KYBERNETIKA — VOLUME 32 (1996), NUMBER 2, PAGES 139-158

TIME-DISCRETIZATION FOR CONTROLLED MARKOV PROCESSES PART II: A Jump and Diffusion Application

NICO M. VAN DIJK AND ARIE HORDIJK

In a first Part I ([24]) a method of time-discretization was investigated in order to approximate continuous-time stochastic control problems over a finite time horizon.

This approximation was based on using recursive discrete-time dynamic programming. To this end, three conditions are to be fulfilled:

- *Smoothness* of the continuous-time functions
- *Consistency* or convergence of the discrete-time generators
- *Stability* or uniform boundedness of the discrete-time constructions.

In this Part II, these conditions will be verified for two practical applications:

- A *controlled infinite server queue*, as example of a controlled Markov jump process
- A *controlled cash-balance model*, as example of a controlled diffusion model.

For both applications it is shown and illustrated that the discrete-time constructions lead to a computational feasible scheme to approximate the optimal cost function as well as to construct an ε-optimal control.

CONTENTS

Introduction and summary

6. Controlled infinite server queue
   - 6.1 Model description and introduction
   - 6.2 Continuous- and discrete-time control objects
   - 6.3 Finite horizon cost function; approximation
   - 6.4 Finite horizon optimal cost function; approximation
   - 6.5 ε-optimal piecewise stationary control; construction
7. Controlled cash-balance model
   - 7.1 Model description and introduction
   - 7.2 Continuous- and discrete-time control objects
   - 7.3 Finite horizon cost function; approximation
   - 7.4 Finite horizon optimal cost function; approximation
   - 7.5 ε-optimal piecewise constant control; construction.

References

INTRODUCTION AND SUMMARY

This paper is a continuation of Part I as in [24], which dealt with the discrete-time approximation of finite horizon cost functions for controlled continuous-time Markov processes. Since the results of Part I will be applied, it is without further saying that in this second part we adopt all notation of Part I and that we will frequently refer to numbered expressions or statements from Part I. Sections 1–5 are in Part I, sections 6 and 7 are in Part II. For example, relation (3.2.7) or Theorem 5.2.1 can be found in Part I while relation (6.2.4) or Theorem 7.4.2 in Part II.

In Part I general approximation results were established provided a number of conditions were fulfilled. Roughly speaking, these conditions are:

- Sufficient smoothness of the continuous-time functions with respect to the time parameter (*smoothness*).
- A discrete-time approximation of the infinitesimal operators by means of appropriate one-step transition probabilities (*consistency*).
- Sufficient boundedness of discrete-time constructions (*stability*).

In this second Part these conditions will be verified for two applications and specifically chosen discretizations. Consequently, the discretization results obtained in Part I can be adopted. In addition, we will consider several modifications of these results that can be developed as well and we will put special attention to error bounds for feasible approximations. In both applications our main objective is the approximation of the optimal cost function and the construction of an ε-optimal control over a finite time horizon. More precisely, that is:

- To *approximate the optimal cost function* within some order of the length of the step-size h of the discrete-time parameter.
- To provide a way of *constructing ε-optimal controls* for a given ε by using discrete-time dynamic programming.

In order to obtain these results, we first need to guarantee the existence of a unique and sufficiently smooth solution of the continuous-time optimality (Bellman) equation. In neither of the two applications this is a standard result.

The first application, presented in Section 6, concerns a *controlled infinite server queue*. The underlying process of this model fits in the framework of controlled Markov jump processes. Moreover, as special complication the jump rates are unbounded. Approximation results given in Section 7, Chapter II of Van Dijk [20] for controlled Markov jump processes with bounded jump rates can therefore not be applied. Nevertheless, a constructive approximation of the optimal cost function will be provided.

The second application, given in Section 7, deals with a *controlled cash-balance (or investment) model*. The underlying process, in this case, is a controlled diffusion process. A special complication here is that the decision set is assumed to be discrete, so that natural controls are unlikely to satisfy standard Lipschitz conditions with respect to the state variable. Nevertheless, under the assumption of a sufficiently smooth solution of the Bellman equation, a discrete-time computational approximation of the optimal cost function will be established. As our primary purpose is the approximation of the optimal cost function and the construction of an implementable ε-optimal control, we will restrict our attention to results for the class of piecewise stationary Markov controls for the first application and of piecewise constant and so-called almost Markov controls for the second. Similar results that can be obtained for arbitrary controls are left to the reader. As the details for the verification of the necessary conditions can be rather technical and tedious, we will often restrict the proofs to essential steps and refer to Van Dijk [20] for these remaining technical details.

6. CONTROLLED INFINITE SERVER QUEUE

6.1. Model description and introduction

Consider a service facility where customers arrive according to a Poisson process with parameter λ. The number of servers is controlled continuously. A customer can only be served by one server at a time, and the number of servers may never exceed the number of customers present. Each customer demands an amount of service according to an exponential distribution with parameter ν. Consequently, for a fixed control, an informal description of the underlying process, with the number of customers present as state, can be given as follows:

Given that at time point t the actual number of customers is i and that during $[t,\, t+\Delta t]$ the number of servers is j, where $j \leq i$, then with probability

$$\left.\begin{array}{ll} \lambda\,\Delta t + O_i([\Delta t]^2) & \text{there is only one arrival} \\ [j\nu]\;\Delta t + O_i([\Delta t]^2) & \text{there is only one departure} \\ 1 - [\lambda + j\nu]\;\Delta t + O_i([\Delta t]^2) & \text{there is no arrival nor departure} \end{array}\right\} \qquad (6.1.1)$$

during $[t,\, t+\Delta t]$ and where (as can be shown similarly to p. 42 of Van Dijk [20]) for some positive constant C and all $i \in \mathbb{N}$, $\Delta t > 0$:

$$|O_i([\Delta t]^2)| \leq C[i]^2\,[\Delta t]^2. \qquad (6.1.2)$$

Costs are taken into account by means of a cost-rate function $\boldsymbol{L}$ depending on the actual number of customers i and active servers j. We assume that for certain constants C_0, C_1, C_2, p_1, and p_2 and all i, $j \in \mathbb{N}$:

$$\boldsymbol{L}(i,j) \leq C_0 + C_1[i]^{p_1} + C_2[j]^{p_2}. \qquad (6.1.3)$$

Hence, with $j \leq i$ the cost-rate function can be bounded by a polynomial only depending on i and of order p, where:

$$p = \max(p_1, p_2). \qquad (6.1.4)$$

Note that this condition of $j \leq i$ is natural if there are no switching costs involved when changing the number of servers, and for a cost rate function which is nondecreasing in the number of servers.

To the best of our knowledge, this particular model has not been dealt with in the literature. However, in view of applying time-discretization, there is a relation with the analysis for a controlled $M|M|1$-queue given by Hordijk and Van der Duyn Schouten [10] and [11] as well as that for a controlled $M|G|1$-queue given by Mitchell [14] and Doshi [5]. Let us briefly elaborate on the results of these references.

These references also deal with the continuous control of the service capacity based on the actual workload. To this end, time-discretization is applied in order to derive structural results for optimal policies in the continuous-time model. Particularly, by transferring structural results for discrete-time models by means of weak convergence arguments, Hordijk and Van der Duyn Schouten [10] and [11] have been able to prove the optimality of *bang-bang policies* for the continuous-time model. The application in this section does not allow a controllable service capacity. But in contrast:

- The service capacity is not assumed to be bounded.
- Also the computational aspect of the discretization-method is considered by explicit error bounds for the approximation of the optimal value function and also by constructing ε-optimal controls.

More precisely, the objectives in this section are:

(i) To approximate the finite horizon optimal cost function for some specified accuracy (Section 6.4).

(ii) To show how to construct ε-optimal controls for some given ε by using discrete-time dynamic programming (Section 6.5).

Since the cost rate is bounded by a polynomial the results will be given in a weighted supremum norm with an appropriate polynomial as weighting function. As an implication of (6.1.3), this polynomial will be of order $p+2$.

6.2. Continuous- and discrete-time control objects

6.2.1. *Continuous-time*

Define

$$B^{\mu_p} = \{f : I\!N \to I\!R \mid |f(i)| \le K_p(1+i)^p \text{ for all } i \in I\!N \text{ and some constant } K_p\}. \tag{6.2.1}$$

Then, associated with the informal description in Section 6.1, we will consider the control object $(S,\ \Gamma,\ \Delta,\ \mu,\ D_A,\ \{\boldsymbol{A}^\delta \mid \delta \in \Delta\},\ \boldsymbol{L})$ as defined in Section 2.1, specified by

$$\left.\begin{array}{l} S = I\!N,\ \Gamma = I\!N;\ \Delta = \{\delta \in I\!N \to I\!N \mid \delta(i) \le i,\ i \in I\!N\} \\ \text{With } p = \max(p_1, p_2): \\ \mu_p(i) = (1+i)^{p+2} \\ D_A = B^{\mu_p} \end{array}\right\} \tag{6.2.2}$$

and

$$\boldsymbol{A}^\delta f(i) = \lambda[f(i+1) - f(i)] + \delta(i)\,\nu[f(i-1) - f(i)], \quad i \in I\!N, \tag{6.2.3}$$

and $\boldsymbol{L}$ as in Section 6.1.

The particular choices of the bounding function μ, the domain D_A and the infinitesimal operators $\boldsymbol{A}^\delta$ will be justified by Lemma 6.2.1 given below.

6.2.2. Discrete-time h-control object

With $h > 0$ denoting the step size of the discrete-time parameter, we will consider the h-control object $(S, \Gamma, \Delta, \{\boldsymbol{P}_h^\delta | \delta \in \Delta\}, \boldsymbol{L})$, as defined in Section 3.1, with $\{\boldsymbol{P}_h^\delta | \delta \in \Delta\}$ defined by:

$$\boldsymbol{P}_h^\delta(i,j) = \begin{cases} 1 - [h(\lambda + \nu\,\delta(i)) \wedge 1], & j = i, \\ [h(\lambda + \nu\,\delta(i)) \wedge 1]\,\frac{\lambda}{[\lambda + \nu\,\delta(i)]}, & j = i+1, \\ [h(\lambda + \nu\,\delta(i)) \wedge 1]\,\frac{\nu\,\delta(i)}{[\lambda + \nu\,\delta(i)]}, & j = i-1. \end{cases} \tag{6.2.4}$$

The particular choice of $\boldsymbol{P}_h^\delta$ results from the fact that h times the departure rate $\nu\,\delta(i)$ is not necessarily bounded by 1. Therefore, we cannot take one-step transition probabilities as products of h and jump intensities as in Hordijk and Van der Duyn Schouten [11] and Van Dijk [20].

6.2.3. Consistency and stability

In order to obtain approximation results, this subsection presents the basic inequalities as direct consequences of the above defined discretization. In view of the literature on numerical analysis, the inequalities (6.2.6) and (6.2.7) may be referred to as the consistency and the stability relation, respectively.

First recall the notation $\boldsymbol{A}_h^\delta$ for the one-step operator given by

$$\boldsymbol{A}_h^\delta = [\boldsymbol{P}_h^\delta - I]\,h^{-1}$$

with $\boldsymbol{P}_h^\delta$ defined by (6.2.4). Further, for any $m \in \mathbb{N}$ let the polynomial bounding function $\mu_m : \mathbb{N} \to \mathbb{R}$ be defined by:

$$\mu_m(i) = (1+i)^m, \quad i \in \mathbb{N}. \tag{6.2.5}$$

The following lemma will guarantee the necessary *consistency* and *stability* of the discrete-time construction.

Lemma 6.2.1. For constants C and K_Δ and all $f \in B^{\mu_p}$:

$$\sup_{\delta \in \Delta} \|(\boldsymbol{A}_h^\delta - \boldsymbol{A}^\delta)\,f\|_{\mu_{p+2}} \le h\,C\|f\|_{\mu_p} \tag{6.2.6}$$

$$\sup_{\delta \in \Delta} \left\| \sum_j \mu_{p+2}(j)\,\boldsymbol{P}_h^\delta(\cdot;j) \right\|_{\mu_{p+2}} \le (1 + h\,K_\Delta). \tag{6.2.7}$$

Proof. See the Lemmas 5.2.8 and 5.2.9, Chapter I of Van Dijk [20].

The proof of (6.2.6) and our use of a polynomial bounding function of order $p+2$ instead of an order p, as one might expect, follows from the fact that for some constant C and all $\delta \in \Delta$:

$$\{[h(\lambda+\nu\,\delta(i)) \wedge 1] - h(\lambda+\nu\delta(i))\} \;/\; h(1+i)^2 \leq hC, \qquad i \in \mathbb{N}. \tag{6.2.8}$$

The proof of (6.2.7) is a straightforward result from the monotonicity of μ_{p+2} and the boundedness of the (in fact constant) arrival rate λ. □

6.3. Finite horizon cost function: Approximation

Let $h > 0$ and consider a piecewise constant control $\pi \in \Pi$ satisfying $\pi(t) = \pi(n\,h), n = \lfloor t\,h^{-1} \rfloor$. The following lemma shows that π is admissible and μ-bounding such as defined in Section 2.2.

Lemma 6.3.1. There exists a unique family of transition expectation operators $\{\boldsymbol{T}^{\pi}_{s,t} \mid s,\, t \leq Z\}$ corresponding to the control object from Section 6.2.1 and satisfying (2.2.4) and (2.2.5). For any $m \in \mathbb{N}$, there is some constant $M(m)$ such that for all $s,\, t \leq Z$:

$$\left\|\boldsymbol{T}^{\pi}_{s,t}\, f\right\|_{\mu_m} \leq M(m)\, \|f\|_{\mu_m}. \tag{6.3.1}$$

Furthermore, for all $s,\, t \leq Z,\ g \in B^{\mu_p+2},\ f \in B^{\mu_p};\ s+\Delta s \leq \lfloor s\,h^{-1} \rfloor\, h + h;\ m = p, p+1, p+2;\ n \in \mathbb{N};\ \delta \in \Delta$ and some constant C:

$$\|\boldsymbol{T}^{\pi}_{s,t}\, g\|_{\mu_m} \leq C\|g\|_{\mu_m}. \tag{6.3.2}$$

$$\|\boldsymbol{A}^{\delta}\, g\|_{\mu_{m+1}} \leq C\|g\|_{\mu_m}. \tag{6.3.3}$$

$$\|[\boldsymbol{T}^{\pi}_{s,\,s+\Delta s} - I]\, f\|_{\mu_{p+2}} \leq \Delta s\, C\|f\|_{\mu_p}. \tag{6.3.4}$$

$$\|[(\boldsymbol{T}^{\pi}_{nh,\,nh+h} - I]\, h^{-1} - \boldsymbol{A}^{\pi(nh)})\, f\|_{\mu_{p+2}} \leq h\, C\|f\|_{\mu_p}. \tag{6.3.5}$$

Proof. By standard construction of the minimal jump process, the quadratic order in i given in (6.1.2) and technical steps similar to Lemma 5.2.7, Chapter 1 of Van Dijk [20]. □

Note that the stability relation (6.2.7) implies (3.2.2), hence we have:

$$\pi^h = (\pi(0),\, \pi(h), \ldots, \pi(\ell\, h)) \in \Pi^h(AB).$$

Recall $\boldsymbol{V}^{\pi}_t$ and $\boldsymbol{V}^h_j$ as defined by (2.2.6) and (3.2.6), where the justification of these definitions will be given below. Then, the following approximation theorem is an application of Theorem 5.2.1.

Theorem 6.3.2. For some constant C and all $n \le \ell$:

$$\|\boldsymbol{V}_n^h - \boldsymbol{V}_{nh}^\pi\|_{\mu_{p+2}} \le h\,C. \tag{6.3.6}$$

Proof. We will apply Theorem 5.2.1. First of all, we will verify Assumptions 2.2.3 and 3.2.3, which guarantee the finiteness of $\boldsymbol{V}_t^\pi$ and $\boldsymbol{V}_j^h$. According to the polynomial boundedness of $\boldsymbol{L}$,

$$\sup_{\delta\in\Delta} \|\boldsymbol{L}^\delta\|_{\mu_p} \le C. \tag{6.3.7}$$

Relation (6.3.7) together with (6.3.2) implies the μ_p- (and hence μ_{p+2}-) boundedness of $\{\boldsymbol{T}_{t,s}^\pi\, \boldsymbol{L}^{\pi(s)} \,|\, s \le Z\}$ as well as of $\{\boldsymbol{V}_t^\pi \,|\, t \le Z\}$. Further, for $t \le s \le Z$ and with $s,\ s+\Delta s \in [nh,\, nh+h]$, the fact that $\pi(s+\Delta s) = \pi(s)$ and the relations (6.3.2) and (6.3.4) yield:

$$\left\|\boldsymbol{T}_{t,s}^\pi\, \boldsymbol{L}^{\pi(s+\Delta s)} - \boldsymbol{T}_{t,s}^\pi\, \boldsymbol{L}^{\pi(s)}\right\|_{\mu_{p+2}} \le \left\|\boldsymbol{T}_{t,s+\Delta s}^\pi \left(\boldsymbol{L}^{\pi(s+\Delta s)} - \boldsymbol{L}^{\pi(s)}\right)\right\|_{\mu_{p+2}}$$

$$+\|\boldsymbol{T}_{t,s}^\pi \left(\boldsymbol{T}_{s,\,s+\Delta s}^\pi - I\right)\, \boldsymbol{L}^{\pi(s)}\|_{\mu_{p+2}} \le \Delta s\, C. \tag{6.3.8}$$

This implies that $\boldsymbol{T}_{t,s}^\pi\, \boldsymbol{L}^{\pi(s)}$ is piecewise μ_{p+2}-continuous and thus integrable. Consequently, Assumption 2.2.3 is verified with $D_A = B^{\mu_p}$.

Obviously, also Assumption 3.2.3 is guaranteed by (6.3.7).

Next, expression (2.2.8) for $R_{nh}^\pi(\boldsymbol{V}, h)$, relation (6.3.5), together with the μ_p-boundedness of $\{\boldsymbol{V}_t^\pi \,|\, t \le Z\}$ yield:

$$\begin{aligned} &\left\|R_{nh}^\pi(\boldsymbol{V}, h)\, h^{-1}\right\|_{\mu_{p+2}} \\ \le\ &\left\|\left[\int_{nh}^{nh+h} \boldsymbol{T}_{nh,\,s}^\pi\, \boldsymbol{L}^{\pi(s)} \mathrm{d}s - h\boldsymbol{L}^{\pi(s)}\right]\, h^{-1}\right\|_{\mu_{p+2}} \\ +\ &\left\|\left([\boldsymbol{T}_{nh,\,nh+h}^\pi - I]\, h^{-1} - \boldsymbol{A}^{\pi(nh)}\right)\, \boldsymbol{V}_{nh}^\pi\right\|_{\mu_{p+2}} \le h\,C. \end{aligned} \tag{6.3.9}$$

Finally, the proof is completed by applying Theorem 5.2.1, using again relation (6.3.5) and the μ_p-boundedness of $\{\boldsymbol{V}_t^\pi \,|\, t \le z\}$. □

6.4. Finite horizon optimal cost function: Approximation

Before we can present the main approximation result on the discrete-time approximation of the continuous-time optimal cost function $\boldsymbol{\Phi}_t$, we first need to justify the existence and sufficient boundedness of $\boldsymbol{\Phi}_t$.

Lemma 6.4.1. There exist a unique family $\{\boldsymbol{\Phi}_t \,|\, t \le Z\}$ satisfying:

(i) The continuous-time optimality equation (2.3.2).

(ii) The family $\{\boldsymbol{\Phi}_t \,|\, t \le Z\}$ is μ_p-bounded.

Proof. The proof is technical and can be found in Van Dijk [21]. It is an extension of existence results in Pliska [15] and Yushkevich [25] for the pure bounded and the partially bounded nonnegative case. □

Note that in the proof of the next theorem we use the same notation C for possibly different bounding constants in the various relations we derive.

Theorem 6.4.2. For some constant C and all $n \le \ell$:

$$\left\| \boldsymbol{\Phi}_n^h - \boldsymbol{\Phi}_{nh} \right\| \le h\,C. \tag{6.4.1}$$

Proof. We will apply Theorem 5.3.1. First of all, we need to verify Assumptions 2.3.1 and 2.3.2 for the continuous-time model as well as Assumption 3.3.1 for the discrete-time model. Assumption 2.3.1 directly follows from Lemma 6.4.1. From the μ_p-boundedness of $\boldsymbol{L}$ it follows from (6.3.3) that for all $g,\, g_1,\, g_2 \in B^{\mu_p},\; m = p,\, p+1$:

$$||\boldsymbol{J}(g)||_{\mu_{m+1}} \le C(1 + ||g||_{\mu_m}), \tag{6.4.2}$$

$$||\boldsymbol{J}(g_1) - \boldsymbol{J}(g_2)||_{\mu_{m+1}} \le C||g_1 - g_2||_{\mu_m}. \tag{6.4.3}$$

Since also: $||g||_{\mu_{m+1}} \le ||g||_{\mu_m}$, Lemma 6.4.1 together with (6.4.2) implies the μ_{p+2}-boundedness of $\{\boldsymbol{J}(\boldsymbol{\Phi}_t) \,|\, t \le Z\}$ which together with Lemma 6.4.1 guarantees Assumption 2.3.2 with $D_A = B^{\mu_p}$ and $\mu = \mu_{p+2}$. Similarly, by using (6.2.7) one can easily verify Assumption 3.3.1 with $f \in B^{\mu_p}$.

Next, from inequality (6.4.3) with $m = p+1$ and (6.4.2) with $m = p$, the continuous-time optimality equation (2.3.2) and the μ_p-boundedness of $\{\boldsymbol{\Phi}_t \,|\, t \le Z\}$ we obtain:

$$\begin{aligned} ||\boldsymbol{J}(\boldsymbol{\Phi}_{t+\Delta t}) - \boldsymbol{J}(\boldsymbol{\Phi}_t)||_{\mu_{p+2}} &\le ||\boldsymbol{\Phi}_{t+\Delta t} - \boldsymbol{\Phi}_t||_{\mu_{p+1}} C \\ &= \left\| \int_t^{t+\Delta t} \boldsymbol{J}(\boldsymbol{\Phi}_s)\,\mathrm{d}s \right\|_{\mu_{p+1}} C \le \Delta t\, C. \end{aligned} \tag{6.4.4}$$

Expression (2.3.3) for $R_{nh}(\boldsymbol{\Phi}, h)$ and (6.4.4) yield:

$$||R_{nh}(\boldsymbol{\Phi}, h)||_{\mu_{p+2}}\, h^{-1} \le h\,C. \tag{6.4.5}$$

Finally, the proof is completed by applying Theorem 5.3.1, using Lemma 6.2.1, together with the μ_p-boundedness of $\{\boldsymbol{\Phi}_t \,|\, t \le Z\}$. □

6.5. ε-optimal piecewise stationary controls; construction

In this section we show that the results of Sections 6.3 and 6.4 enable us to construct a piecewise stationary control which is ε-optimal for the continuous-time model, where ε may be chosen arbitrarily small. We must first establish a discrete-time control, for instance by using discrete-time dynamic programming, which is optimal or γ-optimal for an h-discrete time model. Next, this control will be implemented in the continuous-time model as a piecewise stationary control π, which is constant on the intervals $[nh,\, nh+h)$. Finally, by combining the approximation Theorems 6.3.2 and 6.4.2 we obtain a bound for the difference $||\boldsymbol{V}^{\pi} - \boldsymbol{\Phi}||$. More precisely:

Theorem 6.5.1. Let $\pi^h = (\delta(0),\, \delta(1), \ldots, \delta(\ell)) \in \Pi^h$ such that for some $\gamma \geq 0$:

$$||\boldsymbol{V}_n^h - \boldsymbol{\Phi}_n^h||_{\mu_{p+2}} \leq \gamma, \quad n \leq \ell. \tag{6.5.1}$$

Then with $\pi \in \Pi$ defined by: $\pi(t) = \delta(n)$ for $t \in [nh,\, nh+n),\ t \leq Z$:

$$||\boldsymbol{V}_{nh}^{\pi} - \boldsymbol{\Phi}_{nh}||_{\mu_{p+2}} \leq \gamma + h\,C, \quad n \leq \ell. \tag{6.5.2}$$

Proof. Since

$$||\boldsymbol{V}_{nh}^{\pi} - \boldsymbol{\Phi}_{nh}||_{\mu_{p+2}} \leq ||\boldsymbol{V}_{nh}^{\pi} - \boldsymbol{V}_n^h||_{\mu_{p+2}} + ||\boldsymbol{V}_n^h - \boldsymbol{\Phi}_n^h||_{\mu_{p+2}} + ||\boldsymbol{\Phi}_n^h - \boldsymbol{\Phi}_{nh}||_{\mu_{p+2}}, \tag{6.5.3}$$

the proof follows directly from Theorems 6.3.2 and 6.4.2 and relation (6.5.1). □

Remark 6.5.2. The above theorem shows that for any given $\varepsilon > 0$ an ε-optimal control is obtained by finding a discrete-time control which is γ-optimal for an h-discrete-time model such that

$$\gamma + h\,C \leq \varepsilon, \tag{6.5.4}$$

where C is some constant, following from (6.3.6) and (6.4.1), which does not depend on h. Consequently, in order to guarantee (6.5.4) an upper bound of C must be known. Such a bound can be obtained theoretically from the several inequalities used in proving (6.3.6) and (6.4.1), such as inequalities (6.3.2), (6.3.3), (6.3.4), (6.3.5) and (6.3.7). This approach, however, would be cumbersome and very inaccurate due to the many steps involved. An easier and, most likely, much more accurate way for obtaining an upper bound of C is to deduct C from numerically obtained values $h\,C$ for a number of different values of h. Once an upper bound of C is established, the finding of a sufficiently small h and γ, in order to satisfy (6.5.4), still contains the difficulty that according to (6.5.1) these are interrelated. It is well-known from dynamic programming that $\boldsymbol{V}_n^h = \boldsymbol{\Phi}_n^h$, thus $\gamma = 0$, if

$$\boldsymbol{\Phi}_j^h = \inf_{\delta \in \Delta} [h\,\boldsymbol{L}^{\delta} + \boldsymbol{T}_h^{\delta}\,\boldsymbol{\Phi}_{j+1}^{\delta}] = [h\,\boldsymbol{L}^{\delta(j)} + \boldsymbol{T}_h^{\delta(j)}\,\boldsymbol{\Phi}_{j+1}^h], \quad j < \ell. \tag{6.5.5}$$

This dynamic programming equation, however, requires the exact computation of the infima, whereas numerical computations will involve inaccuracies. For this reason as well as for its own right, Lemma 6.5.3 given below, as a direct application of the approximation Lemma 4.1, may be helpful in finding a more convenient discrete-time control satisfying (6.5.1).

Lemma 6.5.3. Suppose that for $\eta > 0$ and all $j < \ell$:

$$\left\| (h\,\boldsymbol{L}^{\delta(j)} + \boldsymbol{T}_h^{\delta(j)}\,\boldsymbol{V}_{j+1}^h) - \inf_{\delta\in\Delta}(h\,\boldsymbol{L}^{\delta} + \boldsymbol{T}_h^{\delta}\,\boldsymbol{V}_{j+1}^h) \right\|_{\mu_{p+2}} \leq \eta\, h. \tag{6.5.6}$$

Then, with K_Δ given by (6.2.7), relation (6.5.1) holds with $\gamma = \eta \exp(Z\,K_\Delta)$.

P r o o f. The systems (3.2.6) and (3.3.2) guarantee system 4.1 specified by:

$$\left.\begin{aligned} &\boldsymbol{U}_{jh} = \boldsymbol{\Phi}_{jh};\ \boldsymbol{U}_j^h = \boldsymbol{V}_j^h;\ \boldsymbol{B} = B^{\mu_{p+2}};\\ &\boldsymbol{C}_{jh}(f) = \inf_{\delta\in\Delta}[h\,\boldsymbol{L}^{\delta} + \boldsymbol{T}_h^{\delta}(f)];\\ &\boldsymbol{C}_j^h(f) = [h\,\boldsymbol{L}^{\delta(j)} + \boldsymbol{T}_h^{\delta(j)}(f)]. \end{aligned}\right\} \tag{6.5.7}$$

Consequently, relation (6.5.6) guarantees (4.2). Furthermore, similarly to relation (5.3.7), it follows that relation (6.2.7) implies (4.3) with $K = K_\Delta$. Finally, $\boldsymbol{U}_{\ell h} = \boldsymbol{\Phi}_\ell^h = \boldsymbol{U}_\ell^h = \boldsymbol{V}_\ell^h = \overline{0}$. Application of Lemma 4.1 completes the proof. □

Remark 6.5.4. Combination of Theorem 6.5.1 and Lemma 6.5.3 together with the recursive system (3.2.6) for calculating $\boldsymbol{V}_j^h$, yield the following algorithm for computing an ε-optimal control:

$$\left\{\begin{array}{ll} \underline{\text{ALGORITHM}}: & \\ \underline{\text{Start}}: & \boldsymbol{V}_\ell^h = \overline{0}. \\ & \text{For } j = \ell - 1 \text{ down to 0 do:} \\ \underline{\text{Step 1}}: & \text{Determine } \delta(j) \text{ such that (6.5.6) holds.} \\ \underline{\text{Step 2}}: & \text{Compute } \boldsymbol{V}_j^h \text{ according to (3.2.6).} \end{array}\right\} \tag{6.5.8}$$

Note that this algorithm provides functions $\boldsymbol{V}_n^h$ as approximations of $\boldsymbol{\Phi}_{nh}$ within an accuracy of order $O(h)$ without explicitly knowing $\boldsymbol{\Phi}$ itself.

Remark 6.5.5. In contrast with results in Pliska [14] and Yushkevich [25], note that Lemma 6.5.3 enables one to construct ε-optimal controls with

(i) simple one-step transition probabilities as per (6.2.4)

(ii) prespecified accuracy-value ε.

7. CONTROLLED CASH–BALANCE MODEL

7.1. Model description and introduction

Consider a controlled stochastic equation of the form

$$X_{t+\Delta t} = x + \gamma_1\,\Delta t + \gamma_2\,W_{\Delta t}, \quad t \geq 0 \tag{7.1.1}$$

to indicate the state of a process $\{X_t\,|t \geq 0\}$ at time $t+\Delta t$ given that the system is in state x at time t and that during $[t,\, t+\Delta t]$ the process is continuously controlled by a decision pair (γ_1, γ_2) from an available set of decision pairs O. Here $W_{\Delta t}$ denotes a Wiener increment and the decision set O is assumed to be finite. This equation, which for $\Delta t \to 0$ can be seen as a stochastic differential equation, applies to each of the following economic models.

(i) The value (profit) of an investment of fixed amount is continuously controlled by allocating an investment opportunity (γ_1, γ_2), where γ_1 denotes the rate of return and $[\gamma_2]^2$ indicates a value of risk given by its variance per unit of time. Costs are involved expressed by a cost-rate function $\boldsymbol{L}$ depending on the value of the investment (negative reward rate), the rate of return and the value of risk.

(ii) The cash-balance of a bank is continuously controlled by allocating a transfer rate γ_1 (positive, zero or negative). Fluctuations, due to deposits and withdrawals, which may strongly vary in frequency and size per time unit, are modelled by a Wiener process with variance $[\gamma_2]^2$ per time unit. Here, γ_2 may have a fixed, and thus uncontrolled, value. Costs are taken into account expressed by a cost-rate function $\boldsymbol{L}$ depending on the actual cash-balance X_t according to a holding rate if $X_t > 0$ and a shortage rate if $X_t < 0$, and on the transfer rate γ_1 (buy and sell rates).

For either of the descriptions the following assumption on the cost-rate function $\boldsymbol{L}$ is made:

$$\left.\begin{array}{l}\boldsymbol{L}(x, \gamma_1, \gamma_2) \text{ is three times continuously differentiable in } x \\ \text{for any fixed } (\gamma_1, \gamma_2) \in \mathrm{O}, \text{ and for some constant } p \in I\!N, \\ K_L > 0 \text{ and all } (x, \gamma_1, \gamma_2): \\ \left|\frac{\partial^k}{\partial x^k}\boldsymbol{L}(x, \gamma_1, \gamma_2)\right| \leq K_L(1 + |x|^p), \quad k = 0, 1, 2, 3.\end{array}\right\} \tag{7.1.2}$$

We note that although the system behaviour, or more precisely the drift and diffusion coefficient γ_1 and γ_2 respectively, does not depend on the actual state variable explicitly, in view of a state dependent cost-rate function the total costs from a point of time onward do depend on the actual state, and so do natural controls, such as (ε-) optimal controls.

Investment and particularly cash-balance models have been studied extensively in the literature (cf. Pliska [16], Constantinides [3], Constantinides and Richard [4], Harrison and Taksar [8]). On one hand the models presented in the literature are sometimes more complex, such as by investing a whole fund so that the return rate and variance of risk will be linear in the actual fund value (Pliska [16]) or by dealing with transactions instead of a transfer rate so that actual control will be impulsive (Constantinides [3], Constantinides and Richard [4], Harrison and Taksar [8]). On the other hand, the above references concentrate on the stationary situation and a specific cost structure which simplify the calculation of an optimal control. Furthermore, general results for controlled diffusion processes as given by Fleming and Rishel [6], Puterman [18] and Krylov [12] do not directly apply. In this respect we note that in investment model (i) the diffusion coefficient γ_2 is controllable and that, in view of a discrete decision set, it is unnatural to impose too strong Lipschitz conditions upon controls. Therefore, at least from a theoretical point of view, the model under consideration seems to be of interest for investigation. More importantly, the approximation results that will be presented give a first indication of an obtainable accuracy that can be obtained. Further, they yield the construction of ε-optimal controls for any desired precision ε.

As in Section 6, the main objectives in this section are:

(i) To approximate the finite horizon cost function (Section 7.3).

(ii) To establish a construction of ε-optimal controls for a given ε by means of time-discretization (Section 7.4). Two discretizations are investigated: one which corresponds to discretizations as used in the literature (cf. Gihman and Skorohod [7], Christopheit [2], Bensoussan and Robin [1]), Hausmann [9] and Kushner [13], and one which is computationally more efficient.

Since the cost rate function is bounded by a polynomial, the results will be given in supremum norms weighted by a polynomial. Since the coefficients γ_1 and γ_2 are bounded themselves, the order of that polynomial can be taken equal to that of the cost-rate function. In Section 7.2 we formally introduce the continuous- and discrete-time structure by presenting the corresponding control objects as given in sections 2.1 and 3.1. Further, the essential inequalities for the approximation analysis are given.

7.2. Continuous- and discrete-time control objects

7.2.1. Continuous-time

Define

$$\left.\begin{aligned} C^{3;p} = \Big\{ f : \mathbb{R} \to \mathbb{R} \,\Big|\, & \tfrac{\mathrm{d}^k}{\mathrm{d}x^k} f(x) \text{ exist and are continuous for } k = 1,2,3; \\ & \text{and } \tfrac{\mathrm{d}^k}{\mathrm{d}x^k} f(x) \leq K_f(1+|x|^p), \text{ for all } x \in \mathbb{R},\ k = 0,1,2,3 \\ & \text{and some constant } K_f \Big\} \text{ for any } p \in \mathbb{N}. \end{aligned}\right\} \tag{7.2.1}$$

Then, associated with the informal descriptions given in Section 7.1, we will consider the control objects $(S,\, \Gamma,\, \Delta,\, \mu,\, D_A,\, \{\boldsymbol{A}^\delta | \delta \in \Delta\},\, \boldsymbol{L})$, as defined in Section 2.1:

$$\left.\begin{aligned} & S = \mathbb{R};\ \Gamma = \mathrm{O} \text{ where O is a finite subset of } \mathbb{R}^2 \\ & \Delta = \{\delta : \mathbb{R} \to \mathrm{O} \mid \delta \text{ piecewise continuous}\} \\ & \boldsymbol{L}(\cdot, \gamma_1, \gamma_2) \in C^{3;p} \text{ for any } (\gamma_1, \gamma_2) \in \mathrm{O} \text{ and some } p \in \mathbb{N}, \end{aligned}\right\} \tag{7.2.2}$$

$$\tag{7.2.3}$$

$$\left.\begin{aligned} & \mu(x) = (1+|x|^p) \\ & D_A = C^{3;p}, \end{aligned}\right\} \tag{7.2.4}$$

and

$$\left.\begin{aligned} & \boldsymbol{A}^\delta f(x) = [\gamma_1] \tfrac{\mathrm{d}}{\mathrm{d}x} f(x) + \tfrac{1}{2} [\gamma_2]^2 \tfrac{\mathrm{d}^2}{\mathrm{d}x^2} f(x) \\ & \text{for } \delta(x) = (\gamma_1, \gamma_2),\ x \in \mathbb{R} \text{ and } f \in C^{3;p}. \end{aligned}\right\}. \tag{7.2.5}$$

The choices of S and Γ and the differentiability condition on $\boldsymbol{L}$ are direct consequences of the description and condition (7.1.2) of Section 7.1. The particular choices of the bounding function μ, the domain D_A and the infinitesimal operators $\boldsymbol{A}^\delta$ will be justified by Lemma 7.2.1 and Lemma 7.3.1 below. The choice of Δ guarantees a sufficiently wide class of controls π associated with controlled stochastic differential equations.

7.2.2. *Discrete-time h-control object*

With $h > 0$ denoting the step size in the discrete-time parameter, we will focus on a discrete-time h-control object $(S, \Gamma, \Delta, \mu, h, \{\boldsymbol{P}_h^\delta | \delta \in \Delta\}, \boldsymbol{L})$, as defined in Section 3.1, where $\{\boldsymbol{P}_h^\delta | \delta \in \Delta\}$ is specified by:

$$\boldsymbol{P}_h^\delta(x; \{y\}) = \begin{cases} \frac{1}{2} & \text{for } y = x + \gamma_1 h + \gamma_2 \sqrt{h} \\ \frac{1}{2} & \text{for } y = x + \gamma_1 h - \gamma_2 \sqrt{h} \end{cases}, \text{ where } (\gamma_1, \gamma_2) = \delta(x). \tag{7.2.6}$$

This random walk type discretization is the same as the one given in Kushner [13] if $\gamma_1 = 0$. The corresponding difference method will be computationally slow. However, as opposed to other stochastic discretizations considered in the literature (cf. Gihman and Skorohod [7], Christopeit [2], Bensoussan and Robin [1]), at least the approximations can be actually computed. Compared with the discretizations used by Kushner [13] and Haussmann [9], the simple structure of the discretization (7.2.6) greatly facilitates the form of the discrete-time dynamic programming equation.

7.2.3. *Consistency and stability*

As in subsection 6.2.3, we first present basic inequalities (7.2.8) and (7.2.9), which will guarantee consistency and stability of the approximate scheme. Again, recall the notation $\boldsymbol{A}_h^\delta$ for the one step generator given by (3.1.1) with $\boldsymbol{P}_h^\delta$ defined by (7.2.6). Further, throughout this section the polynomial bounding function $\mu_p : \mathbb{R} \to \mathbb{R}$ is defined by:

$$\mu_p(x) = 1 + |x|^p, \quad x \in \mathbb{R}. \tag{7.2.7}$$

Lemma 7.2.1. For constants C and K_Δ and all $f \in C^{3;p}$:

$$\sup_{\delta \in \Delta} \|(\boldsymbol{A}_h^\delta - \boldsymbol{A}^\delta) f\|_{\mu_p} \leq \sqrt{h}\, C\, K_f. \tag{7.2.8}$$

$$\sup_{\delta \in \Delta} \left\| \int \mu_p(y)\, \boldsymbol{P}_h^\delta(\cdot; \mathrm{d}y) \right\|_{\mu_p} \leq (1 + h\, K_\Delta). \tag{7.2.9}$$

Proof. Similarly to Lemmas 5.3.12 and 5.3.13, Chapter I of Van Dijk [20].

Basically, the proof of (7.2.7) follows from Taylor expansion together with the boundedness of the drift coefficient $\gamma_1 = \delta_1(x)$ and the diffusion coefficient $\gamma_2 = \delta_2(x)$, uniformly in $x \in \mathbb{R}$ and $\delta \in \Delta$. The proof of (7.2.7) is a straightforward result from the boundedness of the coefficients. We note that in the above reference a polynomial bounding function of order $p+3$ instead of p is needed, since there the coefficients satisfy a growth instead of boundedness condition. □

7.3. Finite horizon cost function: Approximation

In this section we will slightly deviate from the formulation of a finite horizon cost function for a given admissible and μ-bounding Markov control as defined in Section 2.2. The main reason for doing so is that conditions have to be imposed upon a Markov (feedback) control in order to guarantee the admissibility, or more precisely to guarantee the existence and uniqueness of corresponding Markov processes, as well as sufficient smoothness of the cost functions in view of the approximation analysis. In Van Dijk [20] the cash-balance model is analyzed under Lipschitz conditions on the coefficients for a given fixed control. In the present model, however, the decision set is discrete and therefore Lipschitz conditions are not satisfied with decision rules other than constant decision rules. Clearly, the class of controls which at each point of time only allow constant decision rules is too restrictive from an optimization point of view. Therefore, instead we will restrict ourselves to the class of controls such that for some fixed h the current control value $\gamma \in \mathrm{O}$ can only change at times $n\,h$ depending on the state at $n\,h$, hence this value remains constant during $[nh,\, nh+h)$ regardless of the change of the state meanwhile. First, we will show that within each interval $[nh,\, nh+h)$ an admissibility, sufficient boundedness and sufficient smoothness are guaranteed under such a control. Let $h > 0$ be fixed.

Lemma 7.3.1. For any fixed $\gamma = (\gamma_1, \gamma_2) \in \mathrm{O}$ and $x \in \mathbb{R}$ there exists a unique Markov process $\{\eta_t^\gamma(x) | t \le h\}$ and associated family of time-homogeneous expectation operators $\{\boldsymbol{T}_s^\gamma | s \le Z\}$ corresponding to the control object of Section 7.2.1 satisfying (2.2.4) and (2.2.5) and defined by

$$\eta_t^\gamma(x) = x + [\gamma_1]\, t + [\gamma_2]\, W_t, \quad t \le h, \tag{7.3.1}$$

$$\boldsymbol{T}_t^\gamma f(x) = \mathbb{E}\, f(\eta_t^\gamma(x)), \quad x \in \mathbb{R}, \tag{7.3.2}$$

where W_t denotes the standard Wiener measure.

Furthermore, for all $\gamma \in \mathrm{O}$; $g \in B^{\mu_p}$; $f \in C^{3;p}$; $\Delta t \le h$ and with C a constant, we have:

$$||\boldsymbol{T}_{\Delta t}^\gamma\, \mu_p||_{\mu_p} \le (1 + h\, C). \tag{7.3.3}$$

$$||\boldsymbol{T}_{\Delta t}^\gamma\, g||_{\mu_p} \le (1 + h\, C)\, ||g||_{\mu_p}. \tag{7.3.4}$$

$$||[\boldsymbol{T}_{\Delta t}^\gamma - I]\, f||_{\mu_p} \le \sqrt{\Delta t}\, C\, K_f. \tag{7.3.5}$$

$$\left\| \left([\boldsymbol{T}_{\Delta t}^\gamma - I]\, (\Delta t)^{-1} - \boldsymbol{A}^\gamma \right) f \right\|_{\mu_p} \le \sqrt{\Delta t}\, C\, K_f. \tag{7.3.6}$$

P r o o f. The proof can be given similarly to that of Lemma 5.3.11, Chapter I and of Lemma 8.2.17, Chapter II of Van Dijk [20].

Basically, the proof results from the fact that for any fixed (γ_1, γ_2) the coefficients in (7.3.1) are Lipschitz (in fact, constant) and bounded uniformly in (γ_1, γ_2). $\square$

Now consider a (discrete-time) control $\pi = (\delta_0, \delta_1, \ldots, \delta_\ell) \in \Pi^h$. Then instead of by (2.2.6), we can recursively define cost functions V^π_{jh} similarly to (3.2.7) such that for all $x \in I\!R$, $j \leq \ell$ and with $\gamma = \delta_j(x)$:

$$\left.\begin{aligned} V^\pi_{jh}(x) &= \int_0^{Z-\ell h} T^\gamma_s\, L^\gamma(x)\,\mathrm{d}s, && j = \ell, \\ V^\pi_{jh}(x) &= \int_0^{h} T^\gamma_s\, L^\gamma(x)\,\mathrm{d}s + T^\gamma_h\, V^\pi_{jh+h}(x), && j < \ell. \end{aligned}\right\} \tag{7.3.7}$$

According to (7.3.1), (7.3.2) and (7.3.7), the function V^π_{jh} can be interpreted as the expected total costs from jh up to Z, associated with a controlled stochastic differential equation under a control $\overline{\pi}$ which prescribes its control value to change at jh according to the decision rule δ_j and the current state x and to remain constant during $[jh, jh+h)$ regardless of the state evolution. (A precise formulation would require the notion of history dependent controls, which we prefer not to include.) Since at any time t the control value of such a control depends on the actual state at $\lfloor th^{-1} \rfloor h$, where h may be thought of as being small, we will refer to such a control as an h-almost Markov control $\overline{\pi}$.

As will be shown in Section 7.5 or can be found in Section 8.2.4, Chapter II of Van Dijk [20], for any given $\varepsilon > 0$ and under the assumption of a sufficiently smooth solution of the Bellman equation, an ε-optimal control can be found among the class of all h-almost Markov controls by taking the minimum over J in the right hand sides of (7.3.7).

From a computational point of view, however, such an construction may still be unsatisfactory since the computation (or approximation) of the expressions in (7.3.7) still involves the computation (or approximation) of the Wiener increment W_h. It is therefore that in the next two sections we will also investigate the approximation of the optimal cost function and the construction of ε-optimal controls by means of the computational more direct discrete-time structure induced by (7.2.6). In the remainder of this section we will first show that for a given h-almost Markov control the cost functions defined by (7.3.7) can be approximated by their discrete-time analogues associated with (7.2.6).

Consider V^π_{jh} defined by (7.3.7) and V^h_j defined by (3.2.6) with one-step transition probabilities P^δ_h given by (7.2.6) and $\pi = (\delta_0, \delta_1, \ldots, \delta_\ell)$. An essential problem which arises in applying the approximation results for comparing the functions V^h_j and V^π_{jh}, is that the function V^π_{jh} is not necessarily sufficiently smooth, as a result from the fact that decision rules δ_j may be non-Lipschitz. We will therefore first present an approximation result under the assumption that either the function V^h_n or V^π_{nh} is sufficiently smooth. Since such an assumption will be difficult to verify or may fail, we will thereafter extend this result to a more natural form allowing sufficiently smooth approximations of V^h_n.

Notation 7.3.2. Let $C^{3;p}\,\{n \leq \ell\}$ denote the set of families $q = \{q_n | n \leq \ell\}$ such that for some constant, denoted by K_q and all $n \leq \ell$:

$$q_n \in C^{3;p} \quad \text{and} \quad K_{q_n} \leq K_q.$$

Theorem 7.3.3. Suppose that either $\{V_n^h | n \le \ell\}$ or $\{V_{nh}^\pi | n \le \ell\}$ is contained in $C^{3;p}\,\{n \le \ell\}$. Then for some constant C and all $n \le \ell$:

$$||V_n^h - V_{nh}^\pi||_{\mu_p} \le \sqrt{h}\, C. \tag{7.3.8}$$

Proof. Since V_{nh}^π is not defined according to (2.2.6), the setting of Theorem 5.2.1 does not apply directly. Moreover, if $V_{nh}^\pi \notin C^{3;p}$ then Assumption 2.2.3 fails. We will therefore proceed similarly to the proof of the approximation Lemma 4.1. To avoid ambiguity for the notation T_h^γ, the operator T_h^δ defined by (3.1.1) will only be denoted in its equivalent form: $(I + h\, A_h^\delta)$. Then, from (7.3.7) and (3.2.7):

$$\left.\begin{aligned} V_{jh}^\pi(x) &= \left[h\, L^\gamma + T_h^\gamma\, V_{jh+h}^\pi\right](x) + \left[\int_0^h T_s^\gamma\, L^\gamma\, \mathrm{d}s - h\, L^\gamma\right](x) \\ V_j^h(x) &= \left[h\, L^\gamma + (I + h\, A_h^\gamma)\, V_{j+1}^h\right](x), \quad \gamma = \delta_j(x). \end{aligned}\right\} \tag{7.3.9}$$

Suppose that $\{V_n^h | n \le \ell\} \in C^{3;p}$. Define $\delta_j = V_{jh}^\pi - V_j^h$. Then from (7.3.2):

$$\begin{aligned} ||\delta_j||_{\mu_p} \;\le\; & \sup_{\gamma\in O} ||T_h^\gamma\, \delta_{j+1}||_{\mu_p} + \sup_{\gamma\in O} \left\|\int_0^h T_s^\gamma\, L^\gamma\, \mathrm{d}s - h\, L^\gamma\right\|_{\mu_p} \\ & + \sup_{\gamma\in O} \left\|(T_h^\gamma - [I + h\, A^\gamma])\, V_{j+1}^h\right\|_{\mu_p} + \sup_{\gamma\in O} \left\|h(A_h^\gamma - A^\gamma)\, V_{j+1}^h\right\|_{\mu_p}. \end{aligned} \tag{7.3.10}$$

Next, recall relation (7.1.2) for L and the fact that $\{V_n^h | n \le \ell\} \in C^{3;p}\,\{n \le \ell\}$. Then from (7.3.10) together with (7.3.4), (7.3.5), (7.3.6) and (7.2.7):

$$||\delta_j||_{\mu_p} \le (1 + h\, C)\, ||\delta_{j+1}||_{\mu_p} + h\sqrt{h}\, C, \quad j \le \ell. \tag{7.3.11}$$

Furthermore, from (7.3.7) and (7.3.4):

$$||\delta_\ell||_{\mu_p} = ||V_{\ell h}^\pi||_{\mu_p} \le h\, C. \tag{7.3.12}$$

Iterating (7.3.11) for $j = n, n+1, \ldots, \ell - 1$ and using (7.3.12) completes the proof for the case that $\{V_n^h | n \le \ell\} \in C^{3;p}\,\{n \le \ell\}$. For the case that $\{V_{nh}^\pi | n \le \ell\} \in C^{3;p}\,\{n \le \ell\}$ the proof can be given similarly by taking $\delta_j = V_j^h - V_{jh}^\pi$ and using (7.2.8) instead of (7.3.4). □

Theorem 7.3.4. Suppose that for some family $\{q_n^h | n \le \ell\} \in C^{3;p}\,\{n \le \ell\}$ and $\varepsilon > 0$:

$$||q_n^h - V_n^h||_{\mu_p} \le h\, \varepsilon, \quad n \le \ell. \tag{7.3.13}$$

Then

$$||V_n^h - V_{nh}^\pi||_{\mu_p} \le \left(\sqrt{h} + \varepsilon\right), \quad n \le \ell.$$

Proof. By virtue of

$$||(I + h\, A_h^\gamma)\, f||_{\mu_p} \le (1 + h\, C)\, ||f||_{\mu_p}, \tag{7.3.14}$$

together with (7.3.13), the second relation in (7.3.9) can be replaced by

$$\left.\begin{aligned} &\boldsymbol{q}_j^h(x) = [h\,\boldsymbol{L}^\gamma + (I + h\,\boldsymbol{A}_h^\gamma)\,\boldsymbol{q}_{j+1}^h](x) + r_j(x), \\ &\text{where } \|r_j\|_{\mu_p} \le h(\varepsilon\, C) \text{ uniformly in } j \le \ell. \end{aligned}\right\} \tag{7.3.15}$$

By reconsidering the proof of Theorem 7.3.3 with $\boldsymbol{\delta}_j = \boldsymbol{V}_{jh}^\pi - \boldsymbol{q}_j^h$ and using (7.3.15) the proof follows similarly. □

Remark 7.3.5. As in Van Dijk [20], Section 8, Chapter II it follows that $\{\boldsymbol{V}_{nh}^\pi | n \le \ell\} \in C^{3;p}\,\{n \le \ell\}$ if we consider an h-almost Markov controls π such that for any $\boldsymbol{\delta} = \pi(nh)$: $\boldsymbol{\delta}$ is Lipschitz, i.e., $|\boldsymbol{\delta}(y) - \boldsymbol{\delta}(x)| \le K|y - x|$. As a particular case this applies to $\boldsymbol{\delta} = \overline{\gamma}$ for some $\gamma \in \mathrm{O}$.

7.4. Finite horizon optimal cost function: Approximation

Before we can present the main approximation result on the discrete-time approximation of the continuous-time optimal function $\boldsymbol{\Phi}_t$, we first need to justify the existence and sufficient smoothness of $\boldsymbol{\Phi}_t$.

Lemma 7.4.1. There exists a unique family $\{\boldsymbol{\Phi}_t | t \le Z\}$ satisfying:

(i) The continuous-time optimality equation (2.3.2).

(ii) $\boldsymbol{\Phi}_t \in C^{3;p}$ with (see (7.2.1)) $K_{\Theta_t} \le K_\Theta$ for all $t \le Z$ and some K_Θ.

(iii) For any $\lambda \in (0,1)$, some C_λ and all s, $s + \Delta s \le Z$:

$$\left\| \frac{\mathrm{d}}{\mathrm{d}x}\boldsymbol{\Phi}_{s+\Delta s} - \frac{\mathrm{d}}{\mathrm{d}x}\boldsymbol{\Phi}_s \right\|_{\mu_p} + \left\| \frac{\mathrm{d}^2}{\mathrm{d}x^2}\boldsymbol{\Phi}_{s+\Delta s} - \frac{\mathrm{d}^2}{\mathrm{d}x^2}\boldsymbol{\Phi}_s \right\|_{\mu_p} \le (\Delta s)^{\lambda/2}\, C_\lambda. \tag{7.4.1}$$

Proof. This can be argued based on results of Krylov [12] and the differentiability condition on $\boldsymbol{L}$ as by (7.1.2). (For more details see Section 8.6, in Van Dijk and Hordijk [23].) □

Theorem 7.4.2. For any $\lambda \in (0,1)$, some C_λ and all $n \le \ell$:

$$\|\boldsymbol{\Phi}_n^h - \boldsymbol{\Phi}_{nh}\|_{\mu_p} \le h^{\lambda/2}\, C_\lambda. \tag{7.4.2}$$

Proof. We will apply Theorem 5.3.1. First of all, we need to verify Assumptions 2.3.1 and 2.3.2 for the continuous-time model and Assumption 3.3.1 for the discrete-time model. Assumption 2.3.1 is immediate. According to (ii) of Lemma 7.4.1 and (7.2.1):

$$\|\boldsymbol{J}(\boldsymbol{\Phi}_t)\|_{\mu_p} \le C, \tag{7.4.3}$$

which together with Lemma 7.4.1 guarantees Assumption 2.3.2. Finally (7.2.1) together with (7.2.9) implies Assumption 3.3.1 with $F = B^{\mu_p}$. Next from (7.4.1):

$$\|\boldsymbol{J}(\boldsymbol{\Phi}_{s+\Delta s}) - \boldsymbol{J}(\boldsymbol{\Phi}_s)\|_{\mu_p} \le \sup_{\gamma\in\mathrm{O}} \|\boldsymbol{A}^\gamma(\boldsymbol{\Phi}_{s+\Delta s} - \boldsymbol{\Phi}_s)\|_{\mu_p} \le (\Delta s)^{\lambda/2}\, C_\lambda. \tag{7.4.4}$$

Expression (2.3.3) for $R_{nh}(\boldsymbol{\Phi}, h)$ and (7.4.4) yield:

$$||R_{nh}(\boldsymbol{\Phi}, h)||_{\mu_p} h^{-1} \leq h\, C. \tag{7.4.5}$$

Finally, the proof is completed by applying Theorem 5.3.1, (ii) of Lemma 7.4.1, (7.2.8) and (7.4.5). □

7.5. ε-optimal piecewise constant controls

Based on the existence of a sufficiently smooth solution of the continuous-time optimality equation, we can present two ways of constructing ε-optimal controls. The first one results from directly minimizing the cost functions defined by (7.3.7). This still requires the calculation (approximation) of the distributions. The second one in contrast results from using the computationally more direct discrete-time structure (7.2.6). This in turn requires a sufficiently smooth (approximation of the) corresponding cost-function.

Theorem 7.5.1. Let $\{\boldsymbol{V}_n^0 | n \leq \ell\} \in B^{\mu_p}$ defined by:

$$\left.\begin{aligned} \boldsymbol{V}_\ell^0(x) &= \min_{\gamma \in O} \left[\int_0^{Z-\ell h} \boldsymbol{T}_s^\gamma\, \boldsymbol{L}^\gamma(x)\, \mathrm{d}s\right] \\ \boldsymbol{V}_j^0(x) &= \min_{\gamma \in O} \left[\int_0^{h} \boldsymbol{T}_s^\gamma\, \boldsymbol{L}^\gamma(x)\, \mathrm{d}s + \boldsymbol{T}_h^\gamma\, \boldsymbol{V}_{j+1}^0(x),\ j < \ell\right]. \end{aligned}\right\} \tag{7.5.1}$$

Let $\pi = (\delta_0, \delta_1, \ldots, \delta_\ell)$ with $\delta_n \in \Delta$ for all $n \leq \ell$ be such that for some $\varepsilon > 0$, and all $n \leq \ell$:

$$||\boldsymbol{V}_{nh}^\pi - \boldsymbol{V}_n^0||_{\mu_p} \leq \varepsilon. \tag{7.5.2}$$

Then, for any $\lambda \in (0,1)$, some constant C_λ and all $n \leq \ell$:

$$||\boldsymbol{V}_{nh}^\pi - \boldsymbol{\Phi}_{nh}||_{\mu_p} \leq \left(h^{\lambda/2} + \varepsilon\right) C_\lambda. \tag{7.5.3}$$

Proof. Since $||\boldsymbol{V}_{nh}^\pi - \boldsymbol{\Phi}_{nh}||_{\mu_p} \leq ||\boldsymbol{V}_{nh}^\pi - \boldsymbol{V}_n^0|| + ||\boldsymbol{V}_n^0 - \boldsymbol{\Phi}_{nh}||_{\mu_p}$ this follows from (7.5.2) and a proof similar to that of Theorem 8.2.16, Chapter II of Van Dijk [20] but with $h^{\lambda/2}$ instead of $\sqrt{h}$. □

Remark 7.5.2. Analogously to Lemmas 8.2.14 and 8.2.15, Chapter II of Van Dijk [20], it can be shown (by using the finiteness of O) that there exist $(\delta_0, \delta_1, \ldots, \delta_\ell) \in \Delta$, i.e., with δ_i piecewise continuous (constant), such that (7.5.2) holds with $\varepsilon = 0$. Further, a characterization of (7.5.2) based upon the (one-step) expectation operators $\boldsymbol{T}_h^\gamma$ can be given similarly to (6.5.6).

Theorem 7.5.3. Let $\pi^h = (\delta_0, \delta_1, \ldots, \delta_\ell) \in \Pi^h$ such that for some $\varepsilon_1 > 0$, $\varepsilon_2 > 0$ as well as some $\{\boldsymbol{q}_n^h | n \leq \ell\} \in C^{3;p}\{n \leq \ell\}$, and all $n \leq \ell$:

$$\left.\begin{aligned} &||\boldsymbol{V}_n^h - \boldsymbol{\Phi}_n^h||_{\mu_p} \leq \varepsilon_1 \\ &||\boldsymbol{q}_n^h - \boldsymbol{V}_n^h||_{\mu_p} \leq \varepsilon_2\, h. \end{aligned}\right\} \tag{7.5.4}$$

Then with $\pi \in \Pi$ defined by $\pi(t) = \delta(n)$, $n = \lfloor t\, h^{-1} \rfloor$, any $\lambda \in (0,1)$, some constant C_λ and all $n < \ell$:

$$||\boldsymbol{V}^{\pi}_{nh} - \boldsymbol{\Phi}_{nh}||_{\mu_p} \leq \left(h^{\lambda/2} + \varepsilon_1 + \varepsilon_2 \right) C_\lambda. \tag{7.5.5}$$

Proof. By using an inequality as (6.5.3) this follows directly from combining Theorem 7.3.3, Theorem 7.4.2 and relation (7.5.4). □

Remark 7.5.4. A similar lemma as 6.5.3 is valid with $\boldsymbol{T}^{\delta}_{h} = (I + h\, \boldsymbol{A}^{\delta}_{h})$ in order to characterize the first inequality of (7.5.4). Particularly, as in Remark 7.5.2 one can show the existence of piecewise constant decision rules such that $\varepsilon_1 = 0$. In this case the function $\boldsymbol{V}^{h}_{n} = \boldsymbol{\Phi}^{h}_{n}$ also appears to be Lipschitz in x and n, so that standard (polynomial) approximation procedures may yield the second inequality with reasonably small ε_2 in (7.5.4).

Remark 7.5.5. As in Remark 6.5.4, also here the application of Theorems 7.4.2 and 7.5.7 together with the recursive scheme (7.3.7) or (7.5.1) lead to an *algorithm* for computing an ε-optimal control, similar to (6.5.8).

Remark 7.5.6. As in Remark 6.5.5, in contrast with results in Bensoussan and Robin [2], Fleming and Rishel [6], Gihman and Shorohod [7], Hausmann [9], Krylov [12], Kushner [13], Pliska [16], Puterman [18], note that the results above lead to the construction of ε-optimal controls with

(i) simple one-step transition probabilities as given in (7.2.6)

(ii) prespecified accuracy-value ε.

ACKNOWLEDGEMENT

The authors are highly grateful to Kybernetika for the opportunity to publish their extensive technical reports [22] and [23] in the present more compact form.

(Received June 7, 1994.)

REFERENCES

[1] A. Bensoussan and M. Robin: On the convergence of the discrete time dynamic programming equation for general groups. SIAM J. Control Optim. *20* (1982), 1, 722–746.

[2] N. Christopeit: Discrete approximation of continuous time stochastic control systems. SIAM J. Control Optim. *21* (1983), 1, 17–40.

[3] G. M. Constantinides: Stochastic cash management with fixed and proportional transaction costs. Management Sci. *22* (1974), 1320–1331.

[4] G. M. Constantinides and S. F. Richard: Existence of optimal simple policies for discounted-costs inventory and cash management in continuous time. Oper. Res. *26* (1978), 620–936.

[5] B. T. Doshi: Optimal control of the service rate in an $M \mid G \mid 1$-queueing system. Adv. in Appl. Probab. *10* (1978), 862–701.

[6] W.H. Fleming and R.W. Rishel: Deterministic and Stochastic Optimal Control. Springer Verlag, Berlin 1975.
[7] I.I. Gihman and A.V. Skorohod: Controlled Stochastic Processes. Springer Verlag, Berlin 1979.
[8] J.M. Harrison and M.I. Taksar: Instantaneous Control of Brownian Motion. Technical Report No. 63, Stanford University 1981.
[9] U.G. Haussmann: A discrete approximation to optimal stochastic controls. In: Analysis and Optimization of Stochastic Systems, Academic Press, London 1980, pp. 229–241.
[10] A. Hordijk and F.A. Van der Duyn Schouten: Average optimal policies in Markov decision drift processes with applications to a queueing and a replacement model. Adv. in Appl. Probab. *15* (1983), 274–303.
[11] A. Hordijk and F.A. Van der Duyn Schouten: Discretization and weak convergence in Markov decision drift processes. Math. Oper. Res. *9* (1984), 1, 112–141.
[12] N.V. Krylov: Controlled Diffusion Processes. Springer Verlag, New York 1980.
[13] H.J. Kushner: Probability Methods for Approximation in Stochastic Control and for Elliptic Equations. Academic Press, New York 1977.
[14] B. Mitchell: Optimal service rate selection in an $M/G/1$ queue. SIAM J. Appl. Math. *24* (1973), 19–35.
[15] S.R. Pliska: Controlled jump processes. Stochastic Process. Appl. *3* (1975), 259–282.
[16] S.R. Pliska: Single and Multi-Person Controlled Diffusions. PhD Thesis, Stanford University 1972.
[17] S.R. Pliska: Single–person controlled diffusion with discounted costs. J. Optim. Theory Appl. *12* (1973), 3, 248–255.
[18] M.L. Puterman: On the Optimal Contrast of Diffusion Processes. PhD Thesis, Stanford University 1972.
[19] M.L. Puterman: Sensitive discount optimality in controlled one-dimensional diffusion. Ann. Probab. *2* (1974), 3, 408–419.
[20] N.M. Van Dijk: Controlled Markov Processes; Time Discretization/Networks of Queues. PhD Thesis, University of Leiden 1983.
[21] N.M. Van Dijk: On the finite horizon Bellman equation for controlled Markov jump models with unbounded characteristics: existence and approximations. Stoch. Proc. Appl. *28* (1988), 141–157.
[22] N.M. Van Dijk and A. Hordijk: Controlled Markov Processes: Time–Discretization. Part I. Research Report, University of British Columbia, Faculty of Commerce and Business Administration, Vancouver 1985.
[23] N.M. Van Dijk and A. Hordijk: Controlled Markov Processes: Time–Discretization. Part II. Research Report, University of British Columbia, Faculty of Commerce and Business Administration, Vancouver 1985.
[24] N.M. Van Dijk and A. Hordijk: Time-discretization for controlled Markov processes. Part I: General approximation results. Kybernetika *32* (1996), 1–16.
[25] A.A. Yushkevich: Controlled jump Markov models. Theory Probab. Appl. *25* (1980), 244–266.

Prof. Dr. Nico M. van Dijk, Faculty of Economic Sciences and Econometrics, University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam. The Netherlands.

Prof. Dr. Arie Hordijk, Department of Mathematics and Computer Science, University of Leiden, Niels Bohrweg 1, P. O. Box 9512, 2300 RA Leiden. The Netherlands.