial condition φ. Notice that X is continuous on $[0, \infty)$ because of the condition on the jump heights of (ξ^M), cf. Theorem 3.10.2 in Ethier and Kurtz (1986: p. 148). Let us define càdlàg processes C^M, C on $[0, \infty)$ by

$$C^M(t) := \;\varphi^M(0) \;+\; \int_0^t b\big(\bar{\xi}^M_s, \bar{u}^M(s)\big)\,ds \;+\; \varepsilon^M_2(t), \qquad t \geq 0,$$

$$C(t) := \;\varphi(0) \;+\; \int_{\Gamma\times[0,t]} b(X_s, \gamma)\,dR(s, \gamma), \qquad t \geq 0.$$

Then C, C^M are bounded on compact time intervals uniformly in $M \in \mathbb{N}$. Invoking Skorohod's representation theorem, one establishes weak convergence of (C^M) to C as in the proof of Proposition 2.1.

The sequence $(\bar{W}^M)$ is of *uniformly controlled variations*, hence a *good* sequence of integrators in the sense of Kurtz and Protter (1991), because the jump heights are uniformly bounded and $\bar{W}^M$ is a martingale for each $M \in \mathbb{N}$. We have weak convergence of $(\bar{W}^M)$ to W. The results in Kurtz and Protter (1991) guarantee weak convergence of the corresponding adapted quadratic variation processes, that is $([\bar{W}^M, \bar{W}^M])$ converges weakly to $[W, W]$ in $\tilde{D}_\infty = D_\mathbb{R}([0,\infty))$, where the square brackets indicate the adapted quadratic (co-)variation. Convergence also holds for the sequence of process pairs $(\bar{W}^M, [\bar{W}^M, \bar{W}^M])$ in $D_{\mathbb{R}^2}([0,\infty))$, see Theorem 36 in Kurtz and Protter (2004).

We now know that each of the sequences $(\bar{\xi}^M)$, (C^M), $(\bar{W}^M)$, $([\bar{W}^M, \bar{W}^M])$ is weakly convergent in $D_\mathbb{R}([0,\infty))$. Actually, we have weak convergence for the sequence of process quadruples $(\bar{\xi}^M, C^M, \bar{W}^M, [\bar{W}^M, \bar{W}^M])$ in $D_{\mathbb{R}^4}([0,\infty))$. To see this, notice that each of the sequences $(\bar{\xi}^M + C^M)$, $(\bar{\xi}^M + \bar{W}^M)$, $(\bar{\xi}^M + [\bar{W}^M, \bar{W}^M])$, $(C^M + \bar{W}^M)$, $(C^M + [\bar{W}^M, \bar{W}^M])$, and $(\bar{W}^M + [\bar{W}^M, \bar{W}^M])$ is tight in $D_\mathbb{R}([0,\infty))$, because the limit processes C, X, W, and $[W, W] = Id_{[0,\infty)}$ are all continuous on $[0,\infty)$. According to Problem 22 in Ethier and Kurtz (1986: p. 153) this implies tightness of the quadruple sequence in $D_{\mathbb{R}^4}([0,\infty))$. Since the four component sequences are all weakly convergent, the four-dimensional sequence must have a unique limit point, namely $(X, C, W, [W, W])$. By virtue of Skorohod's theorem, we may again work under $\mathbf{P}$-almost sure convergence. Since C, X, W, $[W, W]$ are all continuous, it follows that $C^M \to C$, $\bar{\xi}^M \to X$, $\bar{W}^M \to W$, $[\bar{W}^M, \bar{W}^M] \to [W, W]$ uniformly on compact subintervals of $[0,\infty)$ with probability one.

Define the mapping $F\colon D_0 \times \tilde{D}_\infty \to \tilde{D}_\infty$ by

$$F(\varphi, x)(t) := \sigma\left([-r, 0] \ni s \mapsto \begin{cases} x(t+s) & \text{if } t+s \geq 0, \\ \varphi(t+s) & \text{else} \end{cases}\right), \quad t \geq 0.$$

For $M \in \mathbb{N}$, let F^M be the mapping from $\tilde{D}_\infty$ to $\tilde{D}_\infty$ given by $F^M(x) := F(\varphi^M, x)$. Let $H^M : \tilde{D}_\infty \to \tilde{D}_\infty$ be the càdlàg interpolation operator of degree M, that is $H^M(x)$ is the piecewise constant càdlàg interpolation to $x \in \tilde{D}_\infty$ along the time grid of mesh size $\frac{r}{M}$ starting at zero. Define $\bar{F}^M : \tilde{D}_\infty \to \tilde{D}_\infty$ by

$$\bar{F}^M(x)(t) := F\big(\varphi^M, H^M(x)\big)\big(\lfloor t \rfloor_M\big), \quad t \geq 0,$$

where $\lfloor t \rfloor_M := \frac{r}{M} \lfloor \frac{M}{r} t \rfloor$. If $\psi \in D_\infty$, we will take $F^M(\psi)$, $\bar{F}^M(\psi)$ and $F(\psi)$ to equal $F^M(x)$, $\bar{F}^M(x)$ and $F(\varphi, x)$, respectively, where x is the restriction of ψ to $[0,\infty)$. Equation (2.22) translates to

$$\bar{\xi}^M(t) \;=\; C^M(t) \;+\; \int_0^t \bar{F}^M(\bar{\xi}^M)(s-) d\bar{W}^M(s), \quad t \geq 0.$$

Let $\hat{\xi}$ be the unique càdlàg process solving

$$\hat{\xi}(s) \;=\; \varphi(s), \quad s \in [-r, 0), \qquad \hat{\xi}(t) \;=\; C(t) \;+\; \int_0^t F(\hat{\xi})(s-) dW(s), \quad t \geq 0.$$

Fix $T > 0$. Since $\bar{\xi}^M$ converges to X as M goes to infinity uniformly on compacts with probability one, it is enough to show that

$$(*) \qquad \mathbf{E}\Big(\sup_{t \in [-r, T]} \big|\hat{\xi}(t) - \bar{\xi}^M(t)\big|^2\Big) \overset{M \to \infty}{\longrightarrow} 0.$$

First observe that

$$\mathbf{E}\Big(\sup_{t\in[0,T]}\big|C(t)-C^M(t)\big|^2\Big) \overset{M\to\infty}{\longrightarrow} 0, \qquad \sup_{t\in[-r,0)}\big|\hat{\xi}(t)-\bar{\xi}^M(t)\big|^2 \overset{M\to\infty}{\longrightarrow} 0,$$

because C is uniformly bounded on compact time intervals and φ is càdlàg and continuous on $[-r,0)$. Given $\varepsilon>0$, by Lemma 2.1 in Section 2.3.5 and by Gronwall's lemma we find that there is a positive number $M_0 = M_0(\varepsilon)$ such that for all $M \geq M_0$

$$\mathbf{E}\Big(\sup_{t\in[0,T]}\Big|\int_0^t F(\hat{\xi})(s-)dW(s) - \int_0^t \bar{F}^M(\bar{\xi}^M)(s-)d\bar{W}^M(s)\Big|^2\Big) \;\leq\; 76T\varepsilon(K^2+1)\exp\big(4K_L^2T\big).$$

This yields $(*)$ and the assertion follows. □

If we consider approximations along all equidistant partitions of $[-r,0]$, then the hypothesis about the uniform convergence of the initial conditions implies that φ must be continuous on $[-r,0]\setminus\{0\}$. In case φ has jumps at positions locatable on one of the equidistant partitions, the convergence results continue to hold when we restrict to a sequence of refining partitions.

2.3.4 Convergence of the minimal costs

The objective behind the introduction of sequences of approximating chains was to obtain a device for approximating the value function V of the original problem. At this point we define, for each discretisation degree $M\in\mathbb{N}$, a discrete control problem with cost functional J^M so that J^M is an approximation of the cost functional J of the original problem in the following sense: Given a suitable initial segment $\varphi\in D_0$ and a sequence of discrete admissible controls (u^M) such that $(\bar{u}^M)$ weakly converges to a relaxed control R, we have $J(\varphi,u^M)\to\hat{J}(\varphi,R)$ as M tends to infinity. Under the assumptions introduced above, it will follow that also the value functions associated with the discrete cost functionals converge to the value function of the original problem.

Fix $M\in\mathbb{N}$, and set $h:=\frac{r}{M}$. Denote by $\mathcal{U}_{ad}^M$ the set of discrete admissible controls of degree M. Define the *cost functional of degree M* by

$$J^M(\varphi,u) := \mathbf{E}\left(\sum_{n=0}^{N_h-1}\exp(-\beta nh)\cdot k\big(\xi(n),u(n)\big)\cdot h \;+\; g\big(\xi(N_h)\big)\right), \tag{2.23}$$

where $\varphi\in D_0$, $u\in\mathcal{U}_{ad}^M$ is defined on the stochastic basis $(\Omega,\mathcal{F},(\mathcal{F}_t),\mathbf{P})$ and $(\xi(n))$ is a discrete chain of degree M defined according to p^M and u with initial condition φ. The discrete *exit time step* N_h is given by

$$N_h := \min\{n\in\mathbb{N}_0 \mid \xi(n)\notin I_h\} \;\wedge\; \lfloor\tfrac{T}{h}\rfloor. \tag{2.24}$$

Denote by $\bar{\tau}^M := h\cdot N_h$ the exit time for the corresponding interpolated processes. The *value function of degree M* is defined as

$$V^M(\varphi) := \inf\big\{J^M(\varphi,u) \;\big|\; u\in\mathcal{U}_{ad}^M\big\}, \quad \varphi\in D_0. \tag{2.25}$$

We are now in a position to state the result about convergence of the minimal costs. Proposition 2.3 and Theorem 2.3 are comparable to Theorems 10.5.1 and 10.5.2 in Kushner and Dupuis (2001: pp. 292-295). Let us suppose that the initial condition $\varphi \in D_0$ and the sequence of partitions of $[-r, 0]$ are such that the discretised initial conditions converge to φ uniformly on $[-r, 0]$.

Proposition 2.3. *Assume (A1) – (A5). If the sequence* $(\bar{\xi}^M, \bar{u}^M, \bar{W}^M, \bar{\tau}^M)$ *of interpolated processes converges weakly to a limit point* (X, R, W, τ)*, then* X *is a solution to Equation* (2.18) *under relaxed control* (R, W) *with initial condition* φ*,* τ *is the exit time for* X *as given by* (2.15)*, and we have*

$$J^M(\varphi, u^M) \overset{M\to\infty}{\longrightarrow} \hat{J}(\varphi, R).$$

Proof. The convergence assertion for the costs is a consequence of Proposition 2.2, the fact that, by virtue of Assumption (A5), the exit time $\hat{\tau}$ defined in (2.21) is Skorohod-continuous, and the definition of J^M and J (or $\hat{J}$). □

Theorem 2.3. *Assume (A1) – (A5). Then we have* $\lim_{M\to\infty} V^M(\varphi) = V(\varphi)$*.*

Proof. First notice that $\liminf_{M\to\infty} V^M(\varphi) \geq V(\varphi)$ as a consequence of Propositions 2.2 and 2.3. In order to show $\limsup_{M\to\infty} V^M(\varphi) \leq V(\varphi)$ choose a relaxed control (R, W) so that $\hat{J}(\varphi, R) = V(\varphi)$ according to Proposition 2.1. Given $\varepsilon > 0$, one can construct a sequence of discrete admissible controls (u^M) such that $((\bar{\xi}^M, \bar{u}^M, \bar{W}^M, \bar{\tau}^M))$ is weakly convergent, where $(\bar{\xi}^M)$, $(\bar{W}^M)$, $(\bar{\tau}^M)$ are constructed as above, and

$$\limsup_{M\to\infty} |J^M(\varphi, u^M) - \hat{J}(\varphi, R)| \leq \varepsilon.$$

The existence of such a sequence of discrete admissible controls is guaranteed, cf. the discussion at the end of Subsection 2.3.2. By definition, $V^M(\varphi) \leq J^M(\varphi, u^M)$ for each $M \in \mathbb{N}$. Using Proposition 2.3 we find that

$$\limsup_{M\to\infty} V^M(\varphi) \;\leq\; \limsup_{M\to\infty} J^M(\varphi, u^M) \;\leq\; V(\varphi) + \varepsilon,$$

and since ε was arbitrary, the assertion follows. □

The assertion of Theorem 2.3 corresponds to the convergence statement of Theorem 2.1 in Subsection 2.2.2. Let us check whether the hypotheses of Theorem 2.1 are satisfied.

Hypothesis (H1) is met since the cost functional $\hat{J}$ given by (2.19) may be regarded as a mapping $D_0 \times \mathcal{B}(D_\infty \times \mathcal{R} \times \tilde{D}_\infty \times [0, \infty]) \to (-\infty, \infty)$, which is continuous with respect to the topology of weak convergence on the second component and uniform convergence on D_0. Hypothesis (H2) is satisfied because of (2.23), the definition of the discrete cost functionals. Hypothesis (H4) is a consequence of the first part of Proposition 2.2 and the compactness of the space of relaxed control processes. Hypothesis (H5) follows from the second part of Proposition 2.2.

Lastly, Hypothesis (H3), which we have skipped so far, is implied by Proposition 2.3 and the fact that continuous-time relaxed control processes can be approximated (in the sense of weak convergence) by ordinary control processes which are piecewise constant on uniform grids of mesh size $\frac{1}{M}$. A different criterion for the approximation of continuous-time strategies will be applied in Chapter 3. There, not only the drift coefficient of the state equation, but also the diffusion coefficient may be controlled.

2.3.5 An auxiliary result

The proof of the following lemma makes use of standard techniques. In the context of approximation of SDDEs, it should be compared to Section 7 in Mao (2003).

Lemma 2.1. *In the notation and under the assumptions of Proposition 2.2, it holds that for every $\varepsilon > 0$ there is $M_0 \in \mathbb{N}$ such that for all $M \geq M_0$,*

$$\begin{aligned} &\mathbf{E}\Big(\sup_{t\in[0,T]}\Big|\int_0^t F(\hat{\xi})(s-)dW(s) - \int_0^t \bar{F}^M(\bar{\xi}^M)(s-)d\bar{W}^M(s)\Big|^2\Big) \\ \leq\ & 4K_L^2 \int_0^T \mathbf{E}\Big(\sup_{t\in[-r,s]}\big|\hat{\xi}(t) - \bar{\xi}^M(t)\big|^2\Big)\, ds \ +\ 76T\varepsilon(K^2+1). \end{aligned}$$

Proof. Clearly,

$$\begin{aligned} &\mathbf{E}\Big(\sup_{t\in[0,T]}\Big|\int_0^t F(\hat{\xi})(s-)dW(s) - \int_0^t \bar{F}^M(\bar{\xi}^M)(s-)d\bar{W}^M(s)\Big|^2\Big) \\ \leq\ & 2\,\mathbf{E}\Big(\sup_{t\in[0,T]}\Big|\int_0^t F(\hat{\xi})(s-)dW(s) - \int_0^t \bar{F}^M(\bar{\xi}^M)(s-)dW(s)\Big|^2\Big) \\ &+ 2\,\mathbf{E}\Big(\sup_{t\in[0,T]}\Big|\int_0^t \bar{F}^M(\bar{\xi}^M)(s-)dW(s) - \int_0^t \bar{F}^M(\bar{\xi}^M)(s-)d\bar{W}^M(s)\Big|^2\Big) \end{aligned} \tag{2.26}$$

Using Doob's maximal inequality, Itô's isometry, Fubini's theorem and Assumption (A4), for the first expectation on the right hand side of (2.26) we obtain the estimate

$$\begin{aligned} &\mathbf{E}\Big(\sup_{t\in[0,T]}\Big|\int_0^t F(\hat{\xi})(s-)dW(s) - \int_0^t \bar{F}^M(\bar{\xi}^M)(s-)dW(s)\Big|^2\Big) \\ \leq\ & 4\int_0^T \mathbf{E}\Big(\big|F(\hat{\xi})(s) - \bar{F}^M(\bar{\xi}^M)(s)\big|^2\Big)\, ds \\ \leq\ & 4K_L^2 \int_0^T \mathbf{E}\Big(\sup_{t\in[-r,s]}\big|\hat{\xi}(t) - \bar{\xi}^M(t)\big|^2\Big)\, ds. \end{aligned} \tag{2.27}$$

Fix any $N \in \mathbb{N}$. The second expectation on the right hand side of (2.26) splits up into three terms according to

$$\begin{aligned} &\mathbf{E}\Big(\sup_{t\in[0,T]}\Big|\int_0^t \bar{F}^M(\bar{\xi}^M)(s-)dW(s) - \int_0^t \bar{F}^M(\bar{\xi}^M)(s-)d\bar{W}^M(s)\Big|^2\Big) \\ \leq\ & 4\,\mathbf{E}\Big(\sup_{t\in[0,T]}\Big|\int_0^t \bar{F}^M(\bar{\xi}^M)(s-)dW(s) - \int_0^t \bar{F}^N(\bar{\xi}^M)(s-)dW(s)\Big|^2\Big) \\ &+ 4\,\mathbf{E}\Big(\sup_{t\in[0,T]}\Big|\int_0^t \bar{F}^N(\bar{\xi}^M)(s-)dW(s) - \int_0^t \bar{F}^N(\bar{\xi}^M)(s-)d\bar{W}^M(s)\Big|^2\Big) \\ &+ 4\,\mathbf{E}\Big(\sup_{t\in[0,T]}\Big|\int_0^t \bar{F}^N(\bar{\xi}^M)(s-)d\bar{W}^M(s) - \int_0^t \bar{F}^M(\bar{\xi}^M)(s-)d\bar{W}^M(s)\Big|^2\Big). \end{aligned} \tag{2.28}$$

Again using Doob's maximal inequality and a generalized version of Itô's isometry (cf. Protter, 2003: pp. 73-77), for the first and third expectation on the right hand side of

Inequality (2.28) we obtain

$$
\begin{aligned}
&\mathbf{E}\Big(\sup_{t\in[0,T]}\Big|\int_0^t \bar F^M(\bar\xi^M)(s-)dW(s) - \int_0^t \bar F^N(\bar\xi^M)(s-)dW(s)\Big|^2\Big) \\
\le\ & 4\,\mathbf{E}\Big(\int_0^T \big|\bar F^M(\bar\xi^M)(s) - \bar F^N(\bar\xi^M)(s)\big|^2\,ds\Big)
\end{aligned} \tag{2.29}
$$

and

$$
\begin{aligned}
&\mathbf{E}\Big(\sup_{t\in[0,T]}\Big|\int_0^t \bar F^N(\bar\xi^M)(s-)d\bar W^M(s) - \int_0^t \bar F^M(\bar\xi^M)(s-)d\bar W^M(s)\Big|^2\Big) \\
\le\ & 4\,\mathbf{E}\Big(\int_0^T \big|\bar F^M(\bar\xi^M)(s-) - \bar F^N(\bar\xi^M)(s-)\big|^2\,d\big[\bar W^M,\bar W^M\big](s)\Big).
\end{aligned} \tag{2.30}
$$

Notice that, path-by-path, we have

$$
\begin{aligned}
&\int_0^T \big|\bar F^M(\bar\xi^M)(s-) - \bar F^N(\bar\xi^M)(s-)\big|^2\,d\big[\bar W^M,\bar W^M\big](s) \\
\le\ & \sum_{i=0}^{\lfloor \frac{M}{r}T\rfloor} \big|\bar F^M(\bar\xi^M)\big(\tfrac{r}{M}i\big) - \bar F^N(\bar\xi^M)\big(\tfrac{r}{M}i\big)\big|^2\cdot\Big([\bar W^M,\bar W^M]\big(\tfrac{r}{M}(i+1)\big) - [\bar W^M,\bar W^M]\big(\tfrac{r}{M}i\big)\Big).
\end{aligned}
$$

In order to estimate the second expectation on the right hand side of (2.28), observe that, $\mathbf{P}$-almost surely, for all $t\in[0,T]$

$$
\begin{aligned}
\int_0^t \bar F^N(\bar\xi^M)(s-)dW(s) &= \bar F^N(\bar\xi^M)\big(\lfloor t\rfloor_N\big)\cdot\Big(W(t) - W\big(\lfloor t\rfloor_N\big)\Big) \\
&\quad + \sum_{i=0}^{\lfloor \frac{N}{r}t\rfloor -1} \bar F^N(\bar\xi^M)\big(\tfrac{r}{N}i\big)\cdot\Big(W\big(\tfrac{r}{N}(i+1)\big) - W\big(\tfrac{r}{N}i\big)\Big),
\end{aligned}
$$

as $F^N(\bar\xi^M)$ is piecewise constant on the grid of mesh size $\frac{r}{N}$. On the other hand,

$$
\begin{aligned}
\int_0^t \bar F^N(\bar\xi^M)(s-)d\bar W^M(s) &= \bar F^N(\bar\xi^M)\big(\lfloor t\rfloor_N\big)\cdot\Big(\bar W^M(t) - \bar W^M\big(\lfloor t\rfloor_N\big)\Big) \\
&\quad + \sum_{i=0}^{\lfloor \frac{N}{r}t\rfloor -1} \bar F^N(\bar\xi^M)\big(\tfrac{r}{N}i\big)\cdot\Big(\bar W^M\big(\tfrac{r}{N}(i+1)\big) - \bar W^M\big(\tfrac{r}{N}i\big)\Big).
\end{aligned}
$$

By Assumption (A3), $|\sigma|$ is bounded by a constant K, hence

$$
\begin{aligned}
&\Big|\int_0^t \bar F^N(\bar\xi^M)(s-)dW(s) - \int_0^t \bar F^N(\bar\xi^M)(s-)d\bar W^M(s)\Big| \\
\le\ & 2K\lfloor \tfrac{N}{r}t\rfloor\cdot\sup_{s\in[0,t]}|W(s) - \bar W^M(s)| \quad\le\quad 2K\tfrac{N}{r}T\cdot\sup_{s\in[0,T]}|W(s) - \bar W^M(s)|.
\end{aligned}
$$

Bounded convergence yields for each fixed $N\in\mathbb{N}$

$$
\mathbf{E}\Big(\sup_{t\in[0,T]}\Big|\int_0^t \bar F^N(\bar\xi^M)(s-)dW(s) - \int_0^t \bar F^N(\bar\xi^M)(s-)d\bar W^M(s)\Big|^2\Big) \overset{M\to\infty}{\longrightarrow} 0. \tag{2.31}
$$

Let $x, y \in \tilde{D}_\infty$. By Assumption (A4) we have for all $t \in [0,T]$

$$\begin{aligned} \left|\bar{F}^N(y)(t) - F(\varphi,x)(t)\right| &= \left|F\big(\varphi^N, H^N(y)\big)(\lfloor t\rfloor_N) - F(\varphi,x)(t)\right| \\ \leq\ K_L \cdot \sup_{s\in[-r,0)} \left|\varphi^N(s) - \varphi(s)\right| &+ K_L \cdot \sup_{s\in[0,T]} \left|H^N(y)(s) - x(s)\right| \\ + \left|F(\varphi,x)\big(\lfloor t\rfloor_N\big) - F(\varphi,x)(t)\right|. \end{aligned}$$

By Assumption (A1), the map $[0,T] \ni t \mapsto F(\varphi,x)(t)$ is càdlàg, whence it has only finitely many jumps larger than any given positive lower bound. Thus, given $\varepsilon > 0$, there is a finite subset $A = A(\varepsilon, T, \varphi, x) \subset [0,T]$ such that

$$\limsup_{N\to\infty} \left|F(\varphi,x)\big(\lfloor t\rfloor_N\big) - F(\varphi,x)(t)\right| \quad \leq\ \varepsilon \quad \text{for all } t \in [0,T] \setminus A.$$

Moreover, the convergence is uniform in the following sense (cf. Billingsley, 1999): We can choose the finite set A in such a way that there is $N_0 = N_0(\varepsilon, T, \varphi, x) \in \mathbb{N}$ so that

$$\left|F(\varphi,x)\big(\lfloor t\rfloor_N\big) - F(\varphi,x)(t)\right| \quad \leq\ 2\varepsilon \quad \text{for all } t \in [0,T] \setminus A,\ N \geq N_0.$$

Given $\varepsilon > 0$, we therefore find $N \in \mathbb{N}$ and an event $\tilde{\Omega}$ with $\mathbf{P}(\tilde{\Omega}) \geq 1-\varepsilon$ so that for each $\omega \in \tilde{\Omega}$ there is a finite subset $A_\omega \subset [0,T]$ with $\#A_\omega \leq N\varepsilon$ and such that for all $t \in [0,T] \setminus A_\omega$ and all $M \geq N$ we have

$$\left|\bar{F}^M\big(\bar{\xi}^M(\omega)\big)(t) - F\big(X(\omega)\big)(t)\right|^2 + \left|\bar{F}^N\big(\bar{\xi}^M(\omega)\big)(t) - F\big(X(\omega)\big)(t)\right|^2 \quad \leq \quad \varepsilon.$$

The expression on the right hand side of (2.29) is then bounded from above by $9T\varepsilon(K^2+1)$. For M big enough, also the expression on the right hand side of (2.30) is smaller than $9T\varepsilon(K^2+1)$, and the expectation in (2.31) is smaller than $T\varepsilon$. □

2.4 Discussion

Kushner's method applies to approximation schemes which replace a given optimal control problem, usually one over continuous time and with continuous state space, by a sequence of approximating control problems, usually defined for discrete time and discrete state space. When the dynamics of the original problem are described by some kind of deterministic or stochastic differential equation, conditions of local consistency indicate how to choose the dynamics of the approximating problems in a consistent way; that is, in such a way that the associated value functions converge to the value function of the original problem. Local consistency is easy to check when the dynamics are discretised according to some finite differences scheme.

A crucial assumption on the original problem is that the space of control actions is compact. This assumption is less restrictive than it might appear insofar as, in actual numerical computations, optimisation is often performed only with respect to a finite set of control actions; see, for example, the appendix by M. Falcone in Bardi and Capuzzo Dolcetta (1997). However, for the proof of convergence to work, compactness of the space of control actions must carry over to the space of admissible strategies; at least, compactness for sequences of solutions must hold as required by Hypothesis (H4) of Theorem 2.1, the abstract convergence result.

In the case of dynamics described by a stochastic differential equation, compactness of the space of strategies is achieved by introducing relaxed control processes in a way analogous to the deterministic case provided the diffusion coefficient is independent of the control. This is the case for the control problems of Section 2.3, for instance. When also the diffusion coefficient depends on the control, the situation gets more complicated. Martingale measures on the space of control actions may be introduced to obtain the desired compactness of the space of strategies, see Kushner and Dupuis (2001: Ch. 13).

As far as the structure of the control problem that has to be approximated is concerned, the Markov chain method is extremely general. This was demonstrated in Section 2.2, where we set up an abstract framework which encompasses quite arbitrary optimal control problems. The generality of the approach is also its limitation. In particular, it is not clear how to obtain a priori bounds on the approximation error, in addition to convergence.

A priori bounds on the discretisation error are important for several reasons. Error bounds provide an assurance – though usually over-pessimistic – about the accuracy of the approximations relative to the original problem. They also allow to compare different schemes for the same class of problems; or they may serve as a benchmark for new schemes. Lastly, they give an indication of the computational resources required for solving the discretised problems.

In Chapter 3, we will change attitude and develop a more specific scheme for the approximation of control problems with delay, exploiting, in particular, the additivity of the minimal costs as expressed by the Principle of Dynamic Programming. We will obtain bounds on the error for the discretisation in time. Questions of computational requirements and complexity for the solution of the resulting semi-discrete problems will be discussed. The idea, also present in Kushner's method, to discretise a continuous-time control problem by constructing a sequence of approximating problems will be retained.

Chapter 3

Two-step time discretisation and error bounds

In this chapter, we study a semi-discretisation scheme for stochastic systems with delay. Material of this chapter appears in Fischer and Nappo (2007). The control problems to be approximated are characterised as follows: The system dynamics are given by a multi-dimensional controlled stochastic functional differential equation with bounded memory driven by a Wiener process. The driving noise process and the state process may have different dimensions. The optimal control problem itself is, in general, infinite-dimensional in the sense that the associated value function lives on an infinite-dimensional function space. There will be no need to assume ellipticity of the diffusion matrix so that deterministic control problems are included as special cases. The performance criterion is a cost functional of evolutional type over a finite deterministic time horizon. For simplicity, there will be neither state constraints nor state-dependent control constraints.

Our scheme is based on a time discretisation of Euler-Maruyama type and yields a sequence of finite-dimensional optimal control problems in discrete time. Here, as in Chapter 2, we follow the approach where a given control problem is approximated by a sequence of control problems which are easier to solve numerically – or solvable at all. Under quite natural assumptions, we obtain upper bounds on the discretisation error – or worst-case estimates for the rate of convergence – in terms of differences in supremum norm between the value functions corresponding to the original control problem and the approximating control problems, respectively.

The approximation of the original control problem is carried out in two steps. The idea is to separate the discretisation of the dynamics from that of the strategies. The dynamics are discretised first. By "freezing" the dynamics, the problem of approximating the strategies is reduced to the finite-dimensional "constant coefficients" case and results available in the literature can be applied. Notice that the state processes always have a certain time regularity (they are Hölder continuous like typical trajectories of Brownian motion), while the strategies need not have any regularity in time besides being measurable.

The first discretisation step consists in constructing a sequence of control problems whose coefficients are piecewise constant in both the time and the segment variable. The admissible strategies are the same as those of the original problem. We obtain a rate of convergence for the controlled state processes, which is uniform in the strategies, thanks to the fact that the modulus of continuity of Itô diffusions with bounded coefficients has finite moments of all orders. This result can be found in Słomiński (2001), cf. Appendix A.2 below. The convergence rate for the controlled processes carries over to the approximation of the corresponding value functions.

The second discretisation step consists in approximating the original strategies by control processes which are piecewise constant on a sub-grid of the time grid introduced in the first step. A main ingredient in the derivation of an error bound is the Principle of Dynamic Programming (PDP) or, as it is also known, Bellman's Principle of Optimality. The validity of the PDP for the "non-Markovian" dynamics at hand is proved in Larssen (2002), cf. Appendix A.1 below. A version of the PDP for controlled diffusions with time delay is also proved in Gihman and Skorokhod (1979: Ch. 3); there are differences, though, in the formulation of the control problem.

We apply the PDP to obtain a global error bound from an estimate of the local truncation error. The fact that the value functions of the approximating problems from the first step are Lipschitz continuous under the supremum norm guarantees stability of the method. This way of error localisation and, in particular, the use of the PDP are adapted from Falcone and Ferretti (1994) and Falcone and Rosace (1996), who study deterministic optimal control problems with and without delay. Their proof technique is not confined to such simple approximation schemes as we adopt here; it extends the usual convergence analysis of finite difference methods for initial-boundary value problems, cf. Section 5.3 in Atkinson and Han (2001), for example.

To estimate the local truncation error we only need an error bound for the approximation by piecewise constant strategies of finite-dimensional control problems with "constant coefficients"; that is, the cost rate and the coefficients of the state equation are functions of the control variable only. Such a result is provided by a stochastic mean value theorem due to Krylov (2001). When the space of control actions is finite and the diffusion coefficient is not directly controlled, it is possible to derive an analogous result with an error bound of higher order, namely of order $h^{1/2}$ instead of $h^{1/4}$, where h is the length of the time step. When the control problem is deterministic, the error bound is at least of order $h^{1/2}$; it is of order h if, in addition, the space of control actions is finite. In Appendix A.3, we state a reduced version of Krylov's theorem and provide a detailed proof. The more elementary error bounds for special cases are also given.

In a final step, we put together the two error estimates to obtain bounds on the total approximation error. The error bound in the most general case is of order nearly $h^{1/12}$ with h the length of the time step, see Theorem 3.4 in Section 3.4. To the best of our knowledge, this is the first result on the speed of convergence of a time-discretisation scheme for controlled stochastic systems with delay. We do not expect our worst-case estimates to be optimal; in any case, they may serve as benchmarks on the way towards sharp error bounds. Moreover, the scheme's special structure can be exploited so that the computational requirements are lower than what might be expected by looking at the order of the error bound.

In the finite-dimensional setting, our two-step time-discretisation procedure allows to get from the case of "constant coefficients" to the case of general coefficients, even though it yields a worse rate of convergence in comparison with the results cited in Section 1.3, namely $\frac{1}{12}$ instead of $\frac{1}{6}$ and $\frac{1}{10}$, respectively. This is the price we pay for separating the approximation of the dynamics from that of the strategies. On the other hand, it is this separation that enables us to reduce the problem of strategy approximation to an elementary form. Observe that certain techniques like mollification of the value function employed in the works cited above are not available, because the space of initial values is not locally compact.

Our procedure also allows to estimate the error incurred when using strategies which are nearly optimal for the approximating problems with the dynamics of the original problem. This would be the way to apply the approximation scheme in many practically relevant situations. However, this method of nearly optimally controlling the original system is viable only if the available information includes perfect samples of the underlying noise process. The question is more complicated when information is restricted to samples of the state process.

In Section 3.1, the original control problem is described in detail. The dynamics of the original control problem are discretised in Section 3.2. The second discretisation step, based on the PDP and local error bounds for the approximation of the original strategies, is carried out in Section 3.3. In Section 3.4, bounds on the overall discretisation error are derived. In Section 3.5, a procedure for solving the resulting finite-dimensional problems is outlined. Section 3.6 contains some concluding remarks and open questions.

3.1 The original control problem

The dynamics of the control problems we want to approximate are described by a controlled d-dimensional stochastic delay (or functional) differential equation driven by a Wiener process. Both the drift and the diffusion coefficient may depend on the solution's history a certain amount of time into the past. The *delay length* gives a bound on the maximal time the system is allowed to look back into the past; as before, we take it to be a finite deterministic time $r > 0$. For simplicity, we restrict attention to control problems with finite and deterministic time horizon. The performance of the admissible control processes or strategies will be measured in terms of a cost functional of evolutional type.

Recall that, in general, the solution process of an SDDE does not enjoy the Markov property, while the segment process associated with that solution does. For an $\mathbb{R}^d$-valued stochastic process $(X(t))_{t\geq -r}$ living on $(\Omega, \mathcal{F}, \mathbf{P})$, we denote by $(X_t)_{t\geq 0}$ the associated segment process of delay length r. Thus, for any $t \geq 0$, any $\omega \in \Omega$, $X_t(\omega)$ is the function $[-r, 0] \ni s \mapsto X(t+s, \omega) \in \mathbb{R}^d$. If the original process $(X(t))_{t\geq -r}$ has continuous trajectories, then $(X_t)_{t\geq 0}$ is a stochastic process taking its values in $\mathcal{C} := \mathbf{C}([-r, 0], \mathbb{R}^d)$, the space of all $\mathbb{R}^d$-valued continuous functions on the interval $[-r, 0]$. The space $\mathcal{C}$ comes equipped with the supremum norm, written $\|.\|$, induced by the standard norm on $\mathbb{R}^d$.

Let (Γ, ρ) be a complete and separable metric space, the set of *control actions*. We first state our control problem in the weak Wiener formulation, cf. Larssen (2002) and Yong and Zhou (1999: pp. 176-177). This is to justify our use of the Principle of Dynamic Programming. In subsequent sections we will only need the strong formulation.

Definition 3.1. A *Wiener basis* of dimension d_1 is a triple $((\Omega, \mathbf{P}, \mathcal{F}), (\mathcal{F}_t), W)$ such that

(i) $(\Omega, \mathcal{F}, \mathbf{P})$ is a complete probability space carrying a standard d_1-dimensional Wiener process W,

(ii) $(\mathcal{F}_t)$ is the completion by the $\mathbf{P}$-null sets of $\mathcal{F}$ of the filtration induced by W.

A *Wiener control basis* is a quadruple $((\Omega,\mathbf{P},\mathcal{F}),(\mathcal{F}_t),W,u)$ such that $((\Omega,\mathbf{P},\mathcal{F}),(\mathcal{F}_t),W)$ is a Wiener basis and $u\colon [0,\infty)\times\Omega\to\Gamma$ is progressively measurable with respect to $(\mathcal{F}_t)$. The $(\mathcal{F}_t)$-progressively measurable process u is called a *control process*. Write $\mathcal{U}_W$ for the set of all Wiener control bases.

By abuse of notation, we will often hide the stochastic basis involved in the definition of a Wiener control basis; thus, we may write $(W,u)\in\mathcal{U}_W$ meaning that W is the Wiener process and u the control process of a Wiener control basis.

Let b, σ be Borel measurable functions defined on $[0,\infty)\times\mathcal{C}\times\Gamma$ and taking values in $\mathbb{R}^d$ and $\mathbb{R}^{d\times d_1}$, respectively. The functions b, σ are the *coefficients* of the controlled SDDE that describes the dynamics of the control problem. The SDDE is of the form

$$dX(t) \;=\; b\big(t_0{+}t, X_t, u(t)\big)dt \;+\; \sigma\big(t_0{+}t, X_t, u(t)\big)dW(t), \quad t>0, \tag{3.1}$$

where $t_0\geq 0$ is a deterministic initial time and $((\Omega,\mathbf{P},\mathcal{F}),(\mathcal{F}_t),W,u)$ a Wiener control basis. The assumptions on the coefficients stated below will allow b, σ to depend on the segment variable in different ways. Let $\varphi\in\mathcal{C}$ be a generic segment function. The coefficients b, σ may depend on φ through bounded Lipschitz functions of, for example,

$$\begin{aligned}
&\varphi(-r_1),\dots,\varphi(-r_n), && (\textit{point delay}),\\
&\int_{-r}^0 v_1(s,\varphi(s))w_1(s)ds,\dots,\int_{-r}^0 v_n(s,\varphi(s))w_n(s)ds && (\textit{distributed delay}),\\
&\int_{-r}^0 \tilde v_1(s,\varphi(s))d\mu_1(s),\dots,\int_{-r}^0 \tilde v_n(s,\varphi(s))d\mu_n(s), && (\textit{generalised distributed delay}),
\end{aligned}$$

where $n\in\mathbb{N}$, $r_1,\dots,r_n\in[0,r]$, $w_1,\dots,w_n$ are Lebesgue integrable, $\mu_1,\dots,\mu_n$ are finite Borel measures on $[0,r]$, v_i, $\tilde v_i$ are Lipschitz continuous in the second variable uniformly in the first, $v_i(.,0)w_i(.)$ is Lebesgue integrable and $\tilde v_i(.,0)$ is μ_i-integrable, $i\in\{1,\dots,n\}$. Notice that the *generalised distributed delay* comprises the *point delay* as well as the Lebesgue absolutely continuous *distributed delay*. Let us call *functional delay* any type of delay that cannot be written in integral form. An example of a functional delay, which is also covered by the regularity assumptions stated below, is the dependence on the segment variable φ through bounded Lipschitz functions of

$$\sup_{s,t\in[-r,0]} \bar v_1(s,t,\varphi(s),\varphi(t)),\dots,\sup_{s,t\in[-r,0]} \bar v_n(s,t,\varphi(s),\varphi(t)),$$

where $\bar v_i$ is a measurable function which is Lipschitz continuous in the last two variables uniformly in the first two variables and $\bar v_i(.,.,0,0)$ is bounded, $i\in\{1,\dots,n\}$.

As initial condition for Equation (3.1), in addition to the time t_0, we have to prescribe the values of $X(t)$ for all $t\in[-r,0]$, not only for $t=0$. Thus, a deterministic initial condition for Equation (3.1) is a pair (t_0,φ), where $t_0\geq 0$ is the initial time and $\varphi\in\mathcal{C}$ the initial segment. We understand Equation (3.1) in the sense of an Itô equation. An adapted process X with continuous paths defined on the stochastic basis $(\Omega,\mathbf{P},\mathcal{F},(\mathcal{F}_t))$ of (W,u) is a *solution* with initial condition (t_0,φ) if it satisfies, $\mathbf{P}$-almost-surely,

$$X(t) = \begin{cases}\varphi(0) + \int_0^t b\big(t_0{+}s,X_s,u(s)\big)ds \;+\; \int_0^t \sigma\big(t_0{+}s,X_s,u(s)\big)dW(s), & t>0,\\ \varphi(t), \quad t\in[-r,0].\end{cases} \tag{3.2}$$

Observe that the solution process X always starts at time zero; it depends on the initial time t_0 only through the coefficients b, σ. As far as the control problem is concerned, this formulation is equivalent to the usual one, where the process X starts at time t_0 with initial condition $X_{t_0} = \varphi$ and t_0 does not appear in the time argument of the coefficients.

A solution X to Equation (3.2) under (W, u) with initial condition (t_0, φ) is *strongly unique* if it is indistinguishable from any other solution $\tilde{X}$ satisfying Equation (3.2) under (W, u) with the same initial condition. A solution X to Equation (3.2) under (W, u) with initial condition (t_0, φ) is *weakly unique* if (X, W, u) has the same distribution as $(\tilde{X}, \tilde{W}, \tilde{u})$ whenever $(\tilde{W}, \tilde{u})$ has the same distribution as (W, u) and $\tilde{X}$ is a solution to Equation (3.2) under Wiener control basis $(\tilde{W}, \tilde{u})$ with initial condition (t_0, φ). Here, the space of Borel measurable functions $[0, \infty) \rightarrow \Gamma$ is equipped with the topology of convergence locally in Lebesgue measure.

Definition 3.2. A Wiener control basis $(W, u) \in \mathcal{U}_W$ is called *admissible* or an *admissible strategy* if, for each deterministic initial condition, Equation (3.2) has a strongly unique solution under (W, u) which is also weakly unique. Write $\mathcal{U}_{ad}$ for the set of admissible control bases.

Denote by $T > 0$ the finite deterministic time horizon. Let f, g be Borel measurable real-valued functions with f having domain $[0, \infty) \times \mathcal{C} \times \Gamma$ and g having domain $\mathcal{C}$. They will be referred to as the *cost rate* and the *terminal cost*, respectively. We introduce a *cost functional* J defined on $[0, T] \times \mathcal{C} \times \mathcal{U}_{ad}^J$ by setting

$$J(t_0, \varphi, (W, u)) := \mathbf{E}\left(\int_0^{T-t_0} f\big(t_0+s, X_s, u(s)\big)ds \;+\; g(X_{T-t_0})\right), \tag{3.3}$$

where X is the solution to Equation (3.2) under $(W, u) \in \mathcal{U}_{ad}^J$ with initial condition (t_0, φ) and $\mathcal{U}_{ad}^J \subseteq \mathcal{U}_{ad}$ is the set of all admissible Wiener control bases such that the expectation in (3.3) is well defined for all deterministic initial conditions.

The *value function* corresponding to Equation (3.2) and cost functional (3.3) is the function $V\colon [0, T] \times \mathcal{C} \rightarrow [-\infty, \infty)$ given by

$$V(t_0, \varphi) := \inf\big\{J(t_0, \varphi, (W, u)) \mid (W, u) \in \mathcal{U}_{ad}^J\big\}. \tag{3.4}$$

It is this function that we wish to approximate.

Let us specify the hypotheses we make about the regularity of the coefficients b, σ, the cost rate f and the terminal cost g.

(A1) Measurability: the functions $b\colon [0, \infty) \times \mathcal{C} \times \Gamma \rightarrow \mathbb{R}^d$, $\sigma\colon [0, \infty) \times \mathcal{C} \times \Gamma \rightarrow \mathbb{R}^{d \times d_1}$, $f\colon [0, \infty) \times \mathcal{C} \times \Gamma \rightarrow \mathbb{R}$, $g\colon \mathcal{C} \rightarrow \mathbb{R}$ are jointly Borel measurable.

(A2) Boundedness: $|b|$, $|\sigma|$, $|f|$, $|g|$ are bounded by some constant $K > 0$.

(A3) Uniform Lipschitz and Hölder condition: there is a constant $L > 0$ such that for all $\varphi, \tilde{\varphi} \in \mathcal{C}$, $t, s \geq 0$, all $\gamma \in \Gamma$

$$\begin{aligned} |b(t, \varphi, \gamma) - b(s, \tilde{\varphi}, \gamma)| \vee |\sigma(t, \varphi, \gamma) - \sigma(s, \tilde{\varphi}, \gamma)| &\leq L\big(\|\varphi - \tilde{\varphi}\| + \sqrt{|t-s|}\big) \\ |f(t, \varphi, \gamma) - f(s, \tilde{\varphi}, \gamma)| \vee |g(\varphi) - g(\tilde{\varphi})| &\leq L\big(\|\varphi - \tilde{\varphi}\| + \sqrt{|t-s|}\big). \end{aligned}$$

(A4) Continuity in the control: $b(t,\varphi,.)$, $\sigma(t,\varphi,.)$, $f(t,\varphi,.)$ are continuous functions on Γ for any $t \geq 0$, $\varphi \in \mathcal{C}$.

Here and in the sequel, $|.|$ denotes the Euclidean norm of appropriate dimension and $x \vee y$ denotes the maximum of x and y. The above measurability, boundedness and Lipschitz continuity assumptions on the coefficients b, σ guarantee the existence of a strongly unique solution $X = X^{t_0,\varphi,u}$ to Equation (3.2) for every initial condition $(t_0,\varphi) \in [0,T] \times \mathcal{C}$ and $(W,u) \in \mathcal{U}_W$ any Wiener control basis; see, for example, Theorem 2.1 and Remark 1.1(2) in Chapter 2 of Mohammed (1984). Moreover, weak uniqueness of solutions holds for all deterministic initial conditions. This is a consequence of a theorem due to Yamada and Watanabe, see Larssen (2002) for the necessary generalisation to SDDEs.

Consequently, under Assumptions (A1)–(A3), we have $\mathcal{U}_{ad} = \mathcal{U}_W$. Moreover, since f and g are assumed to be measurable and bounded, the expectation in (3.3) is always well defined, whence it holds that $\mathcal{U}_{ad}^J = \mathcal{U}_{ad} = \mathcal{U}_W$. Assumption (A4) will not be needed before Section 3.3.

The fact that weak uniqueness holds allows us to discard the weak formulation and consider our control problem in the strong Wiener formulation. Thus, we may work with a fixed Wiener basis. Under Assumptions (A1)–(A3), the admissible strategies will be precisely the natural strategies, that is, those that are representable as functionals of the driving Wiener process. From now on, let $((\Omega,\mathbf{P},\mathcal{F}),(\mathcal{F}_t),W)$ be a fixed d_1-dimensional Wiener basis. Denote by $\mathcal{U}$ the set of control processes defined on this stochastic basis.

The dynamics of our control problem are still given by Equation (3.2). Due to Assumptions (A1)–(A3), all control processes are admissible in the sense that Equation (3.2) has a (strongly) unique solution under any $u \in \mathcal{U}$ for every deterministic initial condition. In the definition of the cost functional, the Wiener basis does not vary any more. The corresponding value function

$$[0,T] \times \mathcal{C} \ni (t_0,\varphi) \rightarrow \inf\bigl\{J(t_0,\varphi,u)) \mid u \in \mathcal{U}\bigr\}$$

is identical to the function V determined by (3.4). By abuse of notation, we write $J(t_0,\varphi,u)$ for $J(t_0,\varphi,(W,u))$. We next state some important properties of the value function.

Proposition 3.1. *Assume (A1)–(A3). Then the value function V is bounded and Lipschitz continuous in the segment variable uniformly in the time variable. More precisely, there is $L_V > 0$ such that for all $t_0 \in [0,T]$, $\varphi,\tilde{\varphi} \in \mathcal{C}$,*

$$|V(t_0,\varphi)| \;\leq\; K(T{+}1), \qquad |V(t_0,\varphi) - V(t_0,\tilde{\varphi})| \;\leq\; L_V\, \|\varphi - \tilde{\varphi}\|.$$

The constant L_V need not be greater than $3L(T+1)\exp(3T(T+4d_1)L^2)$. Moreover, V satisfies Bellman's Principle of Dynamic Programming, that is, for all $t \in [0,T-t_0]$,

$$V(t_0,\varphi) \;=\; \inf_{u\in\mathcal{U}} \mathbf{E}\left(\int_0^t f\bigl(t_0{+}s, X_s^u, u(s)\bigr)ds \;+\; V\bigl(t_0{+}t, X_t^u\bigr)\right),$$

where X^u is the solution to Equation (3.2) under control process u with initial condition (t_0,φ).

Proof. For the boundedness and Lipschitz continuity of V see Proposition A.1, for the Bellman Principle see Theorem A.1 in Appendix A.1, where we set $\tilde{r} := r$, $\tilde{b} := b$ and so on. Notice that the Hölder continuity in time of the coefficients b, σ, f as stipulated in Assumption (A3) is not needed in the proofs. $\square$

The value function V has some regularity in the time variable, too. It is Hölder continuous in time with parameter α for any $\alpha \in (0, \frac{1}{2}]$ provided the initial segment is at least α-Hölder continuous. Notice that the coefficients b, σ, f need not be Hölder continuous in time. Except for the role of the initial segment, statement and proof of Proposition 3.2 are analogous to the non-delay case, see Krylov (1980: p. 167), for example.

Proposition 3.2. *Assume (A1) – (A3). Let $\varphi \in \mathcal{C}$. If φ is α-Hölder continuous with Hölder constant not greater than L_H, then the function $V(.,\varphi)$ is Hölder continuous; that is, there is a constant $\tilde{L}_V > 0$ depending only on L_H, K, T and the dimensions such that for all $t_0, t_1 \in [0, T]$,*

$$|V(t_0,\varphi) - V(t_1,\varphi)| \;\leq\; \tilde{L}_V \left(|t_1 - t_0|^{\alpha} \vee \sqrt{|t_1 - t_0|}\right).$$

Proof. Let $\varphi \in \mathcal{C}$ be α-Hölder continuous with Hölder constant not greater than L_H. Without loss of generality, we suppose that $t_1 = t_0 + h$ for some $h > 0$. We may also suppose $h \leq \frac{1}{2}$, because we can choose $\tilde{L}_V$ greater than $4K(T+1)$ so that the asserted inequality certainly holds for $|t_0 - t_1| > \frac{1}{2}$. By Bellman's Principle as stated in Proposition 3.1, we see that

$$\begin{aligned}
&|V(t_0,\varphi) - V(t_1,\varphi)| \;=\; |V(t_0,\varphi) - V(t_0+h,\varphi)| \\
=\;& \Big|\inf_{u\in\mathcal{U}} \mathbf{E}\left(\int_0^h f\big(t_0+s, X_s^u, u(s)\big)ds \;+\; V(t_0+h, X_h^u)\right) \;-\; V(t_0+h,\varphi)\Big| \\
\leq\;& \sup_{u\in\mathcal{U}} \mathbf{E}\left(\int_0^h \big|f\big(t_0+s, X_s^u, u(s)\big)\big|ds\right) \;+\; \sup_{u\in\mathcal{U}} \mathbf{E}\left(\big|V(t_0+h, X_h^u) - V(t_0+h,\varphi)\big|\right) \\
\leq\;& K\,h \;+\; \sup_{u\in\mathcal{U}} L_V\, \mathbf{E}\left(\|X_h^u - \varphi\|\right),
\end{aligned}$$

where K is the constant from Assumption (A2) and L_V the Lipschitz constant for V in the segment variable according to Proposition 3.1. We notice that $\varphi = X_0^u$ for all $u \in \mathcal{U}$ since X^u is the solution to Equation (3.2) under control u with initial condition (t_0, φ). By Assumption (A2), Hölder's inequality, Doob's maximal inequality and Itô's isometry, for arbitrary $u \in \mathcal{U}$ it holds that

$$\begin{aligned}
&\mathbf{E}\left(\|X_h^u - \varphi\|\right) \\
\leq\;& \sup_{t\in[-r,-h]} |\varphi(t+h) - \varphi(t)| \;+\; \sup_{t\in[-h,0]} |\varphi(0) - \varphi(t)| \;+\; \mathbf{E}\left(\int_0^h \big|b\big(t_0+s, X_s^u, u(s)\big)\big|ds\right) \\
&+\; \mathbf{E}\left(\sup_{t\in[0,h]} \left|\int_0^t \sigma\big(t_0+s, X_s^u, u(s)\big)dW(s)\right|^2\right)^{\frac{1}{2}} \\
\leq\;& 2L_H\, h^{\alpha} \;+\; K\,h \;+\; 4K\, d_1\sqrt{h}.
\end{aligned}$$

Putting everything together, we obtain the assertion. □

From the proof of Proposition 3.2 we see that the time regularity of the value function V is independent of the time regularity of the coefficients b, σ, f; it is always $\frac{1}{2}$-Hölder provided the initial segment is at least that regular.

3.2 First discretisation step: Euler-Maruyama scheme

In this section, the dynamics and the cost functional of the original control problem are discretised in time and segment space. More precisely, we define a sequence of approximating control problems where the coefficients of the dynamics, the cost rate, and the terminal cost are piecewise constant functions of the time and segment variable, while the dependence on the strategies remains the same as in the original problem. We will obtain an upper bound on the approximation error which is uniform over all initial segments of a given Hölder continuity.

Let $N \in \mathbb{N}$. In order to construct the N-th approximating control problem, set $h_N := \frac{r}{N}$, and define $\lfloor . \rfloor_N$ by $\lfloor t \rfloor_N := h_N \lfloor \frac{t}{h_N} \rfloor$, where $\lfloor . \rfloor$ is the usual Gauss bracket, that is, $\lfloor t \rfloor$ is the integer part of the real number t. Set $T_N := \lfloor T \rfloor_N$ and $I_N := \{k\, h_N \mid k \in \mathbb{N}_0\} \cap [0, T_N]$. As T is the time horizon for the original control problem, T_N will be the time horizon for the N-th approximating problem. The set I_N is the time grid of discretisation degree N. Denote by Lin_N the operator $\mathcal{C} \to \mathcal{C}$ which maps a function in $\mathcal{C}$ to its piecewise linear interpolation on the grid $\{k\, h_N \mid k \in \mathbb{Z}\} \cap [-r, 0]$.

We want to express the dynamics and the cost functional of the approximating problems in the same form as those of the original problem, so that the Principle of Dynamic Programming as stated in Appendix A.1 can be readily applied; see Propositions 3.5 and 3.6 in Section 3.3. To this end, the segment space has to be enlarged according to the discretisation degree N. Denote by $\mathcal{C}_N$ the space $\mathbf{C}([-r-h_N, 0], \mathbb{R}^d)$ of $\mathbb{R}^d$-valued continuous functions living on the interval $[-r-h_N, 0]$. For a continuous function or a continuous process Z defined on the time interval $[-r-h_N, \infty)$, let $\Pi_N(Z)(t)$ denote the segment of Z at time $t \geq 0$ of length $r+h_N$, that is, $\Pi_N(Z)(t)$ is the function $[-r-h_N, 0] \ni s \mapsto Z(t+s)$.

Given $t_0 \geq 0$, $\psi \in \mathcal{C}_N$ and $u \in \mathcal{U}$, we define the Euler-Maruyama approximation $Z = Z^{N,t_0,\psi,u}$ of degree N of the solution X to Equation (3.2) under control process u with initial condition (t_0, ψ) as the solution to

$$Z(t) = \begin{cases} \psi(0) & + \int_0^t b_N\big(t_0+s, \Pi_N(Z)(s), u(s)\big)ds \\ & + \int_0^t \sigma_N\big(t_0+s, \Pi_N(Z)(s), u(s)\big)dW(s), \quad t > 0, \\ \psi(t), & t \in [-r-h_N, 0], \end{cases} \tag{3.5}$$

where the coefficients b_N, σ_N are given by

$$\begin{aligned} b_N(t, \psi, \gamma) &:= b\left(\lfloor t \rfloor_N, \mathrm{Lin}_N\big([-r, 0] \ni s \mapsto \psi(s+\lfloor t \rfloor_N - t)\big), \gamma\right), \\ \sigma_N(t, \psi, \gamma) &:= \sigma\left(\lfloor t \rfloor_N, \mathrm{Lin}_N\big([-r, 0] \ni s \mapsto \psi(s+\lfloor t \rfloor_N - t)\big), \gamma\right), \quad t \geq 0,\ \psi \in \mathcal{C}_N,\ \gamma \in \Gamma. \end{aligned}$$

Thus, $b_N(t, \psi, \gamma)$ and $\sigma_N(t, \psi, \gamma)$ are calculated by evaluating the corresponding coefficients b and σ at $(\lfloor t \rfloor_N, \hat{\varphi}, \gamma)$, where $\hat{\varphi}$ is the segment in $\mathcal{C}$ which arises from the piecewise linear interpolation with mesh size $\frac{r}{N}$ of the restriction of ψ to the interval $[\lfloor t \rfloor_N - t - r, \lfloor t \rfloor_N - t]$. Notice that the control action γ remains unchanged.

Assumptions (A1) – (A3) guarantee that, given any control process $u \in \mathcal{U}$, Equation (3.5) has a unique solution for each initial condition $(t_0, \psi) \in [0, \infty) \times \mathcal{C}_N$. Thus, the process $Z = Z^{N,t_0,\psi,u}$ of discretisation degree N is well defined. Notice that the approximating coefficients b_N, σ_N are still Lipschitz continuous in the segment variable uniformly in the time and control variables, although they are only piecewise continuous in time.

Define the cost functional $J_N\colon [0,T_N]\times\mathcal{C}_N\times\mathcal{U}\to\mathbb{R}$ of discretisation degree N by

(3.6)
$$J_N(t_0,\psi,u) := \mathbf{E}\left(\int_0^{T_N-t_0} f_N\big(t_0{+}s,\Pi_N(Z)(s),u(s)\big)ds \;+\; g_N\big(\Pi_N(Z)(T_N{-}t_0)\big)\right),$$

where f_N, g_N are given by

$$\begin{aligned} f_N(t,\psi,\gamma) &:= f\left(\lfloor t\rfloor_N, \mathrm{Lin}_N\big([-r,0]\ni s\mapsto\psi(s{+}\lfloor t\rfloor_N{-}t)\big),\gamma\right), \\ g_N(\psi) &:= g\left(\mathrm{Lin}_N\big(\psi_{|[-r,0]}\big)\right), \qquad\qquad t\geq 0,\ \psi\in\mathcal{C}_N,\ \gamma\in\Gamma. \end{aligned}$$

As b_N, σ_N above, f_N, g_N are Lipschitz continuous in the segment variable (uniformly in time and control) under the supremum norm on $\mathcal{C}_N$. The value function V_N corresponding to (3.5) and (3.6) is the function $[0,T_N]\times\mathcal{C}_N\to\mathbb{R}$ determined by

$$V_N(t_0,\psi) := \inf\big\{J_N(t_0,\psi,u)\mid u\in\mathcal{U}\big\}. \tag{3.7}$$

If $t_0\in I_N$, then $\lfloor t_0{+}s\rfloor_N = t_0+\lfloor s\rfloor_N$ for all $s\geq 0$. Thus, the solution Z to Equation (3.5) under control process $u\in\mathcal{U}$ with initial condition $(t_0,\psi)\in I_N\times\mathcal{C}_N$ satisfies

$$\begin{aligned} Z(t) = \;& \psi(0) + \int_0^t b\big(t_0{+}\lfloor s\rfloor_N, \mathrm{Lin}_N\big(Z_{\lfloor s\rfloor_N}\big), u(s)\big)ds \\ &+ \int_0^t \sigma\big(t_0{+}\lfloor s\rfloor_N, \mathrm{Lin}_N\big(Z_{\lfloor s\rfloor_N}\big), u(s)\big)dW(s) \quad \text{for all } t\geq 0. \end{aligned}\tag{3.8}$$

Moreover, $(Z(t))_{t\geq 0}$ depends on the initial segment ψ only through the restriction of ψ to the interval $[-r,0]$. In analogy, whenever $t_0\in I_N$, the cost functional J_N takes on the form

(3.9)
$$J_N(t_0,\psi,u) \;=\; \mathbf{E}\left(\int_0^{T_N-t_0} f\big(t_0{+}\lfloor s\rfloor_N, \mathrm{Lin}_N\big(Z_{\lfloor s\rfloor_N}\big), u(s)\big)ds \;+\; g\big(\mathrm{Lin}_N\big(Z_{T_N-t_0}\big)\big)\right).$$

Hence, if $t_0\in I_N$, then $J_N(t_0,\psi,u) = J_N(t_0,\psi_{|[-r,0]},u)$ for all $\psi\in\mathcal{C}_N$, $u\in\mathcal{U}$; that is, $J_N(t_0,.,.)$ coincides with its projection onto $\mathcal{C}\times\mathcal{U}$. Consequently, if $t_0\in I_N$, then $V_N(t_0,\psi) = V_N(t_0,\psi_{|[-r,0]})$ for all $\psi\in\mathcal{C}_N$; that is, $V_N(t_0,.)$ can be interpreted as a function with domain $\mathcal{C}$ instead of $\mathcal{C}_N$. If $t_0\in I_N$, by abuse of notation, we will write $V_N(t_0,.)$ also for this function. Notice that, as a consequence of Equations (3.8) and (3.9), in this case we have $V_N(t_0,\varphi) = V_N(t_0,\mathrm{Lin}_N(\varphi))$ for all $\varphi\in\mathcal{C}$.

By Proposition 3.2, we know that the original value function V is Hölder continuous in time provided the initial segment is Hölder continuous. It is therefore enough to compare V and V_N on the grid $I_N\times\mathcal{C}$. This is the content of the next two statements. Again, the order of the error will be uniform only over those initial segments which are α-Hölder continuous for some $\alpha>0$; the constant in the error bound also depends on the Hölder constant of the initial segment. We start with comparing solutions to Equations (3.2) and (3.5) for initial times in I_N.

Proposition 3.3. *Assume (A1) – (A3). Let $\varphi\in\mathcal{C}$ be Hölder continuous with parameter $\alpha>0$ and Hölder constant not greater than L_H. Then there is a constant C depending only on α, L_H, L, K, T and the dimensions such that for all $N\in\mathbb{N}$ with $N\geq 2r$, all $t_0\in I_N$, $u\in\mathcal{U}$ it holds that*

$$\mathbf{E}\left(\sup_{t\in[-r,T]}|X(t)-Z^N(t)|\right) \;\leq\; C\left(h_N^{\alpha}\vee\sqrt{h_N\ln\big(\tfrac{1}{h_N}\big)}\right),$$

where X is the solution to Equation (3.2) *under control process u with initial condition* (t_0, φ) *and Z^N is the solution to Equation* (3.5) *of discretisation degree N under u with initial condition* (t_0, ψ) *with $\psi \in \mathcal{C}_N$ being such that $\psi_{|[-r,0]} = \varphi$.*

Proof. Notice that $h_N \leq \frac{1}{2}$ since $N \geq 2r$, and observe that $Z := Z^N$ as defined in the assertion satisfies Equation (3.8), as the initial time t_0 lies on the grid I_N. Moreover, Z depends on the initial segment ψ only through $\psi_{|[-r,0]} = \varphi$. Using Hölder's inequality, Doob's maximal inequality, Itô's isometry, Assumption (A3), and Fubini's theorem we find that

$$
\begin{aligned}
\mathbf{E}\left(\sup_{t\in[-r,T]} |X(t) - Z(t)|^2\right) \quad &= \quad \mathbf{E}\left(\sup_{t\in[0,T]} |X(t) - Z(t)|^2\right) \\
\leq \quad & 2T\,\mathbf{E}\left(\int_0^T \left|b\big(t_0{+}s, X_s, u(s)\big) - b\big(t_0{+}\lfloor s\rfloor_N, \mathrm{Lin}_N(Z_{\lfloor s\rfloor_N}), u(s)\big)\right|^2 ds\right) \\
& + \; 8d_1\,\mathbf{E}\left(\int_0^T \left|\sigma\big(t_0{+}s, X_s, u(s)\big) - \sigma\big(t_0{+}\lfloor s\rfloor_N, \mathrm{Lin}_N(Z_{\lfloor s\rfloor_N}), u(s)\big)\right|^2 ds\right) \\
\leq \quad & 4T\,\mathbf{E}\left(\int_0^T \left|b\big(t_0{+}\lfloor s\rfloor_N, X_s, u(s)\big) - b\big(t_0{+}\lfloor s\rfloor_N, \mathrm{Lin}_N(Z_{\lfloor s\rfloor_N}), u(s)\big)\right|^2 ds\right) \\
& + \; 16d_1\,\mathbf{E}\left(\int_0^T \left|\sigma\big(t_0{+}\lfloor s\rfloor_N, X_s, u(s)\big) - \sigma\big(t_0{+}\lfloor s\rfloor_N, \mathrm{Lin}_N(Z_{\lfloor s\rfloor_N}), u(s)\big)\right|^2 ds\right) \\
& + \; 4T(T{+}\,4d_1)L^2 h_N \\
\leq \quad & 4(T{+}\,4d_1)L^2\left(T\,h_N + \int_0^T \mathbf{E}\left(\|X_s - \mathrm{Lin}_N(Z_{\lfloor s\rfloor_N})\|^2\right) ds\right) \\
\leq \quad & 4(T{+}\,4d_1)L^2\left(T\,h_N + \; 3\int_0^T \left(\mathbf{E}\left(\|X_s - X_{\lfloor s\rfloor_N}\|^2\right) + \mathbf{E}\left(\|Z_{\lfloor s\rfloor_N} - \mathrm{Lin}_N(Z_{\lfloor s\rfloor_N})\|^2\right)\right) ds\right) \\
& + \; 12(T{+}\,4d_1)L^2\int_0^T \mathbf{E}\left(\|X_{\lfloor s\rfloor_N} - Z_{\lfloor s\rfloor_N}\|^2\right) ds \\
\leq \quad & 4T(T{+}\,4d_1)L^2\left(h_N + \; 18L_H^2\,h_N^{2\alpha} \; + \; 18C_{2,T}\,h_N\ln\big(\tfrac{1}{h_N}\big)\right) \\
& + \; 12(T{+}\,4d_1)L^2\int_0^T \mathbf{E}\left(\sup_{t\in[-r,s]} |X(t) - Z(t)|^2\right) ds.
\end{aligned}
$$

Applying Gronwall's lemma, we obtain the assertion. In the last step of the above estimate Lemma A.1 from Appendix A.2 and the Hölder continuity of φ have both been used twice. Firstly, to get for all $s \in [0,T]$,

$$
\begin{aligned}
& \mathbf{E}\left(\|X_s - X_{\lfloor s\rfloor_N}\|^2\right) \\
\leq \quad & 2\,\mathbf{E}\left(\sup_{t,\tilde{t}\in[-r,0],|t-\tilde{t}|\leq h_N} |\varphi(t) - \varphi(\tilde{t})|^2\right) \; + \; 2\,\mathbf{E}\left(\sup_{t,\tilde{t}\in[0,T],|t-\tilde{t}|\leq h_N} |X(t) - X(\tilde{t})|^2\right) \\
\leq \quad & 2L_H^2\,h_N^{2\alpha} \; + \; 2C_{2,T}\,h_N\ln\big(\tfrac{1}{h_N}\big).
\end{aligned}
$$

Secondly, to obtain

$$
\begin{aligned}
&\mathbf{E}\left(\left\|Z_{\lfloor s\rfloor_N} - \mathrm{Lin}_N\big(Z_{\lfloor s\rfloor_N}\big)\right\|^2\right) \;=\; \mathbf{E}\left(\sup_{t\in[\lfloor s\rfloor_N - r, \lfloor s\rfloor_N]} \left|Z(t) - \mathrm{Lin}_N\big(Z_{\lfloor s\rfloor_N}\big)(t)\right|^2\right)\\
\leq\;& 2\,\mathbf{E}\left(\sup_{t\in[-r,0)} \left|\varphi(t) - \varphi\big(\lfloor t\rfloor_N\big)\right|^2 + \left|\varphi(t) - \varphi\big(\lfloor t\rfloor_N + h_N\big)\right|^2\right)\\
&+\; 2\,\mathbf{E}\left(\sup_{t\in[0,s)} \left|Z(t) - Z\big(\lfloor t\rfloor_N\big)\right|^2 + \left|Z(t) - Z\big(\lfloor t\rfloor_N + h_N\big)\right|^2\right)\\
\leq\;& 4L_H^2\, h_N^{2\alpha} \;+\; 4\,\mathbf{E}\left(\sup_{t,\tilde t\in[0,s],|t-\tilde t|\leq h_N} |Z(t) - Z(\tilde t)|^2\right)\\
\leq\;& 4L_H^2\, h_N^{2\alpha} \;+\; 4C_{2,T}\, h_N \ln\big(\tfrac{1}{h_N}\big) \quad \text{for all } s\in[0,T].
\end{aligned}
$$

□

The order of the approximation error obtained in Proposition 3.3 for the underlying dynamics carries over to the approximation of the corresponding value functions. This works thanks to the Lipschitz continuity of the cost rate and terminal cost in the segment variable, the bound on the moments of the modulus of continuity from Lemma A.1 in Appendix A.2, and the fact that the error bound in Proposition 3.3 is uniform over all strategies.

Theorem 3.1. *Assume (A1)–(A3). Let $\varphi \in \mathcal{C}$ be Hölder continuous with parameter $\alpha > 0$ and Hölder constant not greater than L_H. Then there is a constant $\tilde C$ depending only on α, L_H, L, K, T and the dimensions such that for all $N \in \mathbb{N}$ with $N \geq 2r$, all $t_0 \in I_N$ it holds that*

$$
\left|V(t_0,\varphi) - V_N(t_0,\varphi)\right| \;\leq\; \sup_{u\in\mathcal{U}} |J(t_0,\varphi,u) - J_N(t_0,\psi,u)| \;\leq\; \tilde C\left(h_N^{\alpha} \vee \sqrt{h_N \ln(\tfrac{1}{h_N})}\right),
$$

where $\psi \in \mathcal{C}_N$ is such that $\psi_{|[-r,0]} = \varphi$.

Proof. To verify the first inequality, we distinguish the cases $V(t_0,\varphi) > V_N(t_0,\varphi)$ and $V(t_0,\varphi) < V_N(t_0,\varphi)$. First suppose that $V(t_0,\varphi) > V_N(t_0,\varphi)$. Then for each $\varepsilon \in (0,1]$ we find a strategy $u^\varepsilon \in \mathcal{U}$ such that $V_N(t_0,\varphi) \geq J_N(t_0,\varphi,u^\varepsilon) - \varepsilon$. Since $V(t_0,\varphi) \leq J(t_0,\varphi,u)$ for all $u \in \mathcal{U}$ by definition, it follows that

$$
\begin{aligned}
&|V(t_0,\varphi) - V_N(t_0,\varphi)| \;=\; V(t_0,\varphi) - V_N(t_0,\varphi) \;\leq\; J(t_0,\varphi,u^\varepsilon) - J_N(t_0,\varphi,u^\varepsilon) \;+\; \varepsilon\\
\leq\;& \sup_{u\in\mathcal{U}} |J(t_0,\varphi,u) - J_N(t_0,\psi,u)| \;+\; \varepsilon.
\end{aligned}
$$

Sending ε to zero, we obtain the asserted inequality provided that $V(t_0,\varphi) > V_N(t_0,\varphi)$. If, on the other hand, $V(t_0,\varphi) < V_N(t_0,\varphi)$, then we choose a sequence of minimising strategies $u^\varepsilon \in \mathcal{U}$ such that $V(t_0,\varphi) \geq J(t_0,\varphi,u^\varepsilon) - \varepsilon$, notice that $|V(t_0,\varphi) - V_N(t_0,\varphi)| = V_N(t_0,\varphi) - V(t_0,\varphi)$ and obtain the asserted inequality as in the first case.

Now, let $u \in \mathcal{U}$ be any control process. Let X be the solution to Equation (3.2) under u with initial condition (t_0,φ) and $Z = Z^N$ be the solution to Equation (3.5) under u with initial condition (t_0, ψ). Using Assumption (A2) and the hypothesis that $t_0 \in I_N$, we get

$$|J(t_0,\varphi,u) - J_N(t_0,\psi,u)| \;\leq\; K\,|T-T_N| \;+\; \mathbf{E}\left(\left|g\big(\mathrm{Lin}_N\big(Z_{T_N-t_0}\big)\big) - g\big(X_{T-t_0}\big)\right|\right)$$
$$+\;\mathbf{E}\left(\int_0^{T_N-t_0} \left|f\big(t_0+\lfloor s\rfloor_N, \mathrm{Lin}_N\big(Z_{\lfloor s\rfloor_N}\big), u(s)\big) - f\big(t_0+s, X_s, u(s)\big)\right| ds\right).$$

Recall that $|T-T_N| = T-\lfloor T\rfloor_N \leq h_N$. Hence, $K\,|T-T_N| \leq K\,h_N$. Now, using Assumption (A3), we see that

$$\mathbf{E}\left(\left|g\big(\mathrm{Lin}_N\big(Z_{T_N-t_0}\big)\big) - g\big(X_{T-t_0}\big)\right|\right)$$
$$\leq\; L\Big(\mathbf{E}\big(\|Z_{T_N-t_0}-X_{T_N-t_0}\|\big) + \mathbf{E}\big(\|\,\mathrm{Lin}_N\big(Z_{T_N-t_0}\big)-Z_{T_N-t_0}\|\big) + \mathbf{E}\big(\|X_{T_N-t_0}-X_{T-t_0}\|\big)\Big)$$
$$\leq\; L\left(C\left(h_N^{\alpha} \vee \sqrt{h_N \ln(\tfrac{1}{h_N})}\right) \;+\; 3L_H\,h_N^{\alpha} + 3C_{1,T}\sqrt{h_N \ln(\tfrac{1}{h_N})}\right),$$

where C is a constant as in Proposition 3.3 and $C_{1,T}$ is a constant as in Lemma A.1 in Appendix A.2. Notice that $(X(t))_{t\geq 0}$ as well as $(Z(t))_{t\geq 0}$ are Itô diffusions with coefficients bounded by the constant K from Assumption (A2). In the same way, also using the Hölder continuity of f in time and recalling that $|s-\lfloor s\rfloor_N| \leq h_N$ for all $s \geq 0$, we see that

$$\mathbf{E}\left(\int_0^{T_N-t_0} \left|f\big(t_0+\lfloor s\rfloor_N, \mathrm{Lin}_N\big(Z_{\lfloor s\rfloor_N}\big), u(s)\big) - f\big(t_0+s, X_s, u(s)\big)\right| ds\right)$$
$$\leq\; L\,(T_N-t_0)\left(\sqrt{h_N} \;+\; 3C_{1,T}\sqrt{h_N \ln(\tfrac{1}{h_N})} \;+\; \big(C+3L_H\big)\left(h_N^{\alpha} \vee \sqrt{h_N \ln(\tfrac{1}{h_N})}\right)\right).$$

Putting the three estimates together, we obtain the assertion. □

In virtue of Theorem 3.1, we can replace the original control problem of Section 3.1 with the sequence of approximating control problems defined above. The error between the problem of degree N and the original problem in terms of the difference between the corresponding value functions V and V_N is not greater than a multiple of $(\frac{r}{N})^{\alpha}$ for α-Hölder continuous initial segments if $\alpha \in (0, \frac{1}{2})$, where the proportionality factor is affine in the Hölder constant; it is less than a multiple of $\sqrt{\ln(N)/N}$ if $\alpha \geq \frac{1}{2}$.

From the proofs of Proposition 3.3 and Theorem 3.1 it is clear that the coefficients b, σ, f of the original problem, instead of being $\frac{1}{2}$-Hölder continuous in time as postulated by Assumption (A3), need only satisfy a bound of the form $\sqrt{|t-s|\ln(\frac{1}{|t-s|})}$, $t,s \in [0,T]$ with $|t-s|$ small, for the error estimates to hold.

Let us assume for a moment that $\sigma \equiv 0$, that is, the diffusion coefficient σ is zero. Then Equation (3.2) becomes a random ordinary differential equation. It is still "random", because the admissible strategies are still Γ-valued stochastic processes adapted to the given Wiener filtration. The minimal costs $V(t_0,\varphi)$ for any deterministic initial condition $(t_0,\varphi) \in [0,T] \times \mathcal{C}$, however, can be arbitrarily well approximated by using deterministic strategies, that is, Borel measurable functions $[0,\infty) \to \Gamma$.

In case $\sigma \equiv 0$, the optimal control problem of Section 3.1 is therefore equivalent to the purely deterministic control problem where minimisation is performed with respect to all deterministic strategies. The cost functional of the deterministic problem is again given by (3.3), but without expectation. The same observation applies to the control problems of degree N, $N \in \mathbb{N}$, introduced in this section. In the sequel, we will not always distinguish between a control problem with zero diffusion matrix and the corresponding purely deterministic problem. If the diffusion coefficient σ is zero and the coefficients b, f are Lipschitz continuous in time, then the error between the value functions V and V_N is of order $\frac{r}{N}$ for all Lipschitz continuous initial segments, as one would expect from the classical Euler scheme.

Corollary 3.1. *Assume (A1) – (A3). Assume in addition that σ is equal to zero and that b, f are Lipschitz continuous also in the time variable with Lipschitz constant not greater than L. Let $\varphi \in \mathcal{C}$ be Hölder continuous with parameter $\alpha \in (0,1]$ and Hölder constant not greater than L_H. Then there is a constant $\tilde{C}$ depending only on L_H, L, K, T such that for all $N \in \mathbb{N}$ with $N \geq r$, all $t_0 \in I_N$ it holds that*

$$\big|V(t_0,\varphi) - V_N(t_0,\varphi)\big| \;\leq\; \tilde{C}\,(h_N^{\alpha} \vee h_N)\,,$$

where $\psi \in \mathcal{C}_N$ is such that $\psi_{|[-r,0]} = \varphi$.

Although we obtain an error bound for the approximation of V by the sequence of value functions $(V_N)_{N\in\mathbb{N}}$ only for Hölder continuous initial segments, the proofs of Proposition 3.3 and Theorem 3.1 show that pointwise convergence of the value functions holds true for all initial segments $\varphi \in \mathcal{C}$. Recall that a function $\varphi : [-r,0] \to \mathbb{R}^d$ is continuous if and only if $\sup_{t,s\in[-r,0],|t-s|\leq h} |\varphi(t) - \varphi(s)|$ tends to zero as $h \searrow 0$. Let us record the result for the value functions.

Corollary 3.2. *Assume (A1) – (A3). Then for all $(t_0,\varphi) \in [0,T] \times \mathcal{C}$,*

$$\big|V(t_0,\varphi) - V_N(\lfloor t_0 \rfloor_N,\varphi)\big| \;\overset{N\to\infty}{\longrightarrow}\; 0.$$

Similarly to the value function of the original problem, also the function $V_N(t_0,.)$ is Lipschitz continuous in the segment variable uniformly in $t_0 \in I_N$ with Lipschitz constant not depending on the discretisation degree N. Since $t_0 \in I_N$, we may interpret $V_N(t_0,.)$ as a function defined on $\mathcal{C}$.

Proposition 3.4. *Assume (A1) – (A3). Let V_N be the value function of discretisation degree N. Then $|V_N|$ is bounded by $K(T+1)$. Moreover, if $t_0 \in I_N$, then $V_N(t_0,.)$ as a function of $\mathcal{C}$ satisfies the following Lipschitz condition:*

$$|V_N(t_0,\varphi) - V_N(t_0,\tilde{\varphi})| \;\leq\; 3L(T+1)\exp\big(3T(T+4d_1)L^2\big)\,\|\varphi - \tilde{\varphi}\| \quad \textit{for all } \varphi,\tilde{\varphi} \in \mathcal{C}.$$

Proof. The assertion is again a consequence of Proposition A.1 in Appendix A.1. To see this, set $\tilde{r} := r + h_N$, $\tilde{T} := T_N$, $\tilde{b} := b_N$, $\tilde{\sigma} := \sigma_N$, $\tilde{f} := f_N$, and $\tilde{g} := g_N$. Equation (3.5) then describes the same dynamics as Equation (A.1), $\tilde{J}$ is the same functional as J_N, whence $V_N = \tilde{V}$. The hypotheses of Appendix A.1 are satisfied. Finally, recall that $T_N \leq T$ and that, since $t_0 \in I_N$, $V_N(t_0,\psi)$ depends on $\psi \in \mathcal{C}_N$ only through $\psi_{|[-r,0]}$. □

3.3 Second discretisation step: piecewise constant strategies

In Section 3.2, we have discretised the time as well as the segment space in time. The resulting control problem of discretisation degree $N \in \mathbb{N}$ has dynamics described by Equation (3.5), cost functional J_N defined by (3.6) and value function V_N given by (3.7). Here, we will also approximate the control processes $u \in \mathcal{U}$, which up to now have been those of the original problem, by introducing further control problems defined over sets of piecewise constant strategies. To this end, for $n \in \mathbb{N}$, set

$$\mathcal{U}_n := \big\{u \in \mathcal{U} \mid u(t) \text{ is } \sigma(W(k\tfrac{r}{n}), k \in \mathbb{N}_0)\text{-measurable and } u(t) = u(\lfloor t \rfloor_n),\ t \geq 0\big\}. \tag{3.10}$$

Recall that $\lfloor t \rfloor_n = \frac{r}{n}\lfloor \frac{n}{r} t \rfloor$. Hence, $\mathcal{U}_n$ is the set of all Γ-valued $(\mathcal{F}_t)$-progressively measurable processes which are right-continuous and piecewise constant in time relative to the grid $\{k\frac{r}{n} \mid k \in \mathbb{N}_0\}$ and, in addition, are $\sigma(W(k\frac{r}{n}), k \in \mathbb{N}_0)$-measurable. In particular, if $u \in \mathcal{U}_n$ and $t \geq 0$, then the random variable $u(t)$ can be represented as

$$u(t)(\omega) \;=\; \theta\big(\lfloor \tfrac{n}{r} t \rfloor, W(0)(\omega), \ldots, W(\lfloor \tfrac{n}{r} t \rfloor)(\omega)\big), \quad \omega \in \Omega,$$

where θ is some Γ-valued Borel measurable function depending on u and n. For the purpose of approximating the control problem of degree N, we will use strategies in $\mathcal{U}_{N \cdot M}$ with $M \in \mathbb{N}$. Let us write $\mathcal{U}_{N,M}$ for $\mathcal{U}_{N \cdot M}$.

With the same dynamics and the same performance criterion as before, for each $N \in \mathbb{N}$, we introduce a family of value functions $V_{N,M}$, $M \in \mathbb{N}$, defined on $[0, T_N] \times \mathcal{C}_N$ by setting

$$V_{N,M}(t_0, \psi) := \inf\big\{J_N(t_0, \psi, u) \mid u \in \mathcal{U}_{N,M}\big\}. \tag{3.11}$$

We will refer to $V_{N,M}$ as the value function of degree (N, M). By construction, it holds that $V_N(t_0, \psi) \leq V_{N,M}(t_0, \psi)$ for all $(t_0, \psi) \in [0, T_N] \times \mathcal{C}_N$. Hence, in estimating the approximation error, we only need an upper bound for $V_{N,M} - V_N$.

As with V_N, if the initial time t_0 lies on the grid I_N, then $V_{N,M}(t_0, \psi)$ depends on ψ only through its restriction $\psi_{|[-r,0]} \in \mathcal{C}$ to the interval $[-r, 0]$. We write $V_{N,M}(t_0, .)$ for this function, too. The dynamics and costs, in this case, can again be represented by Equations (3.8) and (3.9), respectively. And again, if $t_0 \in I_N$, we have $V_{N,M}(t_0, \varphi) = V_{N,M}(t_0, \mathrm{Lin}_N(\varphi))$ for all $\varphi \in \mathcal{C}$.

Propositions 3.5 and 3.6 state Bellman's Principle of Dynamic Programming for the value functions V_N and $V_{N,M}$, respectively. The special case when the initial time as well as the time step lie on the grid I_N is given separately, as it is this representation which will be used in the approximation result; see the proof of Theorem 3.2.

Proposition 3.5. *Assume (A1) – (A3). Let $t_0 \in [0, T_N]$, $\psi \in \mathcal{C}_N$. Then for $t \in [0, T_N - t_0]$,*

$$V_N(t_0, \psi) \;=\; \inf_{u \in \mathcal{U}} \mathbf{E}\left(\int_0^t f_N\big(t_0 + s, \Pi_N(Z^u)(s), u(s)\big) ds \;+\; V_N\big(t_0 + t, \Pi_N(Z^u)(t)\big)\right),$$

where Z^u is the solution to Equation (3.5) of degree N under control process u and with initial condition (t_0, ψ). If $t_0 \in I_N$ and $t \in I_N \cap [0, T_N - t_0]$, then

$$V_N(t_0, \varphi) \;=\; \inf_{u \in \mathcal{U}} \mathbf{E}\left(\int_0^t f\big(t_0 + \lfloor s \rfloor_N, \mathrm{Lin}_N(Z^u_{\lfloor s \rfloor_N}), u(s)\big) ds \;+\; V_N\big(t_0 + t, \mathrm{Lin}_N(Z^u_t)\big)\right),$$

where $V_N(t_0, .)$, $V_N(t_0 + t, .)$ are defined as functionals on $\mathcal{C}$, and φ is the restriction of ψ to the interval $[-r, 0]$.

Proof. Apply Theorem A.1 in Appendix A.1. To this end, let $\tilde{\mathcal{U}}$ be the set of strategies $\mathcal{U}$ and set $\tilde{r} := r + h_N$, $\tilde{T} := T_N$, $\tilde{b} := b_N$, $\tilde{\sigma} := \sigma_N$, $\tilde{f} := f_N$, and $\tilde{g} := g_N$. Observe that Equation (3.5) describes the same dynamics as Equation (A.1), that $\tilde{J} = J_N$, whence $V_N = \tilde{V}$, and verify that the hypotheses of Appendix A.1 are satisfied. □

Proposition 3.6. *Assume (A1) – (A3). Let* $t_0 \in [0,T_N]$, $\psi \in \mathcal{C}_N$. *Then for* $t \in I_{N\cdot M} \cap [0,T_N - t_0]$,

$$V_{N,M}(t_0,\psi) \;=\; \inf_{u\in\mathcal{U}_{N,M}} \mathbf{E}\left(\int_0^t f_N\big(t_0+s,\Pi_N(Z^u)(s),u(s)\big)ds \;+\; V_{N,M}\big(t_0+t,\Pi_N(Z^u)(t)\big)\right),$$

where Z^u *is the solution to Equation* (3.5) *of degree* N *under control process* u *and with initial condition* (t_0,ψ). *If* $t_0 \in I_N$ *and* $t \in I_N \cap [0,T_N - t_0]$, *then*

$$V_{N,M}(t_0,\varphi) \;=\; \inf_{u\in\mathcal{U}_{N,M}} \mathbf{E}\left(\int_0^t f\big(t_0+\lfloor s\rfloor_N,\mathrm{Lin}_N(Z^u_{\lfloor s\rfloor_N}),u(s)\big)ds \;+\; V_{N,M}\big(t_0+t,\mathrm{Lin}_N(Z^u_t)\big)\right),$$

where $V_{N,M}(t_0,.)$, $V_{N,M}(t_0+t,.)$ *are defined as functionals on* $\mathcal{C}$, *and* φ *is the restriction of* ψ *to the interval* $[-r,0]$.

Proof. Apply Theorem A.1 of Appendix A.1 as in the proof of Proposition 3.5, except for the fact that we choose $\mathcal{U}_{N,M} = \mathcal{U}_{N\cdot M}$ instead of $\mathcal{U}$ as the set of strategies $\tilde{\mathcal{U}}$. Notice that, by hypothesis, the intermediate time t lies on the grid $I_{N\cdot M}$. □

The next result gives a bound on the order of the global approximation error between the value functions of degree N and (N,M) provided that the local approximation error is of order greater than one in the discretisation step.

Theorem 3.2. *Assume (A1) – (A3). Let* $N,M \in \mathbb{N}$. *Suppose that for some constants* $\hat{K},\delta > 0$ *the following holds: for any* $t_0 \in I_N$, $\varphi \in \mathcal{C}$, $u \in \mathcal{U}$ *there is* $\bar{u} \in \mathcal{U}_{N,M}$ *such that*

$$(*)\qquad \begin{aligned} &\mathbf{E}\left(\int_0^{h_N} f\big(t_0,\mathrm{Lin}_N(\varphi),\bar{u}(s)\big)ds + V_N(t_0+h_N,\bar{Z}_{h_N})\right)\\ \le\; &\mathbf{E}\left(\int_0^{h_N} f\big(t_0,\mathrm{Lin}_N(\varphi),u(s)\big)ds + V_N(t_0+h_N,Z_{h_N})\right) \;+\; \hat{K}\, h_N^{1+\delta}, \end{aligned}$$

where Z *is the solution to Equation* (3.5) *of degree* N *under control process* u, $\bar{Z}$ *the solution to Equation* (3.5) *of degree* N *under* $\bar{u}$, *both with initial condition* (t_0,ψ) *for some* $\psi \in \mathcal{C}_N$ *such that* $\psi_{|[-r,0]} = \varphi$. *Then*

$$\big|V_{N,M}(t_0,\varphi) - V_N(t_0,\varphi)\big| \;\le\; T\,\hat{K}\,h_N^{\delta} \quad \text{for all } t_0 \in I_N,\ \varphi \in \mathcal{C}.$$

Proof. Let $N,M \in \mathbb{N}$. Recall that $V_{N,M} \ge V_N$ by construction. It is therefore enough to prove the upper bound for $V_{N,M} - V_N$. Suppose Condition $(*)$ is fulfilled for N,M and some constants $\hat{K},\delta > 0$. Observe that $V_N(T_N,.) = g(\mathrm{Lin}_N(.)) = V_{N,M}(T_N,.)$.

Let $t_0 \in I_N \setminus \{T_N\}$. Let $\varphi \in \mathcal{C}$, and choose any $\psi \in \mathcal{C}_N$ such that $\psi_{|[-r,0]} = \varphi$. Given $\varepsilon > 0$, in virtue of Proposition 3.5, we find a control process $u \in \mathcal{U}$ such that

$$V_N(t_0,\varphi) \;\ge\; \mathbf{E}\left(\int_0^{h_N} f\big(t_0,\mathrm{Lin}_N(\varphi),u(s)\big)ds \;+\; V_N\big(t_0+h_N,\mathrm{Lin}_N(Z_{h_N})\big)\right) \;-\; \varepsilon,$$

where Z is the solution to Equation (3.5) of degree N under control process u with initial condition (t_0,ψ). For this u, choose $\bar{u} \in \mathcal{U}_{N,M}$ according to $(*)$, and let $\bar{Z}$ be the solution to Equation (3.5) of degree N under control process $\bar{u}$ with the same initial condition as for Z. Then, using the above inequality and Proposition 3.6, we see that

$$
\begin{aligned}
& V_{N,M}(t_0,\varphi) - V_N(t_0,\varphi) \\
\leq\ & V_{N,M}(t_0,\varphi) \;-\; \mathbf{E}\left(\int_0^{h_N} f\big(t_0, \mathrm{Lin}_N(\varphi), u(s)\big)ds \;+\; V_N\big(t_0{+}h_N, \mathrm{Lin}_N(Z_{h_N})\big)\right) \;+\; \varepsilon \\
\leq\ & \mathbf{E}\left(\int_0^{h_N} f\big(t_0, \mathrm{Lin}_N(\varphi), \bar{u}(s)\big)ds \;+\; V_{N,M}\big(t_0{+}h_N, \mathrm{Lin}_N(\bar{Z}_{h_N})\big)\right) \;+\; \varepsilon \\
& - \mathbf{E}\left(\int_0^{h_N} f\big(t_0, \mathrm{Lin}_N(\varphi), u(s)\big)ds \;+\; V_N\big(t_0{+}h_N, \mathrm{Lin}_N(Z_{h_N})\big)\right) \\
=\ & \mathbf{E}\left(\int_0^{h_N} f\big(t_0, \mathrm{Lin}_N(\varphi), \bar{u}(s)\big)ds \;+\; V_N\big(t_0{+}h_N, \mathrm{Lin}_N(\bar{Z}_{h_N})\big)\right) \\
& - \mathbf{E}\left(\int_0^{h_N} f\big(t_0, \mathrm{Lin}_N(\varphi), u(s)\big)ds \;+\; V_N\big(t_0{+}h_N, \mathrm{Lin}_N(Z_{h_N})\big)\right) \\
& + \mathbf{E}\left(V_{N,M}\big(t_0{+}h_N, \mathrm{Lin}_N(\bar{Z}_{h_N})\big) - V_N\big(t_0{+}h_N, \mathrm{Lin}_N(\bar{Z}_{h_N})\big)\right) \;+\; \varepsilon \\
\leq\ & \hat{K}\, h_N^{1+\delta} \;+\; \sup_{\tilde{\varphi}\in\mathcal{C}}\big\{V_{N,M}(t_0{+}h_N,\tilde{\varphi}) - V_N(t_0{+}h_N,\tilde{\varphi})\big\} \;+\; \varepsilon,
\end{aligned}
$$

where in the last line Condition $(*)$ has been exploited. Since $\varepsilon > 0$ was arbitrary and neither the first nor the last line of the above inequalities depend on u or $\bar{u}$, it follows that for all $t_0 \in I_N \setminus \{T_N\}$,

$$
\sup_{\varphi\in\mathcal{C}}\big\{V_{N,M}(t_0,\varphi) - V_N(t_0,\varphi)\big\} \quad\leq\quad \hat{K}\, h_N^{1+\delta} \;+\; \sup_{\varphi\in\mathcal{C}}\big\{V_{N,M}(t_0{+}h_N,\varphi) - V_N(t_0{+}h_N,\varphi)\big\}.
$$

Recalling the equality $V_{N,M}(T_N,.) = V_N(T_N,.)$, we conclude that for all $t_0 \in I_N$,

$$
\sup_{\varphi\in\mathcal{C}}\big\{V_{N,M}(t_0,\varphi) - V_N(t_0,\varphi)\big\} \quad\leq\quad \frac{1}{h_N}\,(T_N{-}t_0)\,\hat{K}\, h_N^{1+\delta} \quad\leq\quad T\,\hat{K}\, h_N^{\delta},
$$

which yields the assertion. □

Statement and proof of Theorem 3.2 should be compared to Theorem 7 in Falcone and Rosace (1996). We note, though, that the deterministic analogue of Condition $(*)$ in Theorem 3.2 is weaker than the corresponding conditions (37) and (38) in Falcone and Rosace (1996). In particular, it is not necessary to require that any controlled process Z can be approximated with local error of order $h^{1+\delta}$ by some process $\bar{Z}$ using only control processes which are piecewise constant in time on a grid of width h. In the stochastic case, such a requirement would in general be too strong to be satisfiable.

In order to be able to apply Theorem 3.2, we must check whether and how Condition $(*)$ can be satisfied. Given a grid of width $\frac{r}{N}$ for the discretisation in time and segment space, we would expect the condition to be fulfilled provided we choose the sub-grid for the piecewise constant controls fine enough; that is, the time discretisation of the control processes should be of degree M with M sufficiently big in comparison to N. Indeed, if we choose M of any order greater than three in N, then Condition $(*)$ holds. This is the content of Theorem 3.3. The theorem, in turn, relies on a kind of mean value theorem, due to Krylov, which we cite as Theorem A.2 in Appendix A.3.

Theorem 3.3. *Assume (A1)-(A4). Let $\beta > 3$. Then there is a number $\hat{K} > 0$ depending only on K, r, L, T, the dimensions and β such that Condition $(*)$ in Theorem 3.2 is satisfied with constants $\hat{K}$ and $\delta := \frac{\beta-3}{4}$ for all $N, M \in \mathbb{N}$ such that $N \geq r$ and $M \geq N^{\beta}$.*

Proof. Let $N, M \in \mathbb{N}$ be such that $N \geq r$ and $M \geq N^{\beta}$. Let $t_0 \in I_N$, $\varphi \in \mathcal{C}$. Define the following functions:

$$\begin{aligned}
&\tilde{b}\colon \Gamma \to \mathbb{R}^d, && \tilde{b}(\gamma) := b\big(t_0, \mathrm{Lin}_N(\varphi), \gamma\big),\\
&\tilde{\sigma}\colon \Gamma \to \mathbb{R}^{d\times d_1}, && \tilde{\sigma}(\gamma) := \sigma\big(t_0, \mathrm{Lin}_N(\varphi), \gamma\big),\\
&\tilde{f}\colon \Gamma \to \mathbb{R}, && \tilde{f}(\gamma) := f\big(t_0, \mathrm{Lin}_N(\varphi), \gamma\big),\\
&\tilde{g}\colon \mathbb{R}^d \to \mathbb{R}^d, && \tilde{g}(x) := V_N\big(t_0{+}h_N, \mathrm{Lin}_N(S(\varphi, x))\big),
\end{aligned}$$

where $S(\varphi, x)$ is the function in $\mathcal{C}$ given by

$$S(\varphi, x)\colon [-r, 0] \ni s \mapsto \begin{cases} \varphi(s{+}h_N) & \text{if } s \in [-r, -h_N],\\ \varphi(0) + \frac{s+h_N}{h_N}\, x & \text{if } s \in (-h_N, 0]. \end{cases}$$

As a consequence of Assumption (A4), $\tilde{b}$, $\tilde{\sigma}$, $\tilde{f}$ as just defined are continuous functions on (Γ, ρ). By Assumption (A2), $|\tilde{b}|$, $|\tilde{\sigma}|$, $|\tilde{f}|$ are all bounded by K. As a consequence of Proposition 3.4, the function $\tilde{g}$ is Lipschitz continuous and for the Lipschitz constant we have

$$\sup_{x,y\in\mathbb{R}^d, x\neq y} \frac{|\tilde{g}(x) - \tilde{g}(y)|}{|x-y|} \;\leq\; 3L(T{+}1)\exp\big(3T(T{+}4d_1)L^2\big).$$

Let $u \in \mathcal{U}$, and let Z^u be the solution to Equation (3.5) of degree N under control process u with initial condition (t_0, ψ) for some $\psi \in \mathcal{C}_N$ such that $\psi_{|[-r,0]} = \varphi$. As Z also satisfies Equation (3.8), we see that

$$Z^u(t) - \varphi(0) \;=\; \int_0^t \tilde{b}\big(u(s)\big)ds \;+\; \int_0^t \tilde{\sigma}\big(u(s)\big)dW(s) \quad \text{for all } t \in [0, h_N].$$

By Theorem A.2 in Appendix A.3, we find $\bar{u} \in \mathcal{U}_{N,M}$ such that

$$\begin{aligned}
&\mathbf{E}\left(\int_0^{h_N} \tilde{f}\big(\bar{u}(s)\big)ds + \tilde{g}\big(X^{\bar{u}}(h_N)\big)\right) \;-\; \mathbf{E}\left(\int_0^{h_N} \tilde{f}\big(u(s)\big)ds + \tilde{g}\big(Z^u(h_N){-}\varphi(0)\big)\right)\\
\leq\;& \bar{C}(1{+}h_N)\left(\frac{r}{N\cdot M}\right)^{\frac{1}{4}}\left(\left(\frac{r}{N\cdot M}\right)^{\frac{1}{4}}\sup_{\gamma\in\Gamma}|\tilde{f}(\gamma)| \;+\; \sup_{x,y\in\mathbb{R}^d, x\neq y}\frac{|\tilde{g}(x)-\tilde{g}(y)|}{|x-y|}\right),
\end{aligned}$$

where $X^{\bar{u}}$ satisfies

$$X^{\bar{u}}(t) \;=\; \int_0^t \tilde{b}\big(\bar{u}(s)\big)ds \;+\; \int_0^t \tilde{\sigma}\big(\bar{u}(s)\big)dW(s) \quad \text{for all } t \geq 0.$$

Notice that the constant $\bar{C}$ above only depends on K and the dimensions d and d_1. Let $Z^{\bar{u}}$ be the solution to Equation (3.5) of degree N under control process $\bar{u}$ with initial condition (t_0, ψ), where $\psi_{|[-r,0]} = \varphi$ as above. Then, by construction, $Z^{\bar{u}}(t) - \varphi(0) = X^{\bar{u}}(t)$ for all $t \in [0, h_N]$. Set

$$\hat{K} := 2\bar{C}\, r^{-\frac{\beta}{4}}\left(K + 3L(T{+}1)\exp\big(3T(T{+}4d_1)L^2\big)\right).$$

Since $M \geq N^{\beta}$ by hypothesis, $\frac{1+\beta}{4} = 1+\delta > 1$ and $h_N = \frac{r}{N}$, we have

$$r^{\frac{1}{4}}(N\cdot M)^{-\frac{1}{4}} \;\leq\; r^{\frac{1}{4}} \cdot N^{-\frac{1+\beta}{4}} \;=\; r^{-\frac{\beta}{4}} \cdot h_N^{1+\delta}.$$

Recalling the definition of the coefficients $\tilde{b}$, $\tilde{\sigma}$, $\tilde{f}$, $\tilde{g}$, we have thus found a piecewise constant strategy $\bar{u} \in \mathcal{U}_{N,M}$ such that

$$\begin{aligned} &\mathbf{E}\left(\int_0^{h_N} f\big(t_0, \mathrm{Lin}_N(\varphi), \bar{u}(s)\big)ds + V_N\big(t_0{+}h_N, Z^{\bar{u}}_{h_N}\big)\right) \\ \leq\; &\mathbf{E}\left(\int_0^{h_N} f\big(t_0, \mathrm{Lin}_N(\varphi), u(s)\big)ds + V_N\big(t_0{+}h_N, Z^{u}_{h_N}\big)\right) \;+\; \hat{K}\, h_N^{1+\delta}, \end{aligned}$$

where Z^u, $Z^{\bar{u}}$ are the solutions corresponding to u and $\bar{u}$, respectively, as above. □

We note that the constant $\hat{K}$, which appears in Theorem 3.3 and its proof, depends on β only through the factor $r^{-\frac{\beta}{4}}$. Moreover, $\hat{K}$ also depends on the delay length r only through the factor $r^{-\frac{\beta}{4}}$. Theorem 3.2 and Theorem 3.3, together with the above observation, yield the following bound on the difference between the value functions of degree N and degree (N, M), respectively.

Corollary 3.3. *Assume (A1)-(A4). Then there is a positive constant $\bar{K}$ depending only on K, L, T, and the dimensions such that for all $\beta > 3$, all $N \in \mathbb{N}$ with $N \geq r$, all $M \in \mathbb{N}$ with $M \geq N^{\beta}$, all $t_0 \in I_N$, all $\varphi \in \mathcal{C}$ it holds that*

$$\big|V_{N,M}(t_0,\varphi) - V_N(t_0,\varphi)\big| \;\leq\; \bar{K}\, r^{-\frac{\beta}{4}} \left(\frac{r}{N}\right)^{\frac{\beta-3}{4}}.$$

In particular, with $M = \lceil N^{\beta}\rceil$, where $\lceil x\rceil$ is the least integer not smaller than x, the upper bound on the discretisation error can be rewritten as

$$\big|V_{N,\lceil N^{\beta}\rceil}(t_0,\varphi) - V_N(t_0,\varphi)\big| \;\leq\; \bar{K}\, r^{-\frac{\beta}{1+\beta}} \left(\frac{r}{N^{1+\beta}}\right)^{\frac{\beta-3}{4(1+\beta)}}.$$

From Corollary 3.3 we see that, in terms of the total number of time steps $N\lceil N^{\beta}\rceil$, we can achieve any rate of convergence smaller than $\frac{1}{4}$ by choosing the sub-discretisation order β sufficiently large.

When the diffusion coefficient σ is zero or the space of control actions Γ is finite and σ is not directly controlled, then the sub-disretisation degree M may be chosen of an order lower than three in N, and Condition (∗) is still satisfied. For in these special cases, the error bound of Theorem A.2 can be improved on, see Appendix A.3.

Let us first consider the case when $\sigma \equiv 0$, which corresponds to deterministic control problems. To obtain an analogue of Theorem 3.3, we use Lemma A.3 in place of Theorem A.2. The order exponent β must be greater than one, and the order exponent δ in Condition (∗) is taken to be $\frac{\beta-1}{2}$. If, in addition, Γ is finite, instead of Lemma A.3 we invoke Lemma A.4. The analogue of Theorem 3.3 holds true for any $\beta > 0$ and with the choice $\delta := \beta$. These observations in combination with Theorem 3.2 yield the following bounds for deterministic systems on the difference between V_N and $V_{N,M}$; the results are given only for $M = \lceil N^{\beta}\rceil$.

Corollary 3.4. *Assume (A1)-(A4). Assume further that σ is equal to zero. Then there is a positive constant $\bar{K}$ depending only on K, L, T, and the dimension d such that for all $\beta > 1$, all $N \in \mathbb{N}$ with $N \geq r$, all $t_0 \in I_N$, all $\varphi \in \mathcal{C}$ it holds that*

$$\left|V_{N,\lceil N^{\beta}\rceil}(t_0,\varphi) - V_N(t_0,\varphi)\right| \;\leq\; \bar{K}\, r^{-\frac{\beta}{1+\beta}} \left(\frac{r}{N^{1+\beta}}\right)^{\frac{\beta-1}{2(1+\beta)}}.$$

If, in addition, Γ is finite with cardinality N_Γ, then there is a positive constant $\tilde{K}$ depending only on K, L, T such that for all $\beta > 0$, all $N \in \mathbb{N}$ with $N \geq r$, all $t_0 \in I_N$, all $\varphi \in \mathcal{C}$ it holds that

$$\left|V_{N,\lceil N^{\beta}\rceil}(t_0,\varphi) - V_N(t_0,\varphi)\right| \;\leq\; \tilde{K}(1+N_\Gamma)\, r^{-\frac{\beta}{1+\beta}} \left(\frac{r}{N^{1+\beta}}\right)^{\frac{\beta}{1+\beta}}.$$

If the diffusion coefficient σ is not directly controlled, that is, if $\sigma(t,\varphi,\gamma) = \tilde{\sigma}(t,\varphi)$ for some $\tilde{\sigma}$ and all $t \in [0,T]$, $\varphi \in \mathcal{C}$, $\gamma \in \Gamma$, then we may rely on Lemma A.5 in place of Theorem A.2. Observe that the diffusion coefficient for the control problems of degree (N,M) and N, respectively, is constant on time intervals of the form $[(k-1)\frac{r}{N}, k\frac{r}{N})$, $k \in \mathbb{N}$. The order exponent β for the analogue of Theorem 3.3 must be greater than one, and the order exponent δ in Condition $(*)$ is taken to equal $\frac{\beta-1}{2}$. In combination with Theorem 3.2, this implies the following bound.

Corollary 3.5. *Assume (A1)-(A4). Assume in addition that σ does not depend on the control variable and that Γ is finite with cardinality N_Γ. Then there is a positive constant $\tilde{K}$ depending only on K, L, T such that for all $\beta > 1$, all $N \in \mathbb{N}$ with $N \geq r$ and $\lceil N^{\beta}\rceil$ a square number, all $t_0 \in I_N$, all $\varphi \in \mathcal{C}$ it holds that*

$$\left|V_{N,\lceil N^{\beta}\rceil}(t_0,\varphi) - V_N(t_0,\varphi)\right| \;\leq\; \tilde{K}(1+4r\cdot T+N_\Gamma)\, r^{-\frac{\beta}{1+\beta}} \left(\frac{r}{N^{1+\beta}}\right)^{\frac{\beta}{1+\beta}}.$$

The requirement in Corollary 3.5 that $\lceil N^{\beta}\rceil$ be a square number is no serious restriction, as the optimal bound on the total discretisation error will be achieved with $\beta = 2$.

3.4 Bounds on the total error

Here, we put together the error bounds from Sections 3.2 and 3.3 in order to obtain an overall estimate for the rate of convergence, that is, a bound on the discretisation error incurred in passing from the original value function to the value function of degree (N,M). In addition, we address the question of whether and in which sense nearly optimal strategies for the discrete problems can be used as nearly optimal strategies for the original system.

As in Corollary 3.3, we express the error bound in terms of the total number of discretisation steps or, taking into account the presence of the delay length r, in terms of the length of the smallest time step.

Theorem 3.4. *Assume (A1)-(A4). Let $\alpha \in (0,1]$, $L_H > 0$. Then there is a constant $\bar{C}$ depending only on α, L_H, L, K, T and the dimensions such that for all $\beta > 3$, all $N \in \mathbb{N}$ with $N \geq 2r$, all $t_0 \in I_N$, all α-Hölder continuous $\varphi \in \mathcal{C}$ with Hölder constant not greater than L_H, it holds that, with $h = \frac{r}{N^{1+\beta}}$,*

$$\left|V(t_0,\varphi) - V_{N,\lceil N^{\beta}\rceil}(t_0,\varphi)\right| \;\leq\; \bar{C}\left(r^{\frac{\alpha\cdot\beta}{1+\beta}}\, h^{\frac{\alpha}{1+\beta}} \,\vee\, r^{\frac{\beta}{2(1+\beta)}}\sqrt{\ln\!\left(\tfrac{1}{h}\right)}\, h^{\frac{1}{2(1+\beta)}} \;+\; r^{-\frac{\beta}{1+\beta}}\, h^{\frac{\beta-3}{4(1+\beta)}}\right).$$

In particular, with $\beta = 5$ and $h = \frac{r}{N^6}$, it holds that

$$\left|V(t_0,\varphi) - V_{N,N^5}(t_0,\varphi)\right| \;\leq\; \bar{C}\left(r^{\frac{5\alpha}{6}} h^{\frac{2\alpha-1}{12}} \,\vee\, r^{\frac{5}{12}}\sqrt{\ln\left(\tfrac{1}{h}\right)} \;+\; r^{-\frac{5}{6}}\right) h^{\frac{1}{12}}.$$

Proof. Clearly, $|V - V_{N,\lceil N^\beta\rceil}| \leq |V - V_N| + |V_N - V_{N,\lceil N^\beta\rceil}|$. The assertion now follows from Corollary 3.3 and Theorem 3.1, where $\ln(\frac{1}{h_N}) = \ln(\frac{N}{r})$ is bounded by $\ln(\frac{N^{1+\beta}}{r}) = \ln(\frac{1}{h})$. □

The choice $\beta = 5$ in Theorem 3.4 yields the same rate for both summands in the error estimate provided the initial segment is at least $\frac{1}{2}$-Hölder continuous, because $\frac{1}{2} = \frac{\beta-3}{4}$ implies $\beta = 5$. Thus, the best overall error bound we obtain without additional assumptions is of order $h^{1/12}$ up to neglecting the logarithmic term.

The rate $\frac{1}{12}$ is a worst-case estimate. Moreover, better error bounds are obtained in the special situations treated at the end of Section 3.3. In the deterministic case, that is, when the diffusion coefficient σ is zero, two different bounds on the total error – depending on whether or not the space of control actions Γ is finite – can be derived by combining Corollary 3.1 from Section 3.2 with Corollary 3.4 from Section 3.3. The optimal choice of the parameter β is three for a complete and separable metric space Γ, since $1 = \frac{\beta-1}{2}$ implies $\beta = 3$, provided the initial segment as well as the coefficients b and f are Lipschitz continuous in the time variable. If Γ is finite, we choose $\beta = 1$. When the diffusion coefficient is not directly controlled and Γ is finite, we combine the assertions of Theorem 3.1 and Corollary 3.5 to obtain a bound on the overall discretisation error. The optimal choice of β is two, since $\frac{1}{2} = \frac{\beta-1}{2}$ implies $\beta = 2$.

Table 3.1 shows the corresponding bounds on the total error, that is, bounds on the maximal difference between the value functions V and $V_{N,M}$ over all initial segments of a given time regularity. The time regularity of the initial segments and of the coefficients b, σ, f in their time variable is indicated in the first column of the table. A function ψ is *Hölder* $\frac{1}{2}-$ iff $|\psi(t) - \psi(s)| \leq L_H\sqrt{|t-s|\ln(1/|t-s|)}$ for some $L_H > 0$ and all $t, s \geq 0$ with $|t-s|$ small. The second column of the table shows whether the space of control actions Γ is assumed to be finite or not. In the third column, the form of the diffusion coefficient is indicated. The second but last column shows the order of the sub-discretisation degree M in terms of the degree N of the outer discretisation. Notice that M need only be proportional to N^β with β giving the optimal order, not necessarily equal to N^β. The error bounds in terms of the time step $h = \frac{r}{N \cdot M}$ are given in the last column of the table.

Recall that $V_{N,M} \geq V_N$ for all $N, M \in \mathbb{N}$ by construction. If, instead of the two-sided error bound of Theorem 3.4, we were merely interested in obtaining an upper bound for V, we would simply compute $V_{N,M}$ with $M = 1$. Theorem 3.1 implies that we would incur an error of order nearly $\frac{1}{2}$; that is, we would have

$$V \quad\leq\quad V_{N,1} \;+\; \text{constant} \times \sqrt{\frac{\ln(N)}{N}} \qquad \text{for all } N \in \mathbb{N},\; N \geq 2r,$$

where the initial segments are supposed to be Hölder $\frac{1}{2}-$. This direction, however, is the less informative one, since we do not expect the minimal costs for the discretised system to be lower than the minimal costs for the original system.

Up to this point, we have been concerned with convergence of value functions only. A natural question to ask is the following: Suppose we have found a strategy $\bar{u} \in \mathcal{U}_{N,M}$ which

Time regularity	Space Γ	Diffusion coefficient	$M \sim$	Error bound
Lipschitz	finite	$\sigma \equiv 0$	N	$h^{1/2}$
Lipschitz	separable	$\sigma \equiv 0$	N^3	$h^{1/4}$
Hölder $\frac{1}{2}-$	finite	$\sigma(t,\varphi)$	N^2	$h^{1/6}\sqrt{\ln(\frac{1}{h})}$
Hölder $\frac{1}{2}-$	separable	$\sigma(t,\varphi,\gamma)$	N^5	$h^{1/12}\sqrt{\ln(\frac{1}{h})}$

Table 3.1: The table shows bounds on the difference between V and $V_{N,M}$ for some special situations and the general case (last row) in terms of the time step $h = \frac{r}{N\cdot M}$.

is ε-optimal for the control problem of degree (N, M) under initial condition (t_0, φ). Will this same strategy $\bar{u}$ also be nearly optimal for the original control problem?

The hypothesis that $\bar{u}$ be ε-optimal for the problem of degree (N, M) under initial condition (t_0, φ) means that $J_N(t_0, \varphi, \bar{u}) - V_{N,M}(t_0, \varphi) \leq \varepsilon$. Recall that the cost functional for the problem of degree (N, M) is identical to the one for the problem of degree N, namely J_N, and that, by construction, $J_N \geq V_{N,M} \geq V_N$ over the set of strategies $\mathcal{U}_{N,M}$. The strategy $\bar{u}$ is nearly optimal for the original control problem if there is $\tilde{\varepsilon}$ which must be small for ε small and N, M big enough such that $J(t_0, \varphi, \bar{u}) - V(t_0, \varphi) \leq \tilde{\varepsilon}$. Recall that $\mathcal{U}_{N,M} \subset \mathcal{U}$, whence $J(t_0, \varphi, \bar{u})$ is well-defined. The next theorem states that nearly optimal strategies for the approximating problems are nearly optimal for the original problem, too.

Theorem 3.5. *Assume (A1)-(A4). Let $\alpha \in (0, 1]$, $L_H > 0$. Then there is a constant $\bar{C}_r$ depending only on α, L_H, L, K, T, the dimensions and the delay length r such that for all $\beta > 3$, all $N, M \in \mathbb{N}$ with $N \geq 2r$ and $M \geq N^\beta$, all $t_0 \in I_N$, all α-Hölder continuous $\varphi \in \mathcal{C}$ with Hölder constant not greater than L_H the following holds:*

If $\bar{u} \in \mathcal{U}_{N,M}$ is such that $J_N(t_0, \varphi, \bar{u}) - V_{N,M}(t_0, \varphi) \leq \varepsilon$, then, with $h = \frac{r}{N^{1+\beta}}$,

$$J(t_0,\varphi,\bar{u}) - V(t_0,\varphi) \;\leq\; \bar{C}_r \left(h^{\frac{\alpha}{1+\beta}} \vee \sqrt{\ln\left(\tfrac{1}{h}\right)}\, h^{\frac{1}{2(1+\beta)}} \;+\; h^{\frac{\beta-3}{4(1+\beta)}} \right) \;+\; \varepsilon.$$

Proof. Let $\bar{u} \in \mathcal{U}_{N,M}$ be such that $J_N(t_0, \varphi, \bar{u}) - V_{N,M}(t_0, \varphi) \leq \varepsilon$. Then

$$\begin{aligned}
& J(t_0,\varphi,\bar{u}) - V(t_0,\varphi) \\
\leq\;& J(t_0,\varphi,\bar{u}) - J_N(t_0,\varphi,\bar{u}) \;+\; J_N(t_0,\varphi,\bar{u}) - V_{N,M}(t_0,\varphi) \;+\; V_{N,M}(t_0,\varphi) - V(t_0,\varphi) \\
\leq\;& \sup_{u\in\mathcal{U}} \left| J(t_0,\varphi,u) - J_N(t_0,\varphi,u) \right| \;+\; \varepsilon \;+\; V_{N,M}(t_0,\varphi) - V(t_0,\varphi).
\end{aligned}$$

The assertion is now a consequence of Theorem 3.1 and Theorem 3.4. $\square$

Let us suppose we have found a strategy $\bar{u}$ for the problem of degree (N, M) with fixed initial condition $(t_0, \varphi) \in I_N \times \mathcal{C}$ which is ε-optimal or optimal and a *feedback control*. The latter means here that $\bar{u}$ can be written in the form

$$\bar{u}(t)(\omega) \;=\; \bar{u}_0\big(\lfloor t \rfloor_{N\cdot M}, \Pi_N(Z^u)(\lfloor t \rfloor_{N\cdot M})(\omega)\big) \quad \text{for all } \omega \in \Omega,\ t \geq 0,$$

where Z^u is the solution to Equation (3.8) under control $\bar{u}$ and initial condition (t_0, φ) and $\bar{u}_0$ is some measurable Γ-valued function defined on $[0, \infty) \times \mathcal{C}_N$ or, because of the discretisation, on $\{k\frac{r}{N \cdot M} \mid k \in \mathbb{N}_0\} \times \mathbb{R}^{d(N \cdot M + M + 1)}$. We would like to use $\bar{u}_0$ as a feedback control for the original system. It is not clear whether this is possible unless one assumes some regularity like Lipschitz continuity of $\bar{u}_0$ in its segment variable. The problem is that we have to replace solutions to Equation (3.8) with solutions to Equation (3.2).

Something can be said, though. Recall the definition of $\mathcal{U}_{N,M}$ at the beginning of Section 3.3. Strategies in $\mathcal{U}_{N,M}$ are not only piecewise constant, they are also adapted to the filtration generated by $W(k\frac{r}{N \cdot M})$, $k \in \mathbb{N}_0$. Thus, if $\bar{u} \in \mathcal{U}_{N,M}$ is a feedback control, then it can be re-written as

$$\bar{u}(t)(\omega) \;=\; \bar{u}_1\big(\lfloor t \rfloor_{N \cdot M}, W(\lfloor t \rfloor_{N \cdot M} - k\tfrac{r}{N \cdot M})(\omega),\, k = 0, \ldots, (N+1)M\big), \quad \omega \in \Omega,\ t \geq 0,$$

where $\bar{u}_1$ is some measurable Γ-valued function depending on the initial condition (t_0, φ) and defined on $\{k\frac{r}{N \cdot M} \mid k \in \mathbb{N}_0\} \times \mathbb{R}^{d_{N,M}}$ with $d_{N,M} := d(N \cdot M + M + 1)$. The above equality has to be read keeping in mind the convention that $W(t) = 0$ if $t < 0$. The function $\bar{u}_1$ can be used as a noise feedback control for the original problem as it directly depends on the underlying noise process, which is the same for the control problem of degree (N, M) and the original problem. By Theorem 3.5, we then know that $\bar{u}_1$ induces a nearly optimal strategy for the original control problem provided $\bar{u}$ was nearly optimal for the discretised problem.

3.5 Solving the control problems of degree (N, M)

Here, we turn to the question of how to compute the value functions of the control problems resulting from the discretisation procedure analysed above. The value function of degree (N, M) is the value function of a finite-dimensional optimal control problem in discrete time. One time step corresponds to a step of length $\frac{r}{N \cdot M}$ in continuous time. The noise component of the control problem of degree (N, M) is given by a finite sequence of independent Gaussian random variables with mean zero and variance $\frac{r}{N \cdot M}$, because the time horizon is finite and the strategies in $\mathcal{U}_{N,M}$ are not only piecewise constant, but also adapted to the filtration generated by $W(k\frac{r}{N \cdot M})$, $k \in \mathbb{N}_0$.

By construction of the approximation to the dynamics in Section 3.2, the segment space for the problem of degree (N, M) is the subspace of $\mathcal{C}_N$ consisting of all functions which are piecewise linear relative to the grid $\{k\frac{r}{N \cdot M} \mid k \in \mathbb{Z}\} \cap [-r - \frac{r}{N}, 0]$. The segment space of degree (N, M), therefore, is finite-dimensional and isomorphic to $\mathbb{R}^{d_{N,M}}$ with $d_{N,M} := d(N \cdot M + M + 1)$. The functions of interest are actually those whose nodes are multiples of $\frac{r}{N}$ units of time apart, but in each step of the evolution the segment functions (and their nodes) get shifted in time by $\frac{r}{N \cdot M}$ units.

Theoretically, the Principle of Dynamic Programming as expressed in Proposition 3.6 could be applied to compute the value function $V_{N,M}$. Practically, however, it is not possible to use any algorithm based on directly applying one-step Dynamic Programming. This difficulty arises because the state space of the controlled discrete-time Markov chains we are dealing with is $\mathbb{R}^{d_{N,M}}$ and the (semi-)discrete value function $V_{N,M}$ is defined on $I_{N \cdot M} \times \mathbb{R}^{d_{N,M}}$ or, in the fully discrete case, on a $d_{N,M}$-dimensional grid. In view of Theorem 3.4, the dimension $d_{N,M}$ is expected to be very large so that storing the values of $V_{N,M}$

for all initial conditions – as required by the Dynamic Programming method – becomes impossible.

It is well known that the worst-case complexity of solving a $\tilde{d}$-dimensional discrete-time optimal control problem via Dynamic Programming grows exponentially in the dimension $\tilde{d}$. This is related to the famous "curse of dimensionality" (e. g. Bellman and Kabala, 1965: p. 63). The complexity of a problem is here understood in the sense of information-based complexity theory, see Traub and Werschulz (1998) for an overview. For a result in this spirit confirming the presence of the curse of dimensionality see Chow and Tsitsiklis (1989). Observe, though, that the complexity of a problem depends not only on the problem formulation, but crucially also on the error criterion used for determining the accuracy of approximate solutions and on the information available to the admissible algorithms.

The situation in our case is not as desperate as it might seem provided the original control problem has low dimensions d, d_1. Recall that $V_{N,M}$ is an approximation of the value function V_N constructed in Section 3.2, which in turn approximates V, the value function of the original problem, and that the problems of degree N and of degree (N, M), $M \in \mathbb{N}$, have the same dynamics and the same cost functional. Moreover, for any time $t_0 \in I_N$, both $V_N(t_0, .)$ and $V_{N,M}(t_0, .)$ live on the space of all functions $\varphi \in \mathcal{C}$ which are piecewise linear relative to the grid $\{k\frac{r}{N} \mid k \in \mathbb{Z}\} \cap [-r, 0]$. Let us write $\hat{\mathcal{C}}(N)$ for this space. Clearly, $\hat{\mathcal{C}}(N)$ is isomorphic to $\mathbb{R}^{d_N}$ with $d_N := d(N+1)$.

An approximation $\hat{V}_{N,M}(t_0, .)$ to $V_{N,M}(t_0, .)$ for times $t_0 \in I_N$ can be computed by backward iteration starting from time T_N and proceeding in time steps of length $\frac{r}{N}$. Recall that $V_{N,M}(T_N, .) = g(.)$, whence $\hat{V}_{N,M}(T_N, .)$ is determined by g, the function giving the terminal costs. To compute $\hat{V}_{N,M}(t_0, \varphi)$ for any $\varphi \in \hat{\mathcal{C}}(N)$ when $\hat{V}_{N,M}(t_0+\frac{r}{N}, .)$ is available and $t_0 \in I_N$, an "inner" backward iteration can be performed with respect to the grid $\{t_0+ k\frac{r}{N\cdot M} \mid k = 0, \ldots, M\}$.

If $t_0 \in I_N$, then, on the time interval $[t_0, t_0+\frac{r}{N})$, the coefficients b, σ, f are functions of the control variable only, see Equations (3.8) and (3.9), respectively, and the proof of Theorem 3.3. The inner optimisation thus consists in solving a d-dimensional discrete-time optimal control problem with "constant coefficients" and fixed initial condition over M time steps, which correspond to a time horizon of length $\frac{r}{N}$. To be more precise, define, for each $n \in \mathbb{N}_0$, an operator $\mathcal{T}_n^{(N,M)}$ on the space $\mathbf{B}(\hat{\mathcal{C}}(N))$ of all bounded real-valued functions on $\hat{\mathcal{C}}(N)$ by

(3.12)
$$\mathcal{T}_n^{(N,M)}(\Psi)(\varphi) := \inf_{u \in \mathcal{U}_{N,M}} \mathbf{E}\left(\int_0^{\frac{r}{N}} f\big(n\tfrac{r}{N}, \varphi, u(s)\big)ds \;+\; \Psi\big(\mathrm{Lin}_N(Z^u_{\frac{r}{N}})\big)\right), \quad \varphi \in \hat{\mathcal{C}}(N),$$

where $Z^u = Z^{u,n,\varphi}$ is the process defined on the time interval $[-r, \frac{r}{N}]$ by

(3.13)
$$Z^u(t) := \begin{cases} \varphi(0) + \int_0^t b\big(n\frac{r}{N}, \varphi, u(s)\big)ds \;+\; \int_0^t \sigma\big(n\frac{r}{N}, \varphi, u(s)\big)dW(s), & t \in (0, \frac{r}{N}], \\ \varphi(t), & t \in [-r, 0]. \end{cases}$$

The definition of $\mathcal{T}_n^{(N,M)}$ should be compared to Proposition 3.6. Given $\Psi \in \mathbf{B}(\hat{\mathcal{C}}(N))$, let us refer to the evaluation of $\mathcal{T}_n^{(N,M)}(\Psi)$ at $\varphi \in \hat{\mathcal{C}}(N)$ as the *Bellman step* for Ψ at segment φ and time step n. Notice that $\mathrm{Lin}_N(\varphi) = \varphi$ for all $\varphi \in \hat{\mathcal{C}}(N)$. Since any strategy $u \in \mathcal{U}_{N,M}$ is piecewise constant relative to the grid $\{k\frac{r}{N\cdot M} \mid k \in \mathbb{N}_0\}$, the integrals appearing in

(3.12) and (3.13) are really finite sums of random variables; for $n \in \mathbb{N}_0$, $\Psi \in \mathbf{B}(\hat{\mathcal{C}}(N))$, all $\varphi \in \hat{\mathcal{C}}(N)$, it holds that

$$\mathcal{T}_n^{(N,M)}(\Psi)(\varphi) \;=\; \inf_{u \in \mathcal{U}_{N,M}} \mathbf{E}\left(\frac{r}{N \cdot M} \left(\sum_{k=0}^{M-1} f\big(n\tfrac{r}{N}, \varphi, u(k\tfrac{r}{N \cdot M})\big) \right) \;+\; \Psi\big(\mathrm{Lin}_N(Z^u_{\frac{r}{N}})\big) \right),$$

where $\mathrm{Lin}_N(Z^u_{\frac{r}{N}})$ is an element of $\hat{\mathcal{C}}(N)$ and is completely determined by $\varphi(-r+k\frac{r}{N})$, $k \in \{0,\dots,N-1\}$, and

$$\begin{aligned} Z^u(\tfrac{r}{N}) \;=\; \varphi(0) \;+\; & \frac{r}{N \cdot M} \left(\sum_{k=0}^{M-1} b\big(n\tfrac{r}{N}, \varphi, u(k\tfrac{r}{N \cdot M})\big) \right) \\ &+ \sum_{k=0}^{M-1} \sigma\big(n\tfrac{r}{N}, \varphi, u(k\tfrac{r}{N \cdot M})\big) \big(W\big((k+1)\tfrac{r}{N \cdot M}\big) - W\big(k\tfrac{r}{N \cdot M}\big)\big). \end{aligned}$$

If the diffusion coefficient σ is not directly controlled, that is, if $\sigma(t,\varphi,\gamma) = \tilde{\sigma}(t,\varphi)$, then the expression for $Z^u(\frac{r}{N})$ simplifies to

$$Z^u(\tfrac{r}{N}) \;=\; \varphi(0) \;+\; \frac{r}{N \cdot M} \left(\sum_{k=0}^{M-1} b\big(n\tfrac{r}{N}, \varphi, u(k\tfrac{r}{N \cdot M})\big) \right) \;+\; \tilde{\sigma}\big(n\tfrac{r}{N}, \varphi\big) W\big(\tfrac{r}{N}\big).$$

Observe that the operator $\mathcal{T}_n^{(N,M)}$ is a non-expansive mapping in supremum norm on $\mathbf{B}(\hat{\mathcal{C}}(N))$, that is,

$$\sup_{\varphi \in \hat{\mathcal{C}}(N)} |\mathcal{T}_n^{(N,M)}(\Psi)(\varphi) - \mathcal{T}_n^{(N,M)}(\tilde{\Psi})(\varphi)| \;\leq\; \sup_{\varphi \in \hat{\mathcal{C}}(N)} |\Psi(\varphi) - \tilde{\Psi}(\varphi)| \quad \text{for all } \Psi, \tilde{\Psi} \in \mathbf{B}(\hat{\mathcal{C}}(N)).$$

This property, though evident from (3.12), is important in that it guarantees numerical stability when the operators $\mathcal{T}_n^{(N,M)}$, $n \in \mathbb{N}_0$, are repeatedly applied.

The Bellman steps need not necessarily be backward iterations of Dynamic Programming type as was suggested above. We can use any method that solves the arising M-step "constant coefficients" control problems. When the space of control actions Γ is finite, then the coefficients b, σ, f can be evaluated in advance at $(n\frac{r}{N}, \varphi, \gamma)$ for all $\gamma \in \Gamma$, because the time segment pair $(n\frac{r}{N}, \varphi)$ is constant during any Bellman step.

In the deterministic case, it is sometimes possible to optimise directly over the set of deterministic M-step strategies. If Γ has finite cardinality N_Γ, instead of checking N_Γ to the power of M possibilities, we only have to test $\binom{N_\Gamma+M-1}{M}$ possibilities, which is the number of combinations of M objects when there are N_Γ different kinds of objects.

In the stochastic case, a method recently introduced by Rogers (2007) for computing value functions of high-dimensional discrete-time Markovian optimal control problems might prove useful. The method is based on path-wise optimisation and Monte Carlo simulation of trajectories of a reference Markov chain; it uses minimisation over functions which can be interpreted as candidates for the value function. Those candidates should be chosen from a computationally "nice" class so that the value function can be computed at any given point without the need to store its values for the entire state space, although this problem is less acute for low dimensions d, d_1. Unlike schemes directly employing the PDP, Rogers's method does not yield an approximation of the value function over the entire state space, but only its value at the given initial point. This is what is needed for the Bellman step.

Let us return to our procedure for computing $\hat{V}_{N,M}(t_0,.)$, $t_0 \in I_N$. Set $n_T := \lfloor T\frac{N}{r} \rfloor$. The procedure starts by determining $\hat{V}_{N,M}(n_T\frac{r}{N},.) = \hat{V}_{N,M}(T_N,.)$ from g. To this end, choose a finite subset $S_{n_T} \subset \hat{\mathcal{C}}(N)$. For each $\varphi \in S_{n_T}$, set $\hat{V}_{N,M}(n_T\frac{r}{N},\varphi) := g(\varphi)$. The values of $\hat{V}_{N,M}(n_T\frac{r}{N},.)$ at segments not in S_{n_T} are calculated by some interpolation or regression method. Now, suppose that $\hat{V}_{N,M}((n{+}1)\frac{r}{N},.)$ is available for some $n \in \{0,\dots,n_T{-}1\}$. Then the following steps are executed:

1. Choose a finite set $S_n \subset \hat{\mathcal{C}}(N)$.

2. For each segment $\varphi \in S_n$, compute $\hat{V}_{N,M}(n\frac{r}{N},\varphi)$ by executing the Bellman step for $\hat{V}_{N,M}((n{+}1)\frac{r}{N},.)$ at φ and time step n.

3. Compute $\hat{V}_{N,M}(n\frac{r}{N},.)$ by some interpolation or regression method using the data $\{(\varphi, \hat{V}_{N,M}(n\frac{r}{N},\varphi)) \mid \varphi \in S_n\}$.

In this way, by backward iteration, $\hat{V}_{N,M}(n\frac{r}{N},.)$ can be calculated for all $n \in \{0,\dots,n_T\}$. The proposed procedure may be called an application of *approximate Dynamic Programming* or *approximate value iteration*[1] (e. g. Bertsekas, 2005, 2007: I.6, II.1.3). The idea is probably as old as Dynamic Programming itself, cf. Bellman and Kabala (1965).

```
Input: SYSTEM, T, r, N, M
Output: V[0],...,V[T*N/r]

SYSTEM.set_parameters(r,N,M);
SEGMENTS.set_parameters(r,N);
n <- T*N/r;
for i = 0 to n do V[i].set_parameters(r,N);
SEGMENTS.generate(n);
for each x in SEGMENTS do V[n].add(x,SYSTEM.g(x));
V[n].interpolate;
while (n > 0) do begin
  n <- n-1;
  SEGMENTS.generate(n);
  for each x in SEGMENTS do V[n].add(x,SYSTEM.Bellman_step(n,x,V[n+1]));
  V[n].interpolate;
end_while;
```

Figure 3.1: Approximate value iteration: scheme in pseudo code. The object `SYSTEM` contains the coefficient functions b, σ, f, g and provides a method for the Bellman step. The objects `V[0],...,V[T*N/r]` represent approximations to $V_{N,M}(n\frac{r}{N},.)$, $n = 0,\dots,\lfloor T\frac{N}{r} \rfloor$; they possess an interpolation method, as values are calculated only at segments provided by `SEGMENTS`.

[1]The term "value iteration" is usually reserved for the backward iteration in value function space when solving infinite horizon control problems, "Dynamic Programming" for the finite backward iteration when solving problems with finite time horizon.

Figure 3.1 represents the procedure in an object-oriented pseudo code. The object `SYSTEM` contains the coefficient functions b, σ, f, g; the terminal costs g are directly accessible, the other functions are needed for the method `Bellman_step`, which implements the operators $\mathcal{T}_n^{(N,M)}$, $n \in \mathbb{N}$. The object `SEGMENTS` generates and stores the sets S_n of segments at which the Bellman step is carried out. The objects `V[0],...,V[T*N/r]` represent the approximations $\hat{V}_{N,M}(n\frac{r}{N},.)$ to the value functions $V_{N,M}(n\frac{r}{N},.)$, $n = 0,\ldots,n_T$. The method `interpolate` creates an interpolant using the data stored in `V[n]`, that is, it implements the creation of $\hat{V}_{N,M}(n\frac{r}{N},.)$ from the data $\{(\varphi, \hat{V}_{N,M}(n\frac{r}{N},\varphi)) \mid \varphi \in S_n\}$.

We have seen how the Bellman steps can be computed in principle, but will leave open the question of which algorithm should be used. There are two other important questions, here. The first is the choice of the sets of segments $S_n \subset \hat{\mathcal{C}}(N)$, $n \in \{0,\ldots,n_T\}$. The second regards the choice of the interpolation or regression method. Clearly, the two questions are interrelated in that the choice of a certain interpolation method may require a specific choice of the segment sets.

Suppose we have chosen, for each time step n, a set of segments S_n as well as an interpolation method. The latter can be represented as a mapping $\mathcal{A}_n^N : \mathbf{B}(\hat{\mathcal{C}}(N)) \to \mathbf{B}(\hat{\mathcal{C}}(N))$ such that $\mathcal{A}_n^N(\Psi) = \mathcal{A}_n^N(\tilde{\Psi})$ whenever $\Psi(\varphi) = \tilde{\Psi}(\varphi)$ for all $\varphi \in S_n$. The approximate value iteration procedure can then be written as

$$
\begin{aligned}
\hat{V}_{N,M}(n_T\tfrac{r}{N},.) &:= \mathcal{A}_{n_T}^N(g),\\
\hat{V}_{N,M}\big(n\tfrac{r}{N},.\big) &:= \mathcal{A}_n^N \circ \mathcal{T}_n^{N,M}\big(\hat{V}_{N,M}\big((n{+}1)\tfrac{r}{N},.\big)\big), \qquad n \in \{0,\ldots,n_T{-}1\}.
\end{aligned}
$$

An important restriction on the choice of the interpolation method is that the corresponding operators $\mathcal{A}_n^N$ should be non-expansive mappings. This is to preserve the non-expansiveness of the Bellman operator, which in turn guarantees numerical stability of the recursion. Admissible methods are, for example, the nearest neighbour and k nearest neighbour regression, which work with any choice of the segment sets, or interpolation methods using piecewise linear basis functions.

Recall that $\hat{\mathcal{C}}(N)$ is isomorphic to $\mathbb{R}^{d_N}$ with $d_N = d(N{+}1)$. On the other hand, the value function of degree (N,M) is Lipschitz continuous, but not necessarily continuously differentiable. The problem of recovering a Lipschitz continuous function defined on a $\tilde{d}$-dimensional hypercube (to work on a bounded domain) is itself subject to a dimensional curse, at least when the error is measured in supremum norm. Consequently, approximate Dynamic Programming in itself provides no escape from the curse of dimensionality.

Instead of treating the values at the grid points of the segment functions in $\hat{\mathcal{C}}(N)$ as belonging to independent dimensions, we may exploit the fact that they are generated by continuous functions. In view of the error bounds of Sections 3.2 and 3.4, which are uniform only over sets of Lipschitz or Hölder continuous segments with bounded Lipschitz or Hölder constant, it is natural to restrict the domain of the value function of degree (N,M) accordingly. Any function in $\hat{\mathcal{C}}(N)$ is, by construction, Lipschitz continuous, yet its Lipschitz constant may be arbitrarily large. For $\tilde{L} > 0$, let $\hat{\mathcal{C}}_{Lip}(N,\tilde{L})$ denote the (convex) set of all functions in $\hat{\mathcal{C}}(N)$ with Lipschitz constant not greater than $\tilde{L}$. Denote by $\hat{\mathcal{C}}_{1/2-}(N,\tilde{L})$ the (convex) set of all functions in $\varphi \in \hat{\mathcal{C}}(N)$ such that

$$
|\varphi(t) - \varphi(s)| \;\leq\; \tilde{L}\sqrt{|t{-}s|\ln\!\big(\tfrac{e\cdot r}{|t-s|}\big)} \quad \text{for all } t,s \in [-r,0].
$$

In the case of a deterministic system and for bounded drift coefficient b, the segments of all solution trajectories of the original dynamics are Lipschitz continuous with Lipschitz constant not greater than $\tilde{L}$ provided the initial segments are that regular and $\tilde{L}$ was chosen big enough. In the stochastic case, boundedness of b and σ does not guarantee that all trajectory segments are Hölder $\frac{1}{2}-$ for some constant $\tilde{L}$; nevertheless, for all Hölder $\frac{1}{2}-$ initial segments, all trajectory segments are Hölder $\frac{1}{2}-$, and the probability that a trajectory segment has Hölder constant greater than $\tilde{L}$ tends to zero as $\tilde{L}$ goes to infinity, again provided the initial segments are Hölder $\frac{1}{2}-$ with constant $\tilde{L}$. Moreover, the probability that a trajectory segment has Hölder constant greater than $\tilde{L}$ can be estimated by deriving bounds on the moments of the modulus of continuity of Itô diffusions as in Appendix A.2.

These observations can be used in choosing the sets S_n, $n \in \{0, \ldots, n_T\}$ of grid segments. In generating appropriate Lipschitz or Hölder continuous segments, the Brownian bridge construction or a deterministic analogue may be used. The underlying idea is that not all dimensions of the piecewise linear segments are equally important. In particular, the right-most coordinate, which corresponds to the current time, plays a special role in that it provides the initial value for generating the new current state, cf. (3.13).

We leave these observations to future investigation. First numerical experiments have been carried out for the simple deterministic system presented in Subsection 1.2.2. A rough approximation to the true value function can be obtained. The choice of the grid segments and of the interpolation method are seen to be crucial in view of the heavy requirements in memory and computing time.

3.6 Conclusions and open questions

In this chapter, we have presented and analysed a semi-discretisation scheme for finite horizon stochastic control problems with delay. The dependence of the system on its past evolution is allowed to be of a general form; it includes point and distributed delay as well as generalised distributed and functional delay, cf. Section 3.1. Apart from the somewhat restrictive assumption of boundedness, the hypotheses on the system coefficients are quite natural. The state process and the noise process may have different dimensions (d and d_1), and no non-degeneracy assumption on the diffusion coefficient is needed. The space of control actions Γ may be an arbitrary complete separable metric space (only separability is really needed); in particular, Γ need not be compact. The discretisation of time induces a discretisation of the segment space. The discrete-time optimal control problems generated by the scheme are, as a result, finite-dimensional.

Convergence of the scheme has been demonstrated and bounds on the discretisation error have been derived. Under general assumptions, we have a worst-case estimate of the rate of convergence; better bounds have been obtained for important special situations, namely for the deterministic case (finite and separable Γ) and the case of uncontrolled diffusion coefficient (and finite Γ). We stress that the error bounds of Section 3.4 hold without any assumptions on regularity or existence of optimal strategies and without any additional assumptions on the regularity of the value function. Indeed, there are control problems satisfying our hypotheses which either do not possess optimal strategies or where the optimal strategies are Borel measurable, but almost surely discontinuous on any non-empty open time interval, or where the value function is Lipschitz continuous, but not everywhere (Fréchet) differentiable.

The structure of our two-step discretisation scheme can be exploited in designing algorithms for the numerical solution of the discrete-time control problems of degree (N, M). In this way, the memory requirements can be kept within feasible limits. The Bellman steps, that is, the inner optimisation steps of the procedure proposed in Section 3.5, are of "constant coefficients" type, which may be computationally advantageous.

In contrast to Chapter 2, the analysis of this chapter is not confined to proving mere convergence of a discretisation scheme. Kushner's Markov chain method, on the other hand, is applicable to a wide variety of dynamic optimisation problems and discretisation schemes. Notice, however, that some kind of compactness assumption regarding the space of strategies is an essential ingredient of the method, cf. Section 2.2.

In connection with the two-step scheme, there are some open questions. The error bound obtained under general assumptions is a worst-case estimate of the rate of convergence, but it is not clear whether it is sharp. Due to the structure of the scheme, none of the error bounds can be improved beyond the rate of convergence attained by the Euler scheme for the corresponding uncontrolled system – unless the cost functional has some special form.[2]

As far as the numerical solution of the discrete-time control problems of degree (N, M) is concerned, a lot is still to be done. On the one hand, there is the question of the complexity of the problem (in the sense of information based complexity), which depends on the error criterion adopted. On the other hand, there is the question of how to implement the scheme of Section 3.5. Observe that, even if the problem is subject to a dimensional curse (in the discretisation degree N), an approximate Dynamic Programming algorithm can still be useful, as it will produce a first rough approximation to the value function of the original problem. Such an approximation, in turn, can serve as an initial guess of the value function for algorithms of suboptimal control like "limited lookahead" or "rollout" (cf. Bertsekas, 2005: Ch. 6).

The two-step discretisation scheme should be applicable to other types of optimal control problems with delay. Instead of a finite deterministic time, the (random) time of first exit from a compact set (as in Section 2.3) may be taken as time horizon. Other interesting systems are those with reflection at the boundary of a compact polyhedron. The state process would, in both cases, take values in a bounded subset of $\mathbb{R}^d$, which is reasonable also from the point of view of numerical computation. What has to be established is, again, not so much whether the scheme converges, but how fast.

[2] For stochastic systems with general cost functionals, we have the *strong* rate of convergence of the corresponding Euler scheme as bound on the rate of convergence of the two-step scheme. For special cost functionals, the scheme might attain the *weak* rate of convergence of the Euler scheme.

Appendix A

A.1 On the Principle of Dynamic Programming

Let $((\Omega, \mathcal{F}, \mathbf{P}), (\mathcal{F}_t), W)$ be a Wiener basis of dimension d_1. Let $\mathcal{U}$ be the associated set of control processes. For $n \in \mathbb{N}$, define the set $\mathcal{U}_n \subset \mathcal{U}$ of piecewise constant strategies according to (3.10) at the beginning of Section 3.3. Let $\tilde{\mathcal{U}}$ be either $\mathcal{U}$ or $\mathcal{U}_n$ for some $n \in \mathbb{N}$.

Let $\tilde{r} > 0$ and set $\tilde{\mathcal{C}} := \mathbf{C}([-\tilde{r}, 0], \mathbb{R}^d)$. If Y is an $\mathbb{R}^d$-valued process, then the notation Y_t in this subsection denotes the segment of length $\tilde{r}$. Let $\tilde{b}$, $\tilde{\sigma}$, $\tilde{f}$, $\tilde{g}$ be functions satisfying the following hypotheses:

(H1) Measurability: $\tilde{b}\colon [0,\infty) \times \tilde{\mathcal{C}} \times \Gamma \to \mathbb{R}^d$, $\tilde{\sigma}\colon [0,\infty) \times \tilde{\mathcal{C}} \to \mathbb{R}^{d \times d_1}$, $\tilde{f}\colon [0,\infty) \times \tilde{\mathcal{C}} \times \Gamma \to \mathbb{R}$, $\tilde{g}\colon \tilde{\mathcal{C}} \to \mathbb{R}$ are Borel measurable functions.

(H2) Boundedness: $|\tilde{b}|$, $|\tilde{\sigma}|$, $|\tilde{f}|$, $|\tilde{g}|$ are bounded by some positive constant K.

(H3) Uniform Lipschitz condition: there is a constant $L > 0$ such that for all $\varphi, \psi \in \tilde{\mathcal{C}}$, all $t \geq 0$, all $\gamma \in \Gamma$

$$\begin{aligned} |\tilde{b}(t,\varphi,\gamma) - \tilde{b}(t,\psi,\gamma)| \vee |\tilde{\sigma}(t,\varphi,\gamma) - \tilde{\sigma}(t,\psi,\gamma)| &\leq L\,\|\varphi - \psi\|, \\ |\tilde{f}(t,\varphi,\gamma) - \tilde{f}(t,\psi,\gamma)| \vee |\tilde{g}(\varphi) - \tilde{g}(\psi)| &\leq L\,\|\varphi - \psi\|. \end{aligned}$$

Let $\tilde{T} > 0$. Define a cost functional $\tilde{J}\colon [0,\tilde{T}] \times \tilde{\mathcal{C}} \times \mathcal{U} \to \mathbb{R}$ by

$$\tilde{J}(t_0,\psi,u) := \mathbf{E}\left(\int_0^{\tilde{T}-t_0} \tilde{f}\big(t_0{+}s, Y_s, u(s)\big) ds \;+\; \tilde{g}\big(Y_{\tilde{T}-t_0}\big) \right),$$

where $Y = Y^{t_0,\psi,u}$ is the solution to the controlled SDDE

$$\text{(A.1)} \quad Y(t) = \begin{cases} \psi(0) + \int_0^t \tilde{b}\big(t_0{+}s, Y_s, u(s)\big) ds \;+\; \int_0^t \tilde{\sigma}\big(t_0{+}s, Y_s, u(s)\big) dW(s), & t > 0, \\ \psi(t), \quad t \in [-\tilde{r}, 0]. \end{cases}$$

Define the associated value function $\tilde{V}\colon [0,\tilde{T}] \times \tilde{\mathcal{C}} \to \mathbb{R}$ by

$$\tilde{V}(t_0,\psi) := \inf\big\{ \tilde{J}(t_0,\psi,u) \mid u \in \tilde{\mathcal{U}} \big\}.$$

Depending on the choice of $\tilde{\mathcal{U}}$, the function $\tilde{V}$ thus defined gives the minimal costs over the set $\mathcal{U}$ of all control processes or just over a set of strategies which are piecewise constant relative to the grid $\{k\frac{r}{n} \mid k \in \mathbb{N}_0\}$ for some $n \in \mathbb{N}$. The following property of $\tilde{V}$ is useful.

Proposition A.1. *Assume (H1)-(H3). Let $\tilde V$ be the value function defined above. Then $\tilde V$ is bounded and Lipschitz continuous in the segment variable uniformly in the time variable. More precisely, $|\tilde V|$ is bounded by $K(\tilde T+1)$ and for all $t_0 \in [0,\tilde T]$, all $\varphi,\psi \in \tilde{\mathcal{C}}$,*

$$|\tilde V(t_0,\varphi) - \tilde V(t_0,\psi)| \;\leq\; 2\sqrt{2}\,L(\tilde T+1)\exp\bigl(3\tilde T(\tilde T+4d_1)L^2\bigr)\,\|\varphi-\psi\|.$$

Proof. Boundedness of $\tilde V$ is an immediate consequence of its definition and Hypothesis (H2). Let $t_0 \in [0,\tilde T]$, let $\varphi,\psi \in \tilde{\mathcal{C}}$. Recall the inclusion $\tilde{\mathcal{U}} \subseteq \mathcal{U}$ and observe that, in virtue of the definition of $\tilde V$, we have

$$|\tilde V(t_0,\varphi) - \tilde V(t_0,\psi)| \quad\leq\quad \sup_{u\in\mathcal{U}} |\tilde J(t_0,\varphi,u) - \tilde J(t_0,\psi,u)|.$$

By Hypothesis (H3), for all $u \in \mathcal{U}$ we get

$$\begin{aligned}
&|\tilde J(t_0,\varphi,u) - \tilde J(t_0,\psi,u)|\\
\leq\;& \mathbf{E}\left(\int_0^{\tilde T-t_0} \bigl|\tilde f\bigl(t_0+s,X^u_s,u(s)\bigr) - \tilde f\bigl(t_0+s,Y^u_s,u(s)\bigr)\bigr|ds \;+\; \bigl|\tilde g\bigl(X^u_{\tilde T-t_0}\bigr) - \tilde g\bigl(Y^u_{\tilde T-t_0}\bigr)\bigr|\right)\\
\leq\;& L(1+\tilde T-t_0)\,\mathbf{E}\left(\sup_{t\in[-\tilde r,\tilde T]} |X^u(t)-Y^u(t)|^2\right)^{\frac{1}{2}},
\end{aligned}$$

where X^u, Y^u are the solutions to Equation (A.1) under control process u with initial conditions (t_0,φ) and (t_0,ψ), respectively. Now, for every $T \in [0,\tilde T]$,

$$\mathbf{E}\left(\sup_{t\in[-\tilde r,T]} |X^u(t)-Y^u(t)|^2\right) \quad\leq\quad 2\,\mathbf{E}\left(\sup_{t\in[0,T]} |X^u(t)-Y^u(t)|^2\right) \;+\; 2\|\varphi-\psi\|^2,$$

while Hölder's inequality, Doob's maximal inequality, Itô's isometry, Fubini's theorem and Hypothesis (H3) together yield

$$\begin{aligned}
&\mathbf{E}\left(\sup_{t\in[0,T]} |X^u(t)-Y^u(t)|^2\right)\\
\leq\;& 3|\varphi(0)-\psi(0)|^2 \;+\; 3T\,\mathbf{E}\left(\int_0^T \Bigl|\tilde b\bigl(t_0+s,X^u_s,u(s)\bigr) - \tilde b\bigl(t_0+s,Y^u_s,u(s)\bigr)\Bigr|^2 ds\right)\\
&+\; 3d_1 \sum_{i=1}^{d}\sum_{j=1}^{d_1} \mathbf{E}\left(\sup_{t\in[0,T]} \left(\int_0^t \bigl(\tilde\sigma_{ij}\bigl(t_0+s,X^u_s,u(s)\bigr) - \tilde\sigma_{ij}\bigl(t_0+s,Y^u_s,u(s)\bigr)\bigr)\,dW^j(s)\right)^2\right)\\
\leq\;& 3|\varphi(0)-\psi(0)|^2 \;+\; 3T\,L^2 \int_0^T \mathbf{E}\left(|X^u_s - Y^u_s|^2\right) ds\\
&+\; 12d_1\,\mathbf{E}\left(\int_0^T \sum_{i=1}^{d}\sum_{j=1}^{d_1} \bigl(\tilde\sigma_{ij}\bigl(t_0+s,X^u_s,u(s)\bigr) - \tilde\sigma_{ij}\bigl(t_0+s,Y^u_s,u(s)\bigr)\bigr)^2 ds\right)\\
\leq\;& 3|\varphi(0)-\psi(0)|^2 \;+\; 3(T+4d_1)L^2 \int_0^T \mathbf{E}\Bigl(\sup_{t\in[-\tilde r,s]} |X^u(t)-Y^u(t)|^2\Bigr)ds.
\end{aligned}$$

Since $|\varphi(0)-\psi(0)| \leq \|\varphi-\psi\|$, Gronwall's lemma implies that

$$\mathbf{E}\left(\sup_{t\in[-\tilde r,\tilde T]} |X^u(t)-Y^u(t)|^2\right) \quad\leq\quad 8\|\varphi-\psi\|^2\exp\bigl(6\tilde T(\tilde T+4d_1)L^2\bigr).$$

Putting the estimates together, we obtain the assertion. □

Recall that the the value function $\tilde{V}$ has been defined over the set of strategies $\tilde{\mathcal{U}}$. If $\tilde{\mathcal{U}} = \mathcal{U}$, set $\tilde{I} := [0, \infty)$, else if $\tilde{\mathcal{U}} = \mathcal{U}_n$, set $\tilde{I} := \{k\frac{r}{n} \,|\, k \in \mathbb{N}_0\}$. The following version of Bellman's Principle of Optimality or Principle of Dynamic Programming holds.

Theorem A.1 (PDP). *Assume (H1)-(H3). Then for all $t_0 \in [0, \tilde{T}]$, all $t \in \tilde{I} \cap [0, \tilde{T} - t_0]$, all $\psi \in \tilde{\mathcal{C}}$,*

$$\tilde{V}(t_0, \psi) \;=\; \inf_{u \in \mathcal{U}} \mathbf{E}\left(\int_0^t \tilde{f}\big(t_0 + s, Y^u_s, u(s)\big) ds \;+\; \tilde{V}(t_0 + t, Y^u_t) \right),$$

where Y^u is the solution to Equation (A.1) *under control process u with initial condition (t_0, ψ).*

Theorem A.1 is proved in the same way as Theorem 4.2 in Larssen (2002), also see the proof of Theorem 4.3.3 in Yong and Zhou (1999: p. 180). We merely point out the differences in the problem formulation and the hypotheses. Here, all coefficients, those of the dynamics and those of the cost functional, are bounded, while Larssen (2002) also allows for sub-linear growth. Since Equation (A.1) has unique solutions, boundedness of the coefficients guarantees that the cost functional $\tilde{J}$ as well as the value function $\tilde{V}$ are well defined. Notice that we express dependence on the initial time in a different, but equivalent way in comparison with Larssen (2002). Notice further that in Theorem A.1 only deterministic times appear.

We have stated the control problem and given Bellman's principle in the strong Wiener formulation, cf. Section 3.1. Although the weak Wiener formulation is essential for the proof, the resulting value functions are the same for both versions. This is due to the fact that weak uniqueness holds for Equation (A.1). Also the infimum in the Dynamic Programming equation can be taken over all Wiener control bases or just over all control processes associated with a fixed Wiener basis.

There are two respects in which our hypotheses are more general than those of Theorem 4.2 in Larssen (2002). The first is that we do not require the integrand $\tilde{f}$ of the cost functional to be uniformly continuous in its three variables. This assumption is not needed for the Dynamic Programming equation, while it is important for versions of the Hamilton-Jacobi-Bellman partial differential equation. The second is that we allow the optimisation problem to be formulated for certain subclasses of admissible strategies, namely the subclasses $\mathcal{U}_n$ of piecewise constant strategies. The set $\tilde{I}$ and thus the set of allowed intermediate times must be chosen accordingly.

A.2 On the modulus of continuity of Itô diffusions

A typical trajectory of standard Brownian motion is Hölder continuous of any order less than one half. If such a trajectory is evaluated at two different time points $t_1, t_2 \in [0, T]$ with $|t_1 - t_2| \leq h$ small, then the difference between the values at t_1 and t_2 is not greater than a multiple of $\sqrt{h \ln(\frac{1}{h})}$, where the proportionality factor depends on the trajectory and the time horizon T, but not on the choice of the time points t_1, t_2. This is a consequence of Lévy's exact modulus of continuity for Brownian motion. The modulus of continuity of a stochastic process is a random element. Lemma A.1 below shows that the modulus of continuity of Brownian motion and, more generally, that of any Itô diffusion with bounded coefficients has finite moments of any order.

Lemma A.1, which treats the case of Itô diffusions with bounded coefficients, can be found in Słomiński (2001), cf. Lemma A.4 there. It is enough to prove Lemma A.1 for the special case of one-dimensional Brownian motion. The full statement is then derived by a component-wise estimate and a time-change argument (the Dambis-Dubins-Schwarz theorem), cf. Theorem 3.4.6 in Karatzas and Shreve (1991: p. 174), for example.

One way of proving the assertion for Brownian motion – different from the proof in Słomiński (2001) – is to follow the derivation of Lévy's exact modulus of continuity as suggested in Exercise 2.4.8 of Stroock and Varadhan (1979). The main ingredient there is an inequality due to Garsia, Rodemich, and Rumsey, see Theorem 2.1.3 in Stroock and Varadhan (1979: p. 47) and Garsia et al. (1970). For the sake of completeness, we give the two proofs in full detail.

Lemma A.1 (Słomiński). *Let W be a d_1-dimensional Wiener process living on the probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Let $Y = (Y^{(1)}, \dots, Y^{(d)})^{\mathsf{T}}$ be an Itô diffusion of the form*

$$Y(t) \;=\; y_0 + \int_0^t \tilde{b}(s)ds \;+\; \int_0^t \tilde{\sigma}(s)dW(s), \quad t \geq 0,$$

where $y_0 \in \mathbb{R}^d$ and $\tilde{b}$, $\tilde{\sigma}$ are $(\mathcal{F}_t)$-adapted processes with values in $\mathbb{R}^d$ and $\mathbb{R}^{d\times d_1}$, respectively. If $|\tilde{b}|$, $|\tilde{\sigma}|$ are bounded by some positive constant K, then it holds that for every $p > 0$, every $T > 0$ there is a constant $C_{p,T}$ depending only on K, the dimensions, p and T such that

$$\mathbf{E}\left(\sup_{t,s\in[0,T],|t-s|\leq h} \left|Y(t) - Y(s)\right|^p\right) \;\leq\; C_{p,T}\left(h\ln(\tfrac{1}{h})\right)^{\frac{p}{2}} \quad \textit{for all } h \in (0, \tfrac{1}{2}].$$

Proof. Let $T > 0$, $p > 0$. Then for all $t, s \in [0, T]$,

$$\left|Y(t) - Y(s)\right|^p \quad\leq\quad d^{\frac{p}{2}}\left(\left|Y^{(1)}(t) - Y^{(1)}(s)\right|^p + \ldots + \left|Y^{(d)}(t) - Y^{(d)}(s)\right|^p\right),$$

and for the i-th component we have

$$\begin{aligned}\left|Y^{(i)}(t) - Y^{(i)}(s)\right|^p \quad &= \quad \left|\int_s^t \tilde{b}_i(\tilde{s})d\tilde{s} \;+\; \sum_{j=1}^{d_1}\int_s^t \tilde{\sigma}_{ij}(\tilde{s})dW^j(\tilde{s})\right|^p \\ \leq \;\; (d_1{+}1)^p &\left(K^p|t{-}s|^p \;+\; \sum_{j=1}^{d_1}\left|\int_s^t \tilde{\sigma}_{ij}(\tilde{s})dW^j(\tilde{s})\right|^p\right).\end{aligned}$$

Hence, for $h \in (0, \frac{1}{2}]$,

$$\begin{aligned}&\mathbf{E}\left(\sup_{t,s\in[0,T],|t-s|\leq h}\left|Y(t) - Y(s)\right|^p\right) \\ \leq\;\; & d^{\frac{p}{2}}(d_1{+}1)^p\left(d\,K^p\,h^p \;+\; \sum_{i=1}^{d}\sum_{j=1}^{d_1}\mathbf{E}\left(\sup_{t,s\in[0,T],|t-s|\leq h}\left|\int_s^t \tilde{\sigma}_{ij}(\tilde{s})dW^j(\tilde{s})\right|^p\right)\right).\end{aligned}$$

To prove the assertion, it is enough to show that the $d \cdot d_1$ expectations on the right-hand side of the last inequality are of the right order. Let $i \in \{1, \dots, d\}$, $j \in \{1, \dots, d_1\}$, and define the one-dimensional process $M = M^{(i,j)}$ by

$$M(t) := \begin{cases} \int_0^t \tilde{\sigma}_{ij}(\tilde{s})\, dW^j(\tilde{s}) & \text{if } t \in [0,T], \\ M(T) + W^j(t) - W^j(T) & \text{if } t > T. \end{cases}$$

Since $\tilde{\sigma}_{ij}$ is bounded, the process M is a martingale and can be represented as a time-changed Brownian motion. More precisely, by the Dambis-Dubins-Schwarz theorem, see Theorem 3.4.6 in Karatzas and Shreve (1991: p. 174), for example, there is a standard one-dimensional Brownian motion $\tilde{W}$ living on $(\Omega, \mathcal{F}, \mathbf{P})$ such that, $\mathbf{P}$-almost surely,

$$M(t) \;=\; \tilde{W}(\langle M\rangle_t) \quad \text{for all } t \geq 0,$$

where $\langle M\rangle$ is the quadratic variation process associated with M, that is,

$$\langle M\rangle_t \;=\; \begin{cases} \int_0^t \tilde{\sigma}_{ij}^2(\tilde{s})\, d\tilde{s} & \text{if } t \in [0,T], \\ \int_0^T \tilde{\sigma}_{ij}^2(\tilde{s})\, d\tilde{s} + (t{-}T) & \text{if } t > T. \end{cases}$$

Consequently,

$$\begin{aligned} \mathbf{E}\left(\sup_{t,s\in[0,T],|t-s|\leq h}\Big|\int_s^t \tilde{\sigma}_{ij}(\tilde{s})dW^{(j)}(\tilde{s})\Big|^p\right) \;&=\; \mathbf{E}\left(\sup_{t,s\in[0,T],|t-s|\leq h}\big|M(t)-M(s)\big|^p\right) \\ =\; \mathbf{E}\left(\sup_{t,s\in[0,T],|t-s|\leq h}\big|\tilde{W}(\langle M\rangle_t)-\tilde{W}(\langle M\rangle_s)\big|^p\right)& \\ \leq\; \mathbf{E}\left(\sup_{t,s\in[0,(K^2+1)T],|t-s|\leq(K^2+1)h}\big|\tilde{W}(t)-\tilde{W}(s)\big|^p\right)& \end{aligned}$$

as it holds that, $\mathbf{P}$-almost surely, $|\langle M\rangle_t - \langle M\rangle_s| \leq K^2|t{-}s| \leq (K^2{+}1)|t{-}s|$ for all $t, s \in [0,T]$. The assertion is now a consequence of Lemma A.2, which gives an upper bound for the moments of the modulus of continuity for standard one-dimensional Brownian motion. □

Lemma A.2. *Let $\tilde{W}$ be a standard one-dimensional Brownian motion living on the probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Then for every $p > 0$, every $T > 0$ there is a constant $\tilde{C}_{p,T}$ such that*

$$\mathbf{E}\left(\sup_{t,s\in[0,T],|t-s|\leq h}\big|\tilde{W}(t)-\tilde{W}(s)\big|^p\right) \;\leq\; \tilde{C}_{p,T}\left(h\ln(\tfrac{1}{h})\right)^{\frac{p}{2}} \quad \textit{for all } h \in (0,\tfrac{1}{2}].$$

Proof. As announced above, the main ingredient in the proof is an inequality due to Garsia, Rodemich, and Rumsey; it allows us to get an upper bound for $|\tilde{W}(t)(\omega) - \tilde{W}(s)(\omega)|^p$ in terms of $\omega \in \Omega$, T and the distance $|t{-}s|$. To this end, we define two strictly increasing functions Ψ, μ on $[0,\infty)$ by

$$\Psi(x) := \exp\big(\tfrac{x^2}{2}\big) - 1, \qquad \mu(x) := \sqrt{2x}, \qquad x \in [0,\infty).$$

Instead of μ we could have taken any function of the form $x \mapsto c\sqrt{x}$ provided $c > 1$; as one may expect, the resulting constant $\tilde{C}_{p,T}$ would be different. Clearly,

$$\Psi(0) = 0 = \mu(0), \quad \Psi^{-1}(y) = \sqrt{2\ln(y{+}1)} \quad \text{for all } y \geq 0, \quad d\mu(x) = \mu(dx) = \frac{dx}{\sqrt{2x}}.$$

In order to prepare for the application of the Garsia-Rodemich-Rumsey inequality, we set

$$\xi(\omega) := \int_0^T \int_0^T \Psi\left(\frac{|\tilde{W}(t)(\omega) - \tilde{W}(s)(\omega)|}{\mu(|t-s|)}\right) ds\, dt, \quad \omega \in \Omega,$$

thus defining an $\mathcal{F}$-measurable random variable with values in $[0, \infty]$. Since $\tilde{W}(t) - \tilde{W}(s)$ has normal distribution with mean zero and variance $|t-s|$, we see that

$$\begin{aligned}
\mathbf{E}(\xi) &= \mathbf{E}\left(\int_0^T \int_0^T \exp\left(\frac{|\tilde{W}(t) - \tilde{W}(s)|^2}{4|t-s|}\right) ds\, dt\right) - T^2 \\
&= \int_0^T \int_0^T \mathbf{E}\left(\exp\left(\frac{|\tilde{W}(t) - \tilde{W}(s)|^2}{4|t-s|}\right)\right) ds\, dt - T^2 \\
&= \frac{1}{\sqrt{2\pi}} \int_0^T \int_0^T \left(\frac{1}{\sqrt{|t-s|}} \int_{-\infty}^{\infty} \exp\left(\frac{u^2}{4|t-s|} - \frac{u^2}{2|t-s|}\right) du\right) ds\, dt - T^2 \\
&= \frac{1}{\sqrt{2\pi}} \int_0^T \int_0^T \left(\frac{1}{\sqrt{|t-s|}} \sqrt{2\pi}\sqrt{2|t-s|}\right) ds\, dt - T^2 \quad = \quad (\sqrt{2}-1)T^2 \; < \; \infty,
\end{aligned}$$

that is, ξ has finite expectation. In particular, $\xi(\omega) < \infty$ for $\mathbf{P}$-almost all $\omega \in \Omega$. The Garsia-Rodemich-Rumsey inequality now implies that for all $\omega \in \Omega$, all $t, s \in [0, T]$,

$$\left|\tilde{W}(t)(\omega) - \tilde{W}(s)(\omega)\right| \quad \leq \quad 8 \int_0^{|t-s|} \Psi^{-1}\left(\frac{4\xi(\omega)}{x^2}\right) \mu(dx) \; = \; 8 \int_0^{|t-s|} \sqrt{2\ln\left(\tfrac{4\xi(\omega)}{x^2} + 1\right)} \, \frac{dx}{\sqrt{2x}}.$$

Notice that if $\xi(\omega) = \infty$ then the above inequality is trivially satisfied. With $h \in (0, \frac{1}{2}]$ we have

$$\begin{aligned}
&\sup_{t,s \in [0,T], |t-s| \leq h} \left|\tilde{W}(t)(\omega) - \tilde{W}(s)(\omega)\right| \quad \leq \quad 8 \int_0^h \sqrt{\ln\bigl(4\xi(\omega) + x^2\bigr) + 2\ln(\tfrac{1}{x})} \, \frac{dx}{\sqrt{x}} \\
&\leq \; 8\sqrt{\ln(4\xi(\omega)+1)} \int_0^h \frac{dx}{\sqrt{x}} \; + \; 8\sqrt{2} \int_0^h \left(\sqrt{\ln(\tfrac{1}{x})} - \frac{1}{\sqrt{\ln(\frac{1}{x})}} + \frac{1}{\sqrt{\ln(\frac{1}{x})}}\right) \frac{dx}{\sqrt{x}} \\
&= \; 16\sqrt{h}\sqrt{\ln(4\xi(\omega)+1)} \; + \; 16\sqrt{2h\ln(\tfrac{1}{h})} \; + \; 8\sqrt{2} \int_0^h \frac{dx}{\sqrt{x\ln(\frac{1}{x})}} \\
&\leq \; 16\left(\sqrt{\ln(4\xi(\omega)+1)} \; + \; \sqrt{2}\sqrt{\ln(\tfrac{1}{h})} \; + \; \sqrt{\tfrac{2}{\ln(2)}}\right) \sqrt{h} \\
&\leq \; 32\left(\sqrt{\ln(4\xi(\omega)+1)} + 2\right) \sqrt{h\ln(\tfrac{1}{h})}.
\end{aligned}$$

Consequently, for all $p > 0$, all $h \in (0, \frac{1}{2}]$,

$$\mathbf{E}\left(\sup_{t,s \in [0,T], |t-s| \leq h} |\tilde{W}(t) - \tilde{W}(s)|^p\right) \; \leq \; 32^p\, \mathbf{E}\left(\left(\sqrt{\ln(4\xi+1)} + 2\right)^p\right) \left(h\ln(\tfrac{1}{h})\right)^{\frac{p}{2}}.$$

The above inequality yields the assertion provided we can show that the expectation on the right-hand side is finite. But this is the case, because

$$\mathbf{E}\left(\left(\sqrt{\ln(4\xi+1)}+2\right)^p\right) \quad\leq\quad 2^p\,\mathbf{E}\left(\left(\ln(4\xi+1)\right)^{\frac{p}{2}}\right)+4^p$$

and the expectation on the right-hand side of the last inequality is finite, as $\mathbf{E}(\xi)<\infty$ and $\ln(x+1)\leq x^{\frac{2}{p}}$ for all $x\geq 0$ big enough. More precisely, if $p\geq 1$, then $\ln(x)\leq(\ln(p)+1)\cdot x^{\frac{1}{p}}$ for all $x\geq(e\cdot p)^p$, whence

$$\begin{aligned}\mathbf{E}\left(\left(\ln(4\xi+1)\right)^{\frac{p}{2}}\right) \quad&\leq\quad \sqrt{1+\ln(p)}\,\mathbf{E}\left(\sqrt{4\xi+1}\right) \;+\; \big(p\ln(p)+p\big)^{\frac{p}{2}}\\ \leq\quad \sqrt{1+\ln(p)}&\left(\sqrt{1+4(\sqrt{2}-1)T^2}\right) \;+\; \big(p\ln(p)+p\big)^{\frac{p}{2}}\\ \leq\quad 2\big(p\ln(p)+p\big)^{\frac{p}{2}}&\big(1+T\big).\end{aligned}$$

Therefore, the asserted inequality follows for $p\geq 1$, where the constant $C_{p,T}$ need not be greater than

$$256^p\big(p\ln(p)+p\big)^{\frac{p}{2}}\big(1+T\big).$$

On the other hand, if $p\in(0,1)$, then clearly

$$\begin{aligned}\mathbf{E}\left(\left(\sqrt{\ln(4\xi+1)}+2\right)^p\right) \quad&\leq\quad \sqrt{\mathbf{E}\big(\ln(4\xi+1)\big)} \;+\; 2\\ \leq\quad 2\left(\sqrt{\mathbf{E}(\xi)}+1\right) \quad&\leq\quad 2\big(T+1\big),\end{aligned}$$

and the constant $C_{p,T}$, $p\in(0,1)$, need not be greater than $2\cdot 32^p(1+T)$. □

A.3 Proofs of "constant coefficients" error bounds

The first result we give here is a reduced version, adapted to our notation, of Theorem 2.7 in Krylov (2001). It provides an estimate of the error in approximating constant-coefficient controlled Itô diffusions by diffusions with piecewise constant strategies. The error is measured in terms of cost-functional-like expectations with Lipschitz (or Hölder) coefficients; see Section 1 in Krylov (2001) for a discussion of various error criteria. In the deterministic case, better error bounds can be obtained, see Lemmata A.3 and A.4 below.

Let $((\Omega,\mathcal{F},\mathbf{P}),(\mathcal{F}_t),W)$ be a Wiener basis of dimension d_1 in the sense of Definition 3.1. As above, let (Γ,ρ) be a complete and separable metric space, and denote by $\mathcal{U}$ the set of all $(\mathcal{F}_t)$-progressively measurable processes $[0,\infty)\times\Omega\to\Gamma$. For $n\in\mathbb{N}$, let $\mathcal{U}_n$ be the subset of $\mathcal{U}$ given by (3.10). Thus, if $\bar{u}\in\mathcal{U}_n$, then $\bar{u}$ is right-continuous and piecewise constant in time relative to the grid $\{k\frac{r}{n}\,|\,k\in\mathbb{N}_0\}$ and $\bar{u}(t)$ is measurable with respect to the σ-algebra generated by $W(k\frac{r}{n})$, $k=0,\ldots,\lfloor t\frac{n}{r}\rfloor$. We have incorporated the delay length r in the partition in order to be coherent with the notation of Section 3.3. In the original work by Krylov (2001), there is no delay and the time grid has mesh size $\frac{1}{n}$ instead of $\frac{r}{n}$.

Let $\tilde{b}\colon\Gamma\to\mathbb{R}^d$, $\tilde{\sigma}\colon\Gamma\to\mathbb{R}^{d\times d_1}$ be continuous functions with $|\tilde{b}|$, $|\tilde{\sigma}|$ bounded by K. For $u\in\mathcal{U}$ denote by X^u the process

$$X^u(t) := \int_0^t \tilde{b}\big(u(s)\big)ds \;+\; \int_0^t \tilde{\sigma}\big(u(s)\big)dW(s), \quad t\geq 0.$$

Let us write $|.|_\Gamma$ for the supremum norm of a real-valued function over Γ. Let us write $|.|_1$ for the Lipschitz norm of a real-valued function defined on $\mathbb{R}^d$. Thus, if $\tilde{g}$ is a Lipschitz continuous function $\mathbb{R}^d \to \mathbb{R}$, then

$$|\tilde{g}|_1 := \sup_{x,y\in\mathbb{R}^d, x\neq y} \frac{|\tilde{g}(x)-\tilde{g}(y)|}{|x-y|}.$$

The following theorem provides an error estimate for the approximation of a process X^u, where $u \in \mathcal{U}$, by processes X^{u_n}, $n \in \mathbb{N}$, where $u_n \in \mathcal{U}_n$, in terms of suitable cost functionals.

Theorem A.2 (Krylov). *Let $\bar{T} > 0$. There is a constant $\bar{C} > 0$ depending only on K and the dimensions such that the following holds: For any $n \in \mathbb{N}$ such that $n \geq r$, any bounded continuous function $\tilde{f}\colon \Gamma \to \mathbb{R}$, any bounded Lipschitz continuous function $\tilde{g}\colon \mathbb{R}^d \to \mathbb{R}$, any $u \in \mathcal{U}$ there exists $u_n \in \mathcal{U}_n$ such that*

$$\begin{aligned} &\mathbf{E}\left(\int_0^{\bar{T}} \tilde{f}\big(u_n(s)\big)ds + \tilde{g}\big(X^{u_n}(\bar{T})\big)\right) - \mathbf{E}\left(\int_0^{\bar{T}} \tilde{f}\big(u(s)\big)ds + \tilde{g}\big(X^{u}(\bar{T})\big)\right) \\ \leq\ & \bar{C}(1+\bar{T}) \left(\frac{r}{n}\right)^{\frac{1}{4}} \left(\left(\frac{r}{n}\right)^{\frac{1}{4}} |\tilde{f}|_\Gamma + |g|_1 \right). \end{aligned}$$

Note that in Theorem A.2 the difference between the two expectations may be inverted, since we can take $-\tilde{f}$ in place of $\tilde{f}$ and $-\tilde{g}$ in place of $\tilde{g}$.

Proof. Let $n \in \mathbb{N}$ such that $\frac{r}{n} \leq 1$. Define an (extended) cost functional $\bar{J}$ on $\mathbb{R} \times \mathbb{R}^d \times \mathcal{U}$ by

$$\bar{J}(t,x,u) := \begin{cases} \mathbf{E}\left(\int_0^{\bar{T}-t} \tilde{f}\big(u(s)\big)ds + \tilde{g}\big(x+X^u(\bar{T}-t)\big)\right) & \text{if } t < \bar{T}, \\ \tilde{g}(x) & \text{if } t \geq \bar{T}. \end{cases}$$

Let $\bar{V}_n$ be the value function arising from minimising $\bar{J}$ over $\mathcal{U}_n$, that is,

$$\bar{V}_n(t,x) := \inf_{u\in\mathcal{U}_n} \bar{J}(t,x,u), \quad (t,x) \in \mathbb{R}\times\mathbb{R}^d.$$

To prove the assertion it is enough to show that for all $u \in \mathcal{U}$, $x \in \mathbb{R}^d$,

$$\bar{V}_n(0,x) - \bar{J}(0,x,u) \quad \leq \quad \bar{C}(1+\bar{T}) \left(\frac{r}{n}\right)^{\frac{1}{4}} \left(\left(\frac{r}{n}\right)^{\frac{1}{4}} |\tilde{f}|_\Gamma + |\tilde{g}|_1 \right).$$

Indeed, it suffices to verify the above inequality for $x = 0 \in \mathbb{R}^d$, because we may consider the "translated" problem with $\tilde{g}(x + .)$ in place of $\tilde{g}(.)$, leaving the other functions $\tilde{f}$, $\tilde{b}$, $\tilde{\sigma}$ unchanged. Hence, it suffices to show that

$$(\star) \qquad \bar{V}_n(0,0) \quad \leq \quad \bar{J}(0,0,u) + \bar{C}(1+\bar{T}) \left(\frac{r}{n}\right)^{\frac{1}{4}} \left(\left(\frac{r}{n}\right)^{\frac{1}{4}} |\tilde{f}|_\Gamma + |\tilde{g}|_1 \right) \quad \text{for all } u \in \mathcal{U}.$$

We take note of the following properties of the discrete value function $\bar{V}_n$, cf. Lemma 3.1 in Krylov (2001).

1. Lipschitz continuity in space: for all $t \in \mathbb{R}$, $x, y \in \mathbb{R}^d$,

   $$\big|\bar{V}_n(t,x) - \bar{V}_n(t,y)\big| \ \leq\ |\tilde{g}|_1\, |x-y|.$$

   This is clear from the observation that $|\bar{V}_n(t,x) - \bar{V}_n(t,y)|$ is bounded by the supremum of $|\bar{J}(t,x,u) - \bar{J}(t,y,u)|$ over $u \in \mathcal{U}_n$ and the definition of $\bar{J}$.

2. One-step Principle of Dynamic Programming: for all $t \leq \bar{T} - \frac{r}{n}$, $x \in \mathbb{R}^d$,

$$\bar{V}_n(t,x) \;=\; \inf_{\gamma\in\Gamma} \mathbf{E}\left(\frac{r}{n}\,\tilde{f}(\gamma) \;+\; \bar{V}_n\big(t+\tfrac{r}{n}, x+X^{\gamma}\big(\tfrac{r}{n}\big)\big)\right).$$

   This is a consequence of Theorem A.1 in Appendix A.1. As will be seen below, it is actually enough to have an upper bound for $\bar{V}_n$, that is, to have the one-step Dynamic Programming Inequality with "$\leq$" in place of "$=$".

3. Hölder continuity in time: for all $t, s \leq \bar{T}$, $x \in \mathbb{R}^d$,

$$\big|\bar{V}_n(t,x) - \bar{V}_n(s,x)\big| \quad\leq\quad |\tilde{f}|_\Gamma\,|t-s| \;+\; K\,|\tilde{g}|_1\big(\sqrt{|t-s|} + \sqrt{d_1}\big)\sqrt{|t-s|}.$$

   To check this property, notice that $|\bar{V}_n(t,x) - \bar{V}_n(s,x)|$ is bounded by the supremum of $|\bar{J}(t,x,u) - \bar{J}(s,x,u)|$ over $u \in \mathcal{U}_n$. Now, for $u \in \mathcal{U}$, it holds that

$$\begin{aligned}\big|\bar{J}(t,x,u) - \bar{J}(s,x,u)\big| \quad&\leq\quad |\tilde{f}|_\Gamma\,|t-s| \;+\; |\tilde{g}|_1\,\mathbf{E}\left(|X^u(\bar{T}-t) - X^u(\bar{T}-s)|\right)\\ \leq\quad |\tilde{f}|_\Gamma\,|t-s| \;+\; |\tilde{g}|_1&\left(K\,|t-s| + K\sqrt{d_1}\sqrt{|t-s|}\right).\end{aligned}$$

A main difficulty in estimating the error arising from time-discretisation of the strategies is due to the fact that neither the discrete value function $\bar{V}_n$ nor the original value function $\bar{V}$ are necessarily differentiable. Krylov's idea for overcoming this problem is to consider a family of mollified functions $(\bar{V}_n^{(\varepsilon)})_{\varepsilon\in(0,1]}$ in place of $\bar{V}_n$. The Hölder and Lipschitz regularity of $\bar{V}_n$ translate into bounds on the partial derivatives of $\bar{V}_n^{(\varepsilon)}$, which in turn serve to estimate the discretisation error for the mollified value functions; because of the smoothness of the functions $\bar{V}_n^{(\varepsilon)}$, Itô's formula can be applied. Also the error between $\bar{V}_n$ and $\bar{V}_n^{(\varepsilon)}$ has to be estimated. Finally, to equate the two error bounds, one chooses the mollification paramater ε of the right order in $\frac{r}{n}$. The idea of using the Principle of Dynamic Programming to get from a local to a global error bound re-appears.

Let $\eta \in \mathbf{C}^\infty(\mathbb{R})$, $\xi \in \mathbf{C}^\infty(\mathbb{R}^d)$ be non-negative real-valued functions with unit integral and compact support; assume that $\eta(t) = 0$ for $t \in \mathbb{R} \setminus (0,1)$. For $\varepsilon \in (0,1]$ define

$$\begin{aligned}\eta_\varepsilon(t) &:= \varepsilon^{-1}\,\eta\big(\tfrac{1}{\varepsilon}\,t\big), \quad t\in\mathbb{R}, \qquad\qquad \xi_\varepsilon(x) := \varepsilon^{-d}\,\xi\big(\tfrac{1}{\varepsilon}\,x\big), \quad x\in\mathbb{R}^d,\\ \zeta_\varepsilon(t,x) &:= \eta_{\varepsilon^2}(t)\,\xi_\varepsilon(x), \quad (t,x)\in\mathbb{R}\times\mathbb{R}^d.\end{aligned}$$

Notice the different scaling in time and space as regards the functions ζ_ε. Define the mollified discrete value function with parameter ε as

$$\bar{V}_n^{(\varepsilon)} := \bar{V}_n * \zeta_\varepsilon, \text{ i.e., } \bar{V}_n^{(\varepsilon)}(t,x) \;=\; \int_{\mathbb{R}}\int_{\mathbb{R}^d} \zeta_\varepsilon(t-s, x-y)\,\bar{V}_n(s,y)\,dy\,ds, \quad (t,x)\in\mathbb{R}\times\mathbb{R}^d.$$

Denote by $\bar{V}_n^{\varepsilon}$ the discrete value function with parameter ε which is mollified only in the space variable, that is,

$$\bar{V}_n^{\varepsilon}(t,x) \;=\; \int_{\mathbb{R}^d} \xi_\varepsilon(x-y)\,\bar{V}_n(t,y)\,dy, \quad (t,x)\in\mathbb{R}\times\mathbb{R}^d.$$

The function $\bar{V}_n^{(\varepsilon)}$, i. e. the mollification of $\bar{V}_n$ in time and space, is in $\mathbf{C}^\infty(\mathbb{R}\times\mathbb{R}^d)$ and has bounded partial derivatives of all orders. The following estimates on the partial derivatives will be needed. The constants $C_1, \ldots, C_6$ that will appear in the estimates below depend only on K, the dimensions d, d_1 and the choice of the mollifiers η and ξ. Recall that $\varepsilon \in (0,1]$ and that η and ξ are $\mathbf{C}^\infty$-functions with unit integral and compact support, the support of η being contained in $[0,1]$. This implies, in particular, that the integrals $\int_0^1 \eta'(s)ds$, $\int_0^1 \eta''(s)ds$, $\int_{\mathrm{supp}(\xi)} D^l\xi(y)dy$ all equal zero, where $l > 0$ is the order of any partial derivatives in space.

1. Partial derivative in time of second order: for all $t \leq \bar{T}$, $x \in \mathbb{R}^d$,

$$\begin{aligned}
\Big|\frac{\partial^2}{\partial t^2}\bar{V}_n^{(\varepsilon)}(t,x)\Big| \quad &= \quad \Big|\frac{\partial^2}{\partial t^2}\int_{-\infty}^{\infty}\varepsilon^{-2}\,\eta\big(\tfrac{t-s}{\varepsilon^2}\big)\bar{V}_n^{\varepsilon}(s,x)\,ds\Big| \\
= \quad & \varepsilon^{-6}\Big|\int_{t-\varepsilon^2}^{t}\eta''\big(\tfrac{t-s}{\varepsilon^2}\big)\left(\int_{\mathbb{R}^d}\xi_\varepsilon(x-y)\,\bar{V}_n(s,y)\,dy\right)ds\Big| \\
= \quad & \varepsilon^{-6}\Big|\int_0^{\varepsilon^2}\eta''\big(\tfrac{s}{\varepsilon^2}\big)\left(\int_{\mathbb{R}^d}\xi_\varepsilon(x-y)\,\big(\bar{V}_n(t-s,y)-\bar{V}_n(t,y)\big)\,dy\right)ds\Big| \\
\leq \quad & \varepsilon^{-6}\int_0^{\varepsilon^2}\big|\eta''\big(\tfrac{s}{\varepsilon^2}\big)\big|\left(|\tilde{f}|_\Gamma|s| + (1+\sqrt{d_1})K|\tilde{g}|_1\sqrt{|s|}\right)\left(\int_{\mathbb{R}^d}\xi_\varepsilon(y)dy\right)ds \\
\leq \quad & \varepsilon^{-6}\left(\varepsilon|\tilde{f}|_\Gamma + (1+\sqrt{d_1})K|\tilde{g}|_1\right)\int_0^{\varepsilon^2}\big|\eta''\big(\tfrac{s}{\varepsilon^2}\big)\big|\sqrt{|s|}ds \\
\leq \quad & \varepsilon^{-6}\left(\varepsilon|\tilde{f}|_\Gamma + (1+\sqrt{d_1})K|\tilde{g}|_1\right)\varepsilon^3\int_0^1\big|\eta''(s)\big|\sqrt{|s|}\,ds \quad \leq \quad C_1\,\varepsilon^{-3}\left(\varepsilon|\tilde{f}|_\Gamma + |\tilde{g}|_1\right).
\end{aligned}$$

2. Partial derivatives in space of order $l \in \{1,2,3,4\}$: for all $(t,x) \in \mathbb{R}\times\mathbb{R}^d$,

$$\begin{aligned}
\Big|D^l\bar{V}_n^{(\varepsilon)}(t,x)\Big| \quad &\leq \quad \sup_{s\in\mathbb{R}}\Big|D^l\bar{V}_n^{\varepsilon}(s,x)\Big| \\
= \quad & \sup_{s\in\mathbb{R}}\varepsilon^{-l-d}\Big|\int_{\mathbb{R}^d}(D^l\xi)\big(\tfrac{1}{\varepsilon}\,y\big)\,\big(\bar{V}_n(s,x-y)-\bar{V}_n(s,x)\big)\,dy\Big| \\
\leq \quad & \varepsilon^{-l-d}\int_{\mathbb{R}^d}\big|(D^l\xi)\big(\tfrac{1}{\varepsilon}\,y\big)\big|\,|y|\,|\tilde{g}|_1dy \quad = \quad \varepsilon^{-l-d}\,|\tilde{g}|_1\,\varepsilon^d\int_{\mathrm{supp}(\xi)}\big|(D^l\xi)(y)\big|\,|\varepsilon\,y|\,dy \\
\leq \quad & \varepsilon^{-l}\,|\tilde{g}|_1\,\varepsilon\sup_{y\in\mathrm{supp}\,(\xi)}\big|(D^l\xi)(y)\big|\int_{\mathrm{supp}\,(\xi)}|y|\,dy \quad \leq \quad C_2\,\varepsilon^{1-l}\,|\tilde{g}|_1.
\end{aligned}$$

3. Mixed partial derivatives of first order in time and order $l \in \{1,2\}$ in space: for all $(t,x) \in \mathbb{R}\times\mathbb{R}^d$,

$$\begin{aligned}
\Big|\frac{\partial}{\partial t}D^l\bar{V}_n^{(\varepsilon)}(t,x)\Big| \quad &= \quad \varepsilon^{-4}\Big|\int_0^{\varepsilon^2}\eta'\big(\tfrac{s}{\varepsilon^2}\big)\,(D^l\bar{V}_n^{\varepsilon})(t-s,x)\,ds\Big| \\
\leq \quad & \varepsilon^{-4}C_2\,\varepsilon^{1-l}\,|\tilde{g}|_1\,\varepsilon^2\int_0^1|\eta'(s)|ds \quad =: \quad C_3\,\varepsilon^{-l-1}\,|\tilde{g}|_1.
\end{aligned}$$

Itô's formula will presently be applied to get an upper bound for $\bar{V}_n^{(\varepsilon)}(0,0)$. To this purpose, for $\gamma \in \Gamma$, let $\mathcal{L}^\gamma$ be the second order partial differential operator

$$\frac{\partial}{\partial t} \;+\; \sum_{i,j=1}^{d}(\tilde{\sigma}\tilde{\sigma}^{\mathsf{T}})_{ij}(\gamma)\frac{\partial^2}{\partial x_i\partial x_j} \;+\; \sum_{i=1}^{d}\tilde{b}_i(\gamma)\frac{\partial}{\partial x_i}$$

acting on functions in $\mathbf{C}^2(\mathbb{R}\times\mathbb{R}^d)$. Let $u\in\mathcal{U}$ be any strategy. Itô's (or Dynkin's) formula then yields

$$\mathbf{E}\left(\bar V_n^{(\varepsilon)}\big(\bar T-\tfrac{r}{n}, X^u\big(\bar T-\tfrac{r}{n}\big)\big)\right) \;=\; \bar V_n^{(\varepsilon)}(0,0) \;+\; \mathbf{E}\left(\int_0^{\bar T-\frac{r}{n}} \mathcal{L}^{u(t)}\bar V_n^{(\varepsilon)}\big(t,X^u(t)\big)dt\right).$$

Let $t\leq \bar T-\frac{r}{n}$, $x\in\mathbb{R}^d$. As a consequence of the one-step PDP for $\bar V_n$, Fatou's lemma and Fubini's theorem, we have

$$\begin{aligned}
\bar V_n^{(\varepsilon)}(t,x) \;&=\; \int_{\mathbb{R}}\int_{\mathbb{R}^d} \zeta_\varepsilon(t-s,x-y)\,\bar V_n(s,y)\,dy\,ds\\
=\;& \int_{\mathbb{R}}\int_{\mathbb{R}^d} \zeta_\varepsilon(t-s,x-y)\inf_{\gamma\in\Gamma}\mathbf{E}\left(\frac{r}{n}\,\tilde f(\gamma) \;+\; \bar V_n\big(s+\tfrac{r}{n},y+X^\gamma\big(\tfrac{r}{n}\big)\big)\right)dy\,ds\\
\leq\;& \inf_{\gamma\in\Gamma}\left\{\frac{r}{n}\,\tilde f(\gamma) \;+\; \mathbf{E}\left(\int_{\mathbb{R}}\int_{\mathbb{R}^d}\zeta_\varepsilon(t-s,x-y)\,\bar V_n\big(s+\tfrac{r}{n},y+X^\gamma\big(\tfrac{r}{n}\big)\big)\,dy\,ds\right)\right\}\\
\leq\;& \inf_{\gamma\in\Gamma}\left\{\frac{r}{n}\,\tilde f(\gamma) \;+\; \mathbf{E}\Big(\bar V_n^{(\varepsilon)}\big(t+\tfrac{r}{n},x+X^\gamma\big(\tfrac{r}{n}\big)\big)\Big)\right\}.
\end{aligned}$$

Let $\gamma\in\Gamma$. Itô's formula and Fubini's theorem yield

$$\mathbf{E}\left(\bar V_n^{(\varepsilon)}\big(t+\tfrac{r}{n},x+X^\gamma\big(\tfrac{r}{n}\big)\big)\right) \;=\; \bar V_n^{(\varepsilon)}(t,x) \;+\; \int_0^{\frac{r}{n}}\mathbf{E}\left(\mathcal{L}^\gamma\bar V_n^{(\varepsilon)}\big(t+s,x+X^\gamma(s)\big)\right)ds.$$

This, together with the above Dynamic Programming inequality, implies that

$$\int_0^{\frac{r}{n}}\mathbf{E}\left(\mathcal{L}^\gamma\bar V_n^{(\varepsilon)}\big(t+s,x+X^\gamma(s)\big)\right)ds \;\geq\; -\frac{r}{n}\,\tilde f(\gamma).$$

Applying Itô's formula to $\mathcal{L}^\gamma\bar V_n^{(\varepsilon)}(t+.,x+.)$ we see that, for all $s\geq 0$,

$$\mathbf{E}\left(\mathcal{L}^\gamma\bar V_n^{(\varepsilon)}\big(t+s,x+X^\gamma(s)\big)\right) \;=\; \mathcal{L}^\gamma\bar V_n^{(\varepsilon)}(t,x) \;+\; \mathbf{E}\left(\int_0^s\mathcal{L}^\gamma\big(\mathcal{L}^\gamma\bar V_n^{(\varepsilon)}\big)\big(t+\tilde s,x+X^\gamma(\tilde s)\big)d\tilde s\right).$$

Therefore, for all $\gamma\in\Gamma$, $t\leq\bar T-\frac{r}{n}$, $x\in\mathbb{R}^d$ it holds that

$$\mathcal{L}^\gamma\bar V_n^{(\varepsilon)}(t,x) \;+\; \frac{n}{r}\int_0^{\frac{r}{n}}\mathbf{E}\left(\int_0^s\mathcal{L}^\gamma\big(\mathcal{L}^\gamma\bar V_n^{(\varepsilon)}\big)\big(t+\tilde s,x+X^\gamma(\tilde s)\big)\,d\tilde s\right)ds \;\geq\; -\tilde f(\gamma).$$

The differential operator $\mathcal{L}^\gamma\circ\mathcal{L}^\gamma$ is composed of the following partial derivatives: derivative in time of second order, second to fourth order derivatives in space, mixed derivatives of first order in time and first and second order in space. The above bounds on the partial derivatives of $V^{(\varepsilon)}$ therefore imply that, for all $\gamma\in\Gamma$, $s\leq\bar T$, $y\in\mathbb{R}^d$,

$$\mathcal{L}^\gamma\big(\mathcal{L}^\gamma\bar V_n^{(\varepsilon)}\big)(s,y) \;\leq\; C_4\,\varepsilon^{-3}\left(\varepsilon|\tilde f|_\Gamma+|\tilde g|_1\right),$$

where $C_4:=\max\{C_1,C_2,C_3\}$. Notice that $\varepsilon^{-3}\geq\varepsilon^l$ for all $l\geq-3$ since $\varepsilon\in(0,1]$. Using the above inequality, we obtain

$$\mathcal{L}^\gamma\bar V_n^{(\varepsilon)}(t,x) \;\geq\; -\tilde f(\gamma) \;-\; \frac{r}{n}\,C_4\,\varepsilon^{-3}\left(\varepsilon|\tilde f|_\Gamma+|\tilde g|_1\right) \quad \text{for all } \gamma\in\Gamma,\ t\leq\bar T-\tfrac{r}{n},\ x\in\mathbb{R}^d.$$

Recall that

$$\bar V_n^{(\varepsilon)}(0,0) \;=\; -\mathbf{E}\left(\int_0^{\bar T-\frac{r}{n}} \mathcal{L}^{u(t)}\bar V_n^{(\varepsilon)}\big(t,X^u(t)\big)dt\right) \;+\; \mathbf{E}\left(\bar V_n^{(\varepsilon)}\big(\bar T-\tfrac{r}{n},X^u\big(\bar T-\tfrac{r}{n}\big)\big)\right),$$

where $u\in\mathcal{U}$ is an arbitrary strategy. The above lower bound for $\mathcal{L}^\gamma \bar V_n^{(\varepsilon)}$ translates into

$$\begin{aligned}\bar V_n^{(\varepsilon)}(0,0) \quad&\le\quad \mathbf{E}\left(\int_0^{\bar T-\frac{r}{n}} \tilde f\big(u(t)\big)dt\right) \;+\; \mathbf{E}\left(\bar V_n^{(\varepsilon)}\big(\bar T-\tfrac{r}{n},X^u\big(\bar T-\tfrac{r}{n}\big)\big)\right)\\ &\qquad +\; \bar T\,\frac{r}{n}\,C_4\,\varepsilon^{-3}\left(\varepsilon|\tilde f|_\Gamma+|\tilde g|_1\right).\end{aligned}$$

On the other hand, $\bar V_n^{(\varepsilon)}$ is close to $\bar V_n$; more precisely, for $(t,x)\in\mathbb{R}\times\mathbb{R}^d$,

$$\begin{aligned}\left|\bar V_n^{(\varepsilon)}(t,x)-\bar V_n(t,x)\right| \quad&\le\quad \int_{\mathbb{R}}\int_{\mathbb{R}^d}\zeta_\varepsilon(s,y)\left|\bar V_n(t-s,x-y)-\bar V_n(t,x)\right|dy\,ds\\ \le\quad& \int_{\mathbb{R}}\int_{\mathbb{R}^d}\zeta_\varepsilon(s,y)\left(|\tilde g|_1|y|+|\tilde f|_\Gamma|s|+K|\tilde g|_1\big(\sqrt{d_1}+\sqrt{|s|}\big)\sqrt{|s|}\right)dy\,ds\\ \le\quad& C_5\,\varepsilon\left(\varepsilon|\tilde f|_\Gamma+|\tilde g|_1\right).\end{aligned}$$

Combining the last two inequalities we get

$$\begin{aligned}\bar V_n(0,0) \quad&\le\quad \mathbf{E}\left(\int_0^{\bar T}\tilde f\big(u(t)\big)dt\right) \;+\; \mathbf{E}\left(\bar V_n\big(\bar T-\tfrac{r}{n},X^u\big(\bar T-\tfrac{r}{n}\big)\big)\right)\\ &\quad+\;\frac{r}{n}|\tilde f|_\Gamma \;+\; 2C_5\,\varepsilon\left(\varepsilon|\tilde f|_\Gamma+|\tilde g|_1\right) \;+\; \bar T\,\frac{r}{n}\,C_4\,\varepsilon^{-3}\left(\varepsilon|\tilde f|_\Gamma+|\tilde g|_1\right).\end{aligned}$$

Now observe that, for $x,y\in\mathbb{R}^d$,

$$\begin{aligned}\left|\bar V_n\big(\bar T-\tfrac{r}{n},y\big)-\tilde g(x)\right| \;&=\; \left|\bar V_n\big(\bar T-\tfrac{r}{n},y\big)-\bar V_n\big(\bar T,x\big)\right|\\ \le\quad |\tilde g|_1\,|x-y| \;+\; &\frac{r}{n}|\tilde f|_\Gamma \;+\; (1+\sqrt{d_1})K|\tilde g|_1\left(\frac{r}{n}\right)^{\frac12},\end{aligned}$$

whence

$$\begin{aligned}&\mathbf{E}\left(\bar V_n\big(\bar T-\tfrac{r}{n},X^u\big(\bar T-\tfrac{r}{n}\big)\big)\right)-\mathbf{E}\left(\tilde g\big(X^u(\bar T)\big)\right)\\ \le\quad& \frac{r}{n}|\tilde f|_\Gamma \;+\; (1+\sqrt{d_1})K|\tilde g|_1\left(\frac{r}{n}\right)^{\frac12} \;+\; |\tilde g|_1\,\mathbf{E}\left(\left|X^u\big(\bar T-\tfrac{r}{n}\big)-X^u\big(\bar T\big)\right|\right)\\ \le\quad& \frac{r}{n}|\tilde f|_\Gamma \;+\; (1+\sqrt{d_1})K|\tilde g|_1\left(\frac{r}{n}\right)^{\frac12} \;+\; (1+\sqrt{d_1})K|\tilde g|_1\left(\frac{r}{n}\right)^{\frac12}.\end{aligned}$$

Consequently, we have

$$\begin{aligned}\bar V_n(0,0) \quad&\le\quad \mathbf{E}\left(\int_0^{\bar T}\tilde f\big(u(t)\big)dt\right) \;+\; \mathbf{E}\left(\tilde g\big(X^u(\bar T)\big)\right) \;+\; C_6\left(\frac{r}{n}\right)^{\frac12}\left(\left(\frac{r}{n}\right)^{\frac12}|\tilde f|_\Gamma+|\tilde g|_1\right)\\ &\quad+\; 2C_5\,\varepsilon\left(\varepsilon|\tilde f|_\Gamma+|\tilde g|_1\right) \;+\; \bar T\,\frac{r}{n}\,C_4\,\varepsilon^{-3}\left(\varepsilon|\tilde f|_\Gamma+|\tilde g|_1\right).\end{aligned}$$

In order to equate the order of the error in the last two summands, set $\varepsilon := (\frac{r}{n})^{\frac{1}{4}}$. With this choice of ε and recalling the definition of $\bar{J}$, we find that

$$\begin{aligned}
\bar{V}_n(0,0) \;\leq\;& \bar{J}(0,0,u) \;+\; C_6 \left(\frac{r}{n}\right)^{\frac{1}{2}} \left(\left(\frac{r}{n}\right)^{\frac{1}{2}} |\tilde{f}|_\Gamma + |\tilde{g}|_1 \right) +\; 2C_5 \left(\frac{r}{n}\right)^{\frac{1}{4}} \left(\left(\frac{r}{n}\right)^{\frac{1}{4}} |\tilde{f}|_\Gamma + |\tilde{g}|_1 \right) \\
&+\; \bar{T}\,\frac{r}{n}\, C_4 \left(\frac{n}{r}\right)^{\frac{3}{4}} \left(\left(\frac{r}{n}\right)^{\frac{1}{4}} |\tilde{f}|_\Gamma + |\tilde{g}|_1 \right) \\
\leq\;& \bar{J}(0,0,u) \;+\; \bar{C}\,(\bar{T}{+}1) \left(\frac{r}{n}\right)^{\frac{1}{4}} \left(\left(\frac{r}{n}\right)^{\frac{1}{4}} |\tilde{f}|_\Gamma + |\tilde{g}|_1 \right),
\end{aligned}$$

where $u \in \mathcal{U}$ is arbitrary and the constant $\bar{C}$ can be chosen as $\max\{C_4, 2C_5, C_6\}$. Hence Inequality ($\star$) holds. □

We now turn to the deterministic case. Let $\hat{\mathcal{U}}$ denote the set of all *deterministic strategies*, that is, $\hat{\mathcal{U}}$ is the set of all measurable functions $[0,\infty) \to \Gamma$. For $n \in \mathbb{N}$, let $\hat{\mathcal{U}}_n$ be the subset of $\hat{\mathcal{U}}$ consisting of all right-continuous functions $[0,\infty) \to \Gamma$ which are piecewise constant relative to the grid $\{k\frac{r}{n} \,|\, k \in \mathbb{N}_0\}$. Again, we have incorporated the delay length r in the partition in order to be coherent with the notation of Section 3.3.

Let $\tilde{b}\colon \Gamma \to \mathbb{R}^d$ be a measurable function with $|\tilde{b}|$ bounded by K. For $u \in \hat{\mathcal{U}}$, denote by x^u the function

$$x^u(t) := \int_0^t \tilde{b}\big(u(s)\big)ds, \quad t \geq 0.$$

The following results provide error estimates for the approximation of a function x^u, where $u \in \hat{\mathcal{U}}$, by functions x^{u_n}, $n \in \mathbb{N}$, where $u_n \in \hat{\mathcal{U}}_n$, in terms of suitable cost functionals.

The result we state first should be compared to Theorem 2.1 in Falcone and Giorgi (1999) and also to Theorem A.2 above. Recall that the error bound in Theorem A.2 is of order $h^{1/4}$ in the time step $h = \frac{r}{n}$, while the bound for deterministic problems automatically improves to $h^{1/2}$.

Lemma A.3. *Let $\bar{T} > 0$. There is a constant $\bar{C} > 0$ depending only on K and the dimension d such that the following holds: For any $n \in \mathbb{N}$ such that $n \geq r$, any bounded measurable function $\tilde{f} : \Gamma \to \mathbb{R}$, any bounded Lipschitz continuous function $\tilde{g} : \mathbb{R}^d \to \mathbb{R}$, any $u \in \hat{\mathcal{U}}$ there exists $u_n \in \hat{\mathcal{U}}_n$ such that*

$$\begin{aligned}
&\int_0^{\bar{T}} \tilde{f}\big(u_n(s)\big)ds + \tilde{g}\big(x^{u_n}(\bar{T})\big) \;-\; \left(\int_0^{\bar{T}} \tilde{f}\big(u(s)\big)ds + \tilde{g}\big(x^{u}(\bar{T})\big) \right) \\
&\leq\; \bar{C}(1{+}\bar{T}) \left(\frac{r}{n}\right)^{\frac{1}{2}} \left(|\tilde{f}|_\Gamma + |\tilde{g}|_1 \right).
\end{aligned}$$

The proof of Lemma A.3 is – mutatis mutandis – completely parallel to the proof of Theorem A.2. Itô's formula has to be replaced by the usual change-of-variable formula, and the scaling relation between smoothing in time and smoothing in space must be modified, as would be expected, from ε vs. $\sqrt{\varepsilon}$ to ε vs. ε. Observe, however, that the proof of Theorem 2.1 in Falcone and Giorgi (1999) is different, as it relies on the theory of viscosity solutions.

If the space of control actions Γ is finite, then the following elementary arguments show that the approximation error is of order h in the length $h = \frac{r}{n}$ of the time step.

Lemma A.4. *Assume that Γ is finite with cardinality N_Γ. Let $\bar T > 0$. Then for any $n \in \mathbb{N}$ such that $n \cdot \bar T \geq N_\Gamma \cdot r$, any bounded measurable function $\tilde f \colon \Gamma \to \mathbb{R}$, any bounded Lipschitz continuous function $\tilde g \colon \mathbb{R}^d \to \mathbb{R}$, any $u \in \hat{\mathcal{U}}$ there exists $u_n \in \hat{\mathcal{U}}_n$ such that*

$$\begin{aligned} & \int_0^{\bar T} \tilde f\big(u_n(s)\big)ds + \tilde g\big(x^{u_n}(\bar T)\big) \;-\; \left(\int_0^{\bar T} \tilde f\big(u(s)\big)ds + \tilde g\big(x^{u}(\bar T)\big)\right) \\ \leq\;& \frac{r}{n}\,(1+N_\Gamma)\left(|\tilde f|_\Gamma + K|\tilde g|_1\right). \end{aligned}$$

Proof. By hypothesis, Γ has N_Γ elements, say $\Gamma = \{\gamma_1, \ldots, \gamma_{N_\Gamma}\}$. Let $n \in \mathbb{N}$ be such that $n \cdot \bar T \geq N_\Gamma \cdot r$. Clearly, for arbitrary $u \in \hat{\mathcal{U}}$, all $\bar u \in \hat{\mathcal{U}}_n$,

$$\begin{aligned} & \int_0^{\bar T} \tilde f\big(\bar u(s)\big)ds + \tilde g\big(x^{\bar u}(\bar T)\big) \;-\; \left(\int_0^{\bar T} \tilde f\big(u(s)\big)ds + \tilde g\big(x^{u}(\bar T)\big)\right) \\ \leq\;& \left|\int_0^{\bar T} \tilde f\big(\bar u(s)\big)ds - \int_0^{\bar T} \tilde f\big(u(s)\big)ds\right| \;+\; |\tilde g|_1 \left|\int_0^{\bar T} \tilde b\big(\bar u(s)\big)ds - \int_0^{\bar T} \tilde b\big(u(s)\big)ds\right|. \end{aligned}$$

Denoting by λ^1 Lebesgue measure on $\mathbb{R}$, we set

$$a_k :=\; \lambda^1\big\{s \in [0,\bar T] \mid u(s) = \gamma_k\big\}, \quad k \in \{1, \ldots, N_\Gamma\}.$$

Then, by definition of the Lebesgue integral,

$$\int_0^{\bar T} \tilde f\big(u(s)\big)ds \;=\; \sum_{k=1}^{N_\Gamma} a_k\, \tilde f(\gamma_k), \qquad \int_0^{\bar T} \tilde b\big(u(s)\big)ds \;=\; \sum_{k=1}^{N_\Gamma} a_k\, \tilde b(\gamma_k).$$

Notice that the integral over $\tilde f$ is just a real number, while the integral over $\tilde b$ is a point in $\mathbb{R}^d$. On the other hand, setting

$$j_k :=\; \#\big\{i \in \{1, \ldots, \lfloor \bar T \tfrac{n}{r}\rfloor - 1\} \mid \bar u(r\tfrac{i}{n}) = \gamma_k\big\}, \quad k \in \{1, \ldots, N_\Gamma\},$$

we have

$$\begin{aligned} \int_0^{\bar T} \tilde f\big(\bar u(s)\big)ds \;&=\; \left(\sum_{k=1}^{N_\Gamma} j_k \frac{r}{n}\, \tilde f(\gamma_k)\right) \;-\; \tilde f\big(\bar u(\tfrac{r}{n}\lfloor \bar T \tfrac{n}{r}\rfloor)\big)\left(\bar T - \tfrac{r}{n}\lfloor \bar T \tfrac{n}{r}\rfloor\right), \\ \int_0^{\bar T} \tilde b\big(\bar u(s)\big)ds \;&=\; \left(\sum_{k=1}^{N_\Gamma} j_k \frac{r}{n}\, \tilde b(\gamma_k)\right) \;-\; \tilde b\big(\bar u(\tfrac{r}{n}\lfloor \bar T \tfrac{n}{r}\rfloor)\big)\left(\bar T - \tfrac{r}{n}\lfloor \bar T \tfrac{n}{r}\rfloor\right). \end{aligned}$$

Consequently,

$$\begin{aligned} & \left|\int_0^{\bar T} \tilde f\big(\bar u(s)\big)ds - \int_0^{\bar T} \tilde f\big(u(s)\big)ds\right| \;+\; |\tilde g|_1 \left|\int_0^{\bar T} \tilde b\big(\bar u(s)\big)ds - \int_0^{\bar T} \tilde b\big(u(s)\big)ds\right| \\ \leq\;& \left(|\tilde f|_\Gamma + |\tilde g|_1 K\right)\left(\frac{r}{n} \;+\; \sum_{k=1}^{N_\Gamma} \left|a_k - j_k \frac{r}{n}\right|\right), \end{aligned}$$

where the hypothesis that $|b| \leq K$ has been used. Recall that $a_1, \ldots, a_{N_\Gamma}$ depend on $u \in \hat{\mathcal{U}}$, while $j_1, \ldots, j_{N_\Gamma}$ depend on the choice of $\bar u \in \hat{\mathcal{U}}_n$. Let us fix $u \in \hat{\mathcal{U}}$. Clearly, $a_k \geq 0$

and $\sum_{k=1}^{N_\Gamma} a_k = \bar{T}$. Define numbers $j_1, \ldots, j_{N_\Gamma}$ recursively by setting $j_1 := \lfloor \frac{n}{r} a_1 \rfloor$ and, if $N_\Gamma \geq 2$,

$$j_l := \left\lfloor \frac{n}{r} \sum_{k=1}^{l} a_k \right\rfloor - \sum_{k=1}^{l-1} j_k, \quad l \in \{2, \ldots, N_\Gamma\}.$$

With this definition, the numbers $j_1, \ldots, j_{N_\Gamma}$ are in $\{0, \ldots, \lfloor \frac{n}{r} \bar{T} \rfloor\}$ and

$$\sum_{k=1}^{N_\Gamma} j_k \quad = \quad j_{N_\Gamma} + \sum_{k=1}^{N_\Gamma - 1} j_k \quad = \quad \left\lfloor \frac{n}{r} \sum_{k=1}^{N_\Gamma} a_k \right\rfloor \quad = \quad \left\lfloor \frac{n}{r} \bar{T} \right\rfloor .$$

To estimate the difference between a_l and $\frac{r}{n} j_l$, $l \in \{1, \ldots, N_\Gamma\}$, note that

$$\left| a_1 - j_1 \frac{r}{n} \right| \quad = \quad \frac{r}{n} \cdot \left| \frac{n}{r} a_1 - \left\lfloor \frac{n}{r} a_1 \right\rfloor \right| \quad < \quad \frac{r}{n},$$

and observe that for all $a, \hat{a} \geq 0$,

$$\left| \hat{a} - \lfloor a + \hat{a} \rfloor + \lfloor a \rfloor \right| \quad = \quad \left| a + \hat{a} - \lfloor a + \hat{a} \rfloor - (a - \lfloor a \rfloor) \right| \quad < \quad 1.$$

Therefore, for all $l \in \{2, \ldots, N_\Gamma\}$,

$$\begin{aligned} \left| a_l - j_l \frac{r}{n} \right| \quad &= \quad \frac{r}{n} \cdot \left| \frac{n}{r} a_l - \left\lfloor \frac{n}{r} \sum_{k=1}^{l} a_k \right\rfloor - \sum_{k=1}^{l-1} j_k \right| \\ &= \quad \frac{r}{n} \cdot \left| \frac{n}{r} a_l - \left\lfloor \frac{n}{r} \sum_{k=1}^{l} a_k \right\rfloor + \left\lfloor \frac{n}{r} \sum_{k=1}^{l-1} a_k \right\rfloor \right| \quad < \quad \frac{r}{n}. \end{aligned}$$

It is clear that we can choose $\bar{u} \in \hat{\mathcal{U}}_n$ such that

$$j_k \;=\; \#\bigl\{ i \in \{1, \ldots, \lfloor \bar{T} \tfrac{n}{r} \rfloor - 1\} \mid \bar{u}(r \tfrac{i}{n}) = \gamma_k \bigr\} \quad \text{for all } k \in \{1, \ldots, N_\Gamma\}.$$

For example, we may define $\bar{u}$ to be equal to γ_1 on the interval $[0, \frac{r}{n} j_1)$, then to be equal to γ_2 on the interval $[\frac{r}{n} j_1, \frac{r}{n}(j_1 + j_2))$ and so on. In this way, given $u \in \hat{\mathcal{U}}$, we find $\bar{u} \in \hat{\mathcal{U}}_n$ such that

$$\begin{aligned} &\left| \int_0^{\bar{T}} \tilde{f}\bigl(\bar{u}(s)\bigr) ds - \int_0^{\bar{T}} \tilde{f}\bigl(u(s)\bigr) ds \right| \;+\; |\tilde{g}|_1 \left| \int_0^{\bar{T}} \tilde{b}\bigl(\bar{u}(s)\bigr) ds - \int_0^{\bar{T}} \tilde{b}\bigl(u(s)\bigr) ds \right| \\ \leq \quad &\Bigl(|\tilde{f}|_\Gamma + |\tilde{g}|_1 K \Bigr) \Bigl(\frac{r}{n} \;+\; N_\Gamma \frac{r}{n} \Bigr), \end{aligned}$$

which yields the assertion. □

Let us return a last time to the stochastic setting. We are interested in the case when the diffusion matrix is constant and the space of control actions Γ is finite. Let $((\Omega, \mathcal{F}, \mathbf{P}), (\mathcal{F}_t), W)$ be a Wiener basis of dimension d_1, $\mathcal{U}$ the set of all $(\mathcal{F}_t)$-progressively measurable processes $[0, \infty) \times \Omega \to \Gamma$, and $\mathcal{U}_n$ be the subset of strategies which are right-continuous and piecewise constant in time relative to the grid $\{k \frac{r}{n} \mid k \in \mathbb{N}_0\}$ and measurable with respect to the σ-algebra generated by $W(k \frac{r}{n})$, $k \in \mathbb{N}_0$, as above.

Let $\tilde{b}\colon \Gamma \to \mathbb{R}^d$ be a continuous function with $|\tilde{b}|$ bounded by K, and let σ be a $d \times d_1$ matrix. For $u \in \mathcal{U}$, denote by X^u the $\mathbb{R}^d$-valued process

$$X^u(t) := \int_0^t \tilde{b}\bigl(u(s)\bigr) ds \;+\; \sigma W(t), \quad t \geq 0.$$

The following result gives a bound on the discretisation error which is of order $\sqrt{h}$ in the time step $h = \frac{r}{n}$.

Lemma A.5. *Assume that* Γ *is finite with cardinality* N_Γ *and that the diffusion coefficient* σ *is a constant matrix. Let* $\bar{T} > 0$. *Then for any square number* $n \in \mathbb{N}$ *such that* $\sqrt{n} \cdot \bar{T} \geq N_\Gamma$, *any bounded measurable function* $\tilde{f} \colon \Gamma \to \mathbb{R}$, *any bounded Lipschitz continuous function* $\tilde{g} \colon \mathbb{R}^d \to \mathbb{R}$, *any* $u \in \mathcal{U}$ *there exists* $u_n \in \mathcal{U}_n$ *such that*

$$\begin{aligned} &\mathbf{E}\left(\int_0^{\bar{T}} \tilde{f}\big(u_n(s)\big)ds + \tilde{g}\big(X^{u_n}(\bar{T})\big)\right) \;-\; \mathbf{E}\left(\int_0^{\bar{T}} \tilde{f}\big(u(s)\big)ds + \tilde{g}\big(X^{u}(\bar{T})\big)\right) \\ \leq\;& \frac{1}{\sqrt{n}}\left(1 + 4r\cdot\bar{T} + N_\Gamma\right)\left(|\tilde{f}|_\Gamma + K|\tilde{g}|_1\right). \end{aligned}$$

Proof. Let $n \in \mathbb{N}$ be such that $\sqrt{n}\cdot\bar{T} \geq N_\Gamma$. Since σ is constant, we have for arbitrary $u \in \mathcal{U}$, all $\bar{u} \in \mathcal{U}_n$,

$$\begin{aligned} &\mathbf{E}\left(\int_0^{\bar{T}} \tilde{f}\big(\bar{u}(s)\big)ds + \tilde{g}\big(X^{\bar{u}}(\bar{T})\big)\right) \;-\; \mathbf{E}\left(\int_0^{\bar{T}} \tilde{f}\big(u(s)\big)ds + \tilde{g}\big(X^{u}(\bar{T})\big)\right) \\ \leq\;& \mathbf{E}\left(\left|\int_0^{\bar{T}} \tilde{f}\big(\bar{u}(s)\big)ds - \int_0^{\bar{T}} \tilde{f}\big(u(s)\big)ds\right| \;+\; |\tilde{g}|_1\left|\int_0^{\bar{T}} \tilde{b}\big(\bar{u}(s)\big)ds - \int_0^{\bar{T}} \tilde{b}\big(u(s)\big)ds\right|\right). \end{aligned}$$

Let $\omega \in \Omega$. Clearly,

$$\begin{aligned} \int_0^{\bar{T}} \tilde{f}\big(u(s,\omega)\big)ds \;&=\; \left(\sum_{k=1}^{\lfloor\frac{\sqrt{n}}{r}\rfloor} \int_{(k-1)\frac{r\cdot\bar{T}}{\sqrt{n}}}^{k\frac{r\cdot\bar{T}}{\sqrt{n}}} \tilde{f}\big(u(s,\omega)\big)ds\right) \;+\; \int_{\lfloor\frac{\sqrt{n}}{r}\rfloor\frac{r\cdot\bar{T}}{\sqrt{n}}}^{\bar{T}} \tilde{f}\big(u(s,\omega)\big)ds, \\ \int_0^{\bar{T}} \tilde{b}\big(u(s,\omega)\big)ds \;&=\; \left(\sum_{k=1}^{\lfloor\frac{\sqrt{n}}{r}\rfloor} \int_{(k-1)\frac{r\cdot\bar{T}}{\sqrt{n}}}^{k\frac{r\cdot\bar{T}}{\sqrt{n}}} \tilde{b}\big(u(s,\omega)\big)ds\right) \;+\; \int_{\lfloor\frac{\sqrt{n}}{r}\rfloor\frac{r\cdot\bar{T}}{\sqrt{n}}}^{\bar{T}} \tilde{b}\big(u(s,\omega)\big)ds. \end{aligned}$$

By Lemma A.4 and its proof, we can find a deterministic function $\hat{u}_\omega \in \hat{\mathcal{U}}_n$ such that for all $k \in \{1,\ldots,\lfloor\frac{\sqrt{n}}{r}\rfloor - 1\}$,

$$\begin{aligned} &\left|\int_{k\frac{r\cdot\bar{T}}{\sqrt{n}}}^{(k+1)\frac{r\cdot\bar{T}}{\sqrt{n}}} \tilde{f}\big(\hat{u}_\omega(s)\big)ds - \int_{(k-1)\frac{r\cdot\bar{T}}{\sqrt{n}}}^{k\frac{r\cdot\bar{T}}{\sqrt{n}}} \tilde{f}\big(u(s,\omega)\big)ds\right| \\ &+ |\tilde{g}|_1\left|\int_{k\frac{r\cdot\bar{T}}{\sqrt{n}}}^{(k+1)\frac{r\cdot\bar{T}}{\sqrt{n}}} \tilde{b}\big(\hat{u}_\omega(s)\big)ds - \int_{(k-1)\frac{r\cdot\bar{T}}{\sqrt{n}}}^{k\frac{r\cdot\bar{T}}{\sqrt{n}}} \tilde{b}\big(u(s,\omega)\big)ds\right|. \\ \leq\;& \frac{r}{n}\left(1 + N_\Gamma\right)\left(|\tilde{f}|_\Gamma + K|\tilde{g}|_1\right). \end{aligned}$$

Notice that, since n is a square number, the points of the grid of mesh size $\frac{r}{\sqrt{n}}$ are also part of the finer grid of mesh size $\frac{r}{n}$. The functions $\hat{u}_\omega$, $\omega \in \Omega$, can now be chosen in such a way that $\bar{u}(t,\omega) := \hat{u}_\omega(t)$, $t \geq 0$, $\omega \in \Omega$, defines an $\mathcal{F}_t$-progressively measurable piecewise constant Γ-valued process which is also measurable with respect to the σ-algebra generated by $W(k\frac{r}{n})$, $k \in \mathbb{N}_0$. Thus, $\bar{u}$ is a strategy in $\mathcal{U}_n$, and it holds that

$$\begin{aligned} &\mathbf{E}\left(\left|\int_0^{\bar{T}} \tilde{f}\big(\bar{u}(s)\big)ds - \int_0^{\bar{T}} \tilde{f}\big(u(s)\big)ds\right| \;+\; |\tilde{g}|_1\left|\int_0^{\bar{T}} \tilde{b}\big(\bar{u}(s)\big)ds - \int_0^{\bar{T}} \tilde{b}\big(u(s)\big)ds\right|\right) \\ \leq\;& \frac{\sqrt{n}}{r}\cdot\frac{r}{n}\left(1 + N_\Gamma\right)\left(|\tilde{f}|_\Gamma + K|\tilde{g}|_1\right) \;+\; 4\frac{r\cdot\bar{T}}{\sqrt{n}}\left(|\tilde{f}|_\Gamma + K|\tilde{g}|_1\right). \end{aligned}$$

□

Bibliography

S. Ankirchner, P. Imkeller, and A. Popier. Optimal cross hedging for insurance derivatives. Humboldt University, 2007.

K. Atkinson and W. Han. *Theoretical Numerical Analysis*, volume 39 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 2001.

M. Bardi and I. Capuzzo Dolcetta. *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman equations*, volume 4 of *Systems & Control: Foundations & Applications*. Birkhäuser, Boston, 1997.

G. Barles and E. R. Jakobsen. Error bounds for monotone approximation schemes for Hamilton-Jacobi-Bellman equations. *SIAM J. Numer. Anal.*, 43(2):540–558, 2005.

G. Barles and E. R. Jakobsen. Error bounds for monotone approximation schemes for parabolic Hamilton-Jacobi-Bellman equations. *Math. Comput.*, 76(260):1861–1893, 2007.

G. Barles and P. E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Anal.*, 4:271–283, 1991.

H. Bauer and U. Rieder. Stochastic control problems with delay. *Math. Meth. Oper. Res.*, 62(3):411–427, 2005.

R. Bellman and R. Kabala. *Dynamic Programming and Modern Control Theory.* Academic Press, New York, 1965.

A. Bensoussan, G. Da Prato, M. C. Delfour, and S. K. Mitter. *Representation and Control of Infinite-Dimensional Systems.* Systems & Control: Foundations & Applications. Birkhäuser, Boston, 2nd edition, 2007.

D. P. Bertsekas. *Dynamic Programming and Optimal Control*, volume 1. Athena Scientific, Belmont, Massachusetts, 3rd edition, 2005.

D. P. Bertsekas. *Dynamic Programming and Optimal Control*, volume 2. Athena Scientific, Belmont, Massachusetts, 3rd edition, 2007.

D. P. Bertsekas and S. E. Shreve. *Stochastic Optimal Control: The Discrete-Time Case.* Athena Scientific, Belmont, Massachusetts, reprint of the 1978 edition, 1996.

P. Billingsley. *Convergence of Probability Measures.* Wiley series in Probability and Statistics. John Wiley & Sons, New York, 2nd edition, 1999.

R. Boucekkine, O. Licandro, L. A. Puch, and F. del Rio. Vintage capital and the dynamics of the AK model. *J. Econ. Theory*, 120:39–72, 2005.

E. Buckwar. Introduction to the numerical analysis of stochastic delay differential equations. *J. Comput. Appl. Math.*, 125(1-2):297–307, 2000.

A. Calzolari, P. Florchinger, and G. Nappo. Convergence in nonlinear filtering for stochastic delay systems. *SIAM J. Control Optim.*, 46(5):1615–1636, 2007.

I. Capuzzo Dolcetta and M. Falcone. Discrete dynamic programming and viscosity solutions of the Bellman equation. *Ann. Inst. Henri Poincaré, Anal. Non Linéaire*, 6(Suppl.): 161–183, 1989.

I. Capuzzo Dolcetta and H. Ishii. Approximate solutions of the Bellman equation of deterministic control theory. *Appl. Math. Optim.*, 11(2):161–181, 1984.

M.-H. Chang, T. Pang, and M. Pemy. Stochastic optimal control problems with a bounded memory. In X. Zhang, D. Liu, and L. Wu, editors, *Operations Research and Its Applications. Papers from the Sixth International Symposium, ISORA'06, Xinjiang, China, August 8-12*, pages 82–94, Beijing, 2006. World Publishing Corporation.

C.-S. Chow and J. N. Tsitsiklis. The complexity of dynamic programming. *J. Complexity*, 5:466–488, 1989.

G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions*, volume 45 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1992.

I. Elsanosi, B. Øksendal, and A. Sulem. Some solvable stochastic control problems with delay. *Stochastics Stochastics Rep.*, 71(1-2):69–89, 2000.

S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley Series in Probability and Statistics. John Wiley & Sons, New York, 1986.

M. Falcone and R. Ferretti. Discrete time high-order schemes for viscosity solutions of Hamilton-Jacobi-Bellman equations. *Numer. Math.*, 67(3):315–344, 1994.

M. Falcone and T. Giorgi. An approximation scheme for evolutive Hamilton-Jacobi equations. In W. McEneaney, G. Yin, and Q. Zhang, editors, *Stochastic Analysis, Control, Optimization and Applications. A Volume in Honor of W. H. Fleming*, pages 289–303, Boston, 1999. Birkhäuser.

M. Falcone and R. Rosace. Discrete-time approximation of optimal control problems for delayed equations. *Control Cybern.*, 25(3):665–675, 1996.

M. Fischer and G. Nappo. Time discretisation and rate of convergence for the optimal control of continuous-time stochastic systems with delay. *Appl. Math. Optim.*, to appear, 2007.

M. Fischer and M. Reiß. Discretisation of stochastic control problems for continuous time dynamics with delay. *J. Comput. Appl. Math.*, 205(2):969–981, 2007.

W. H. Fleming and H. M. Soner. *Controlled Markov Processes and Viscosity Solutions*, volume 25 of *Applications of Mathematics*. Springer-Verlag, New York, 2nd edition, 2006.

R. Gabasov and F. M. Kirillova. Method of optimal control. *J. Math. Sci. (NY)*, 7(5): 795–849, 1977.

A. M. Garsia, E. Rodemich, and H. Rumsey, Jr. A real variable lemma and the continuity of paths of some Gaussian processes. *Indiana Math. J.*, 20(6):565–578, 1970.

I. I. Gihman and A. V. Skorokhod. *Controlled Stochastic Processes*. Springer-Verlag, New York, 1979.

O. Hernández-Lerma and J. B. Lasserre. *Discrete-Time Markov Control Processes: Basic Optimality Criteria*, volume 30 of *Applications of Mathematics*. Springer-Verlag, New York, 1996.

Y. Hu, S.-E. A. Mohammed, and F. Yan. Discrete-time approximations of stochastic delay equations: the Milstein scheme. *Ann. Prob.*, 32(1A):265–314, 2004.

J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes*, volume 288 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin, 1987.

I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2nd edition, 1991.

V. B. Kolmanovskiĭ and L. E. Shaĭkhet. *Control of Systems with Aftereffect*, volume 157 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, Rhode Island, 1996.

N. V. Krylov. On the rate of convergence of finite-difference approximations for Bellman's equations with variable coefficients. *Probab. Theory Relat. Fields*, 117(1):1–16, 2000.

N. V. Krylov. Mean value theorems for stochastic integrals. *Ann. Probab.*, 29(1):385–410, 2001.

N. V. Krylov. The rate of convergence of finite-difference approximations for Bellman equations with Lipschitz coefficients. *Appl. Math. Optim.*, 52:365–399, 2005.

N. V. Krylov. *Controlled Diffusion Processes*, volume 14 of *Applications of Mathematics*. Springer-Verlag, New York, 1980.

N. V. Krylov. Approximating value functions for controlled degenerate diffusion processes by using piece-wise constant policies. *Electron. J. Probab.*, 4(Paper No. 2):1–19, 1999.

T. G. Kurtz and P. Protter. Weak convergence of stochastic integrals and differential equations. Lecture notes for the 1995 CIME School in Probability, October 2004.

T. G. Kurtz and P. Protter. Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Probab.*, 19(3):1035–1070, 1991.

H. J. Kushner. Numerical approximations for nonlinear stochastic systems with delays. *Stochastics*, 77(3):211–240, 2005.

H. J. Kushner. *Weak Convergence Methods and Singularly Perturbed Stochastic Control and Filtering Problems*, volume 3 of *Systems & Control: Foundations & Applications*. Birkhäuser, Boston, 1990.

H. J. Kushner and P. Dupuis. *Numerical Methods for Stochastic Control Problems in Continuous Time*, volume 24 of *Applications of Mathematics*. Springer-Verlag, New York, 2nd edition, 2001.

B. Larssen. Dynamic programming in stochastic control of systems with delay. *Stochastics Stochastics Rep.*, 74(3-4):651–673, 2002.

B. Larssen and N. H. Risebro. When are HJB-equations for control problems with stochastic delay equations finite dimensional? *Stochastic Anal. Appl.*, 21(3):643–671, 2003.

X. Mao. Numerical solutions of stochastic functional differential equations. *LMS J. Comput. Math.*, 6:141–161, 2003.

S.-E. A. Mohammed. *Stochastic Functional Differential Equations*. Pitman Publishing, London, 1984.

B. Øksendal and A. Sulem. A maximum principle for optimal control of stochastic systems with delay, with applications to finance. In J. Menaldi, E. Rofman, and A. Sulem, editors, *Optimal Control and Partial Differential Equations. In Honour of Professor Alain Bensoussan's 60th Birthday. Proceedings of the Conference, Paris, December 4, 2000*, pages 64–79. IOS Press, Amsterdam, 2001.

P. E. Protter. *Stochastic Integration and Differential Equations*, volume 21 of *Applications of Mathematics*. Springer-Verlag, Berlin, 2nd edition, 2003.

L. C. G. Rogers. Pathwise stochastic optimal control. *SIAM J. Control Optim.*, 46(3): 1116–1132, 2007.

L. Słomiński. Euler's approximations of solutions of SDEs with reflecting boundary. *Stochastic Processes Appl.*, 94:317–337, 2001.

D. W. Stroock and S. R. S. Varadhan. *Multidimensional Diffusion Processes*, volume 233 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin, 1979.

J. F. Traub and A. G. Werschulz. *Complexity and Information*. Lezioni Lincee. Cambridge University Press, Cambridge, 1998.

N. M. van Dijk. *Controlled Markov processes: time discretization*, volume 11 of *CWI Tracts*. Centrum voor Wiskunde en Informatica, Amsterdam, 1984.

F. Yan and S.-E. A. Mohammed. A stochastic calculus for systems with memory. *Stochastic Anal. Appl.*, 23:613–657, 2005.

J. Yong and X. Y. Zhou. *Stochastic Controls. Hamiltonian Systems and HJB Equations*, volume 43 of *Applications of Mathematics*. Springer-Verlag, New York, 1999.