Ergodic and Adaptive Control of Hidden Markov Models*

T. E. Duncan† B. Pasik-Duncan† L. Stettner‡

Abstract

A partially observed stochastic system is described by a discrete time pair of Markov processes. The observed state process has a transition probability that is controlled and depends on a hidden Markov process that also can be controlled. The hidden Markov process is completely observed in a closed set, which in particular can be the empty set and only observed through the other process in the complement of this closed set. Initially an ergodic control problem is formulated and solved by a vanishing discount approach.

For an adaptive control problem, it is assumed that the transition operators for the observed state process and the hidden Markov process depend on a parameter from a compact metric space and the closed set, where the hidden Markov process is completely observed, is nonempty and recurrent. An adaptive control is constructed based on this family of estimates that is almost optimal.

Keywords: stochastic adaptive control, control of hidden Markov models, ergodic control, almost optimal adaptive control.

---

*Research supported in part by NSF Grants DMS 0204669 and ANI 0125410 and PBZ KBN 016 / P03 / 99.

†Department of Mathematics, University of Kansas, Lawrence, KS 66045, duncan@math.ukans.edu and bozenna@math.ukans.edu

‡Institute of Mathematics, Polish Academy of Sciences, Warsaw, Poland, stettner@impan.gov.pl.

1 Introduction

Hidden Markov processes provide an important family of models for partially observed stochastic systems. These models have been used in many applications (e.g., [10] and the references therein). A number of related problems appears in literature also under the name of jump systems (see e.g. [14]). It is assumed that the dynamics of the controlled state process depends on another hidden (jump), (possibly controlled) Markov process. Particular applications of such models are for example: placement of failure prone actuators on large space structures ([15]), control of manufacturing systems ([18]), analysis of transient electrical power networks ([21]), economic policy planning (see [12]) or modelling of asset processes explicitly depending on underlying factors (such as interest rates, unemployment rates and dividend yields) considered in [6].

In this paper, the observed state process and the hidden Markov process take values in two Polish spaces. It is assumed that the hidden Markov process is completely observed in a closed set of its state space. The observed state process and the hidden Markov process are controlled and the state process depends on the hidden Markov process. It is assumed that the hidden Markov process is observed in a closed set. Such situation frequently occurs in various applications. In particular, when manufacturing systems are controlled, the production capacity process (hidden Markov process) is observed completely only in specific states (breakdown of the system, periodical maintenance) (see [18]), while in other states it is observed only through the value of production. Furthermore, in mathematics of finance the dynamics of asset prices (frequently its volatility) depends on various economical factors (hidden Markov processes) e.g. divident yields, rate on inflation, short rate interest rates which are observed partially (frequetly with delay) with the exception to their extremal values, e.g. [1] and [6]. To be more specific consider the following

Example. We are given a market consisting of d securities. The price $S_i(n+1)$ of the i-th security at time $n+1$ depends on the value of economical factor θ_n at time n in the following way

$$\frac{S_i(n+1)}{S_i(n)} = \exp\left\{a(\theta_n) + b(\theta_n)w_n\right\}, \tag{1.1}$$

where (θ_n) is an ergodic Markov chain on the finite space E_0, the drift a and the volatility b are real valued functions on E_0 and (w_n) is a sequence of i.i.d. random variables (in particular Gaussian). We assume that extremal states of the economy are recognizable, while the other are observed through the prices of the securities only. Alternatively one can consider (θ_n) as a credit migration process taking values in a finite set of credit rating classes which describes the credit ability of the issuer of the security (see [2] for the methodology). Given initial capital $V(0)$ we invest in securities do not allowing short selling and short borrowing. Let $h_i(n)$ denote the portion of the capital (wealth process) invested at time n in the i-th security. Then the wealth process at time $n+1$ is of the form

$$\begin{aligned} V(n+1) &= V(n) + \sum_{i=1}^{d} \frac{h_i(n)V(n)}{S_i(n)}(S_i(n+1) - S_i(n)) \\ &= V(n)\left(\sum_{i=1}^{d} h_i(n)\frac{S_i(n+1)}{S_i(n)}\right) \qquad (1.2). \end{aligned}$$

Consequently

$$\ln V(n+1) = \sum_{k=1}^{n} \ln\left(\sum_{i=1}^{d} h_i(k)\frac{S_i(k+1)}{S_i(k)}\right) + \ln V(0),$$

and the long run expected growth of the portfolio is equal to

$$\liminf_{n\to\infty} \frac{1}{n} E\left\{\ln V(n)\right\} = \liminf_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} E\left\{\ln\left(\sum_{i=1}^{d} h_i(k)\frac{S_i(k+1)}{S_i(k)}\right)\right\}. \tag{1.3}$$

We are interested to maximize the functional (1.3), i.e. want to find so called (see section 6.2 of [13]) asymptotically optimal constant rebalanced portfolio. We observe the prices of the securities and the factors (θ_n) are only observed at specific states the set of which we denote by E_0^0. We don't know the transition probabilities of the chain (θ_n). Notice that the state processes

(θ_n) and $(S_i(n))$ do not depend on the control process $(h_i(n))$ which appears in the cost functional only. Trying to solve the problem formulated in the above example we shall develop below a more general theory allowing to control both state processes.

Initially an ergodic control problem is formulated and solved where the solution is obtained by a vanishing discount approach following some arguments of [17] and [19]. For the ergodic control problem, the closed set where the hidden Markov process is completely observed can be the empty set. This ergodic control problem is a generalization of the problem in [19]. Some remarks concerning possible extensions using recent results of [3] and [4] are also added at the end of Section 3.

For the adaptive control problem, it is assumed that the unknown parameter is a random variable and the closed set where the hidden Markov process is completely observed is required to be nonempty and ergodic. A sequence of estimates is given by a conditioning and this family of estimates is modified at random times by modifying the control at these random times by choosing a control from a fixed finite family of controls. This resulting sequence of estimates is shown to converge to the true random parameter. A family of adaptive controls using this family of estimates is shown to be almost optimal. The adaptive control of partially observed Markov models is a difficult problem and there are few results (e.g., [9], [21]).

A brief outline of this paper is given. In Section 2, the stochastic model is described. An ergodic cost and a discounted cost are given. In Section 3, the discounted cost control problem is solved. A Bellman equation is given for the solution of this control problem. When the discount factor tends to 1, it is shown that an equation for the solution of the ergodic control problem is obtained. In Section 4, some continuity properties of the ergodic control problem are obtained, e.g., the optimal cost is a continuous function of the parameter. In Section 5, the parameter is assumed to be a random variable with a known probability law. A sequence of estimates is obtained by conditioning and the sequence of controls that is used is modified at random times. This resulting sequence of estimates is shown to converge to the true value. In Section 6, a sequence of randomized controls based on the parameter estimates is shown to be almost self-optimal.

2 A Stochastic Model

The stochastic system is described by the pair $((\theta_n, X_n),\, n \in \mathbb{N})$ on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n), P)$. The process $(\theta_n,\, n \in \mathbb{N})$ is a (partially) hidden Markov process with the controlled transition probability $P_0^{v_n}(\theta_n, d\theta)$ at the generic time $n \in \mathbb{N}$ and the control v_n. This process has state space E_0 that is a Polish space. This process is completely observed in the closed set $E_0^0 \subset E_0$ and only observed through the state process $(X_n,\, n \in \mathbb{N})$ in the set $E_0^1 = E_0 \setminus E_0^0$. The state space for the process $(X_n,\, n \in \mathbb{N})$ is the Polish space E with the controlled transition probability

$$P^{\theta_n, v_n}(X_n, dx) = r(X_n, \theta_n, v_n, x)\eta(dx)$$

where η is a fixed probability measure and v_n is a control with values in a compact set U and is adapted to the available observations of the processes $(\theta_j,\, j \in \{0, \ldots, n\})$ and $(X_j,\, j \in \{0, \ldots, n\})$.

To describe precisely the evolution of the stochastic system, an auxiliary process $(\bar{\theta}_n,\, n \in \mathbb{N})$ is defined as follows

$$\bar{\theta}_n = \begin{cases} \theta_n & \text{if } \theta_n \in E_0^0 \\ \Delta & \text{if } \theta_n \in E_0^1 \end{cases} \tag{2.1}$$

where Δ is an artificial state that is not in E. Let $X^n = \sigma\{X_0, \ldots, X_n\}$, $\theta^n = \sigma\{\theta_0, \ldots, \theta_n\}$ and $\bar{\theta}^n = \sigma\{\bar{\theta}_0, \ldots, \bar{\theta}_n\}$. Clearly the control v_n is adapted to the σ-field $X^n \vee \bar{\theta}^n$. Let $\mathcal{B}(E_0)$ and $\mathcal{B}(E)$ be the Borel σ-fields of E_0 and E respectively. For $A \in \mathcal{B}(E_0)$ and $n \in \mathbb{N}$ define π_n as follows

$$\pi_n(A) = \bar{M}^{v_{n-1}\bar{\theta}_n}(X_{n-1}, X_n, \pi_{n-1})(A) \tag{2.2}$$

where

$$\bar{M}^{v\bar{\theta}}(x, z, \mu)(A) := 1_{E_0^0}(\bar{\theta})\delta_{\bar{\theta}}(A) + 1_{\Delta}(\bar{\theta})M^v(x, z, \mu)(A)$$

$$M^v(x, z, \mu)(A) = \frac{\displaystyle\int_{E_0} P_0^v(\theta, A \cap E_0^1) r(x, \theta, v, z)\mu(d\theta)}{\displaystyle\int_{E_0} P_0^v(\theta, E_0^1) r(x, \theta, v, z)\mu(d\theta)}$$

and δ_θ is the Dirac measure at θ. If ξ is the law of θ_0, then

$$\pi_0(A) = 1_{E_0^0}(\bar{\theta}_0)\delta_{\bar{\theta}_0}(A) + 1_\Delta(\bar{\theta}_0)\frac{\xi(A \cap E_0^1)}{\xi(E_0^1)}. \tag{2.3}$$

The following lemma is easily verified by standard methods (e.g., Lemma 1.1 in [17]).

Lemma 1 *For $n \in \mathbb{N} = \{0, 1, 2, \ldots\}$ the following equality is satisfied*

$$\pi_n(A) = P(\theta_n \in A \mid X^n, \bar{\theta}^n) \qquad a.s.\ P \tag{2.4}$$

The behavior of the stochastic system with the observation at time n that is given by $X^n \vee \bar{\theta}^n = \sigma(X_0, \ldots, X_n, \bar{\theta}_0, \ldots, \bar{\theta}_n)$ is described by the triple $(\pi_n, \bar{\theta}_n, X_n)$. This triple is a Markov process.

Lemma 2 *The process $((\pi_n, \bar{\theta}_n, X_n),\ n \in \mathbb{N})$ is a controlled Markov process with respect to the filtration $(X^n \vee \bar{\theta}^n,\ n \in \mathbb{N})$ and the transition operator Π^v is given by*

$$\begin{aligned}
&\Pi^v(\mu, \bar{\theta}, x, f) \\
&\quad = 1_{E_0^0}(\bar{\theta})\left[\int_E \int_{E_0} f(\delta_\theta, \theta, z) P_0^v(\bar{\theta}, d\theta) r(x, \bar{\theta}, v, z)\eta(dz)\right. \\
&\qquad\qquad \left. + \int_E f(M^v(x, z, \delta_{\bar{\theta}}), \Delta, z) P_0^v(\bar{\theta}, E_0^1) r(x, \bar{\theta}, v, z)\eta(dz)\right] \\
&\qquad + 1_\Delta(\bar{\theta})\left[\int_{E_0^1}\int_E \int_{E_0^0} f(\delta_{\theta'}, \theta', z) P_0^v(\theta, d\theta') r(x, \theta, v, z)\eta(dz)\mu(d\theta)\right. \\
&\qquad\qquad \left. + \int_{E_0^1}\int_E f(M^v(x, z, \mu), \Delta, z) P_0^v(\theta, E_0^1) r(x, \theta, v, z)\eta(dz)\mu(d\theta)\right].
\end{aligned} \tag{2.5}$$

for $f \in \mathcal{B}(\mathcal{P}(E_0) \times E_0^0 \cup \{\Delta\} \times E)$ that is the space of all bounded Borel measurable functions on $\mathcal{P}(E_0) \times E_0^0 \cup \{\Delta\} \times E$.

Proof. The following equalities are easily verified.

$$\mathbb{E}\left[f(\pi_{k+1}, \bar{\theta}_{n+1}, X_{n+1}) \mid X^n \vee \bar{\theta}^n\right]$$

$$
\begin{aligned}
&= 1_{E_0^0}(\bar\theta_n)\mathbb{E}\left[f\left(\bar M^{v_n\bar\theta_{n+1}}(X_n,X_{n+1},\delta_{\bar\theta_n}),\bar\theta_{n+1},X_{n+1}\right)\mid X^n\vee\bar\theta^n\right]\\
&\quad + 1_\Delta(\bar\theta_n)\mathbb{E}\left[f\left(\bar M^{v_n\bar\theta_{n+1}}(X_n,X_{n+1},\pi_n),\bar\theta_{n+1},X_{n+1}\right)\mid X^n\vee\bar\theta^n\right]\\
&= 1_{E_0^0}(\bar\theta_n)\mathbb{E}\Bigg[\int_{E_0^0} f(\delta_\theta,\theta,X_{n+1})P^{v_n}(\bar\theta_n,d\theta)\\
&\qquad\qquad + f\left(M^{v_n}(X_n,X_{n+1},\delta_{\bar\theta_n}),\Delta,X_{n+1}\right)P_0^{v_n}(\bar\theta_n,E_0^1)\mid X^n\vee\bar\theta^n\Bigg]\\
&\quad + 1_\Delta(\bar\theta_n)\mathbb{E}\Bigg[\int_{E_0^0} f(\delta_\theta,\theta,X_{n+1})P^{v_n}(\theta_n,d\theta)\\
&\qquad\qquad + f\left(M^{v_n}(X_n,X_{n+1},\pi_n),\Delta,X_{n+1}\right)P_0^{v_n}(\bar\theta_n,E_0^1)\mid X^n\vee\bar\theta^n\Bigg]\\
&= 1_{E_0^0}(\bar\theta_n)\mathbb{E}\Bigg[\int_E\int_{E_0^0} f(\delta_\theta,\theta,z)P_0^{v_n}(\bar\theta_n,d\theta)r(X_n,\bar\theta_n,v_n,z)\eta(dz)\\
&\qquad\qquad + \int_E f\left(M^{v_n}(X_n,z,\delta_{\bar\theta_n}),\Delta,z\right)P_0^{v_n}(\bar\theta_n,E_0^1)\\
&\qquad\qquad\qquad \times r(X_n,\bar\theta_n,v_n,z)\eta(dz)\Bigg]\\
&\quad + 1_\Delta(\bar\theta_n)\Bigg[\int_{E_0^1}\int_E\int_{E_0^0} f(\delta_\theta,\theta,z)P_0^{v_n}(\theta',d\theta)r(X_n,\theta',v_n,z)\eta(dz)\pi_n(d\theta)\\
&\qquad\qquad + \int_{E_0^1}\int_E f\left(M^{v_n}(X_n,z,\pi_n),\Delta,z\right)P_0^{v_n}(\theta,E_0^1)\\
&\qquad\qquad\qquad \times r(X_n,\theta,v_n,z)\eta(dz)\pi_n(d\theta)\Bigg].
\end{aligned}
$$

This verifies the Markov property. □

Note that the process $(\pi_n,\, n\in\mathbb{N})$ has the state space $\mathcal{P}(E_0^1)\cup\{\delta_\xi,\,\xi\in E_0^0\}$. The following lemma is important to obtain regularity properties of the discounted cost value function in the next section.

Lemma 3 *If $f:\mathcal{P}(E_0)\times E_0^0\cup\{\Delta\}\times E\to\mathbb{R}$ is bounded and Borel measurable and for each $x\in E$ the function $f(\cdot,\Delta,x):\mathcal{P}(E_0)\to\mathbb{R}$ is concave, then for* *each $x \in E$ and $v \in U$ the function*

$$\Pi^v(\cdot, \Delta, x, f) : \mathcal{P}(E_0^1) \to \mathbb{R}$$

is also concave.

Proof. Assume that the initial law, ξ, of ξ_0 is random such that for some $\mu_1, \mu_2 \in \mathcal{P}(E_0^1)$, $\mathbb{P}(\xi = \mu_1) = \gamma > 0$ and $\mathbb{P}(\xi = \mu_2) = 1 - \gamma > 0$. Let $\sigma(\xi)$ be the Borel σ-field generated by ξ. For $A \in \mathcal{B}(E_0)$, $\bar{\theta}_0 = \Delta$ and fixed deterministic $v_n = v \in U$ define $\hat{\pi}$ as follows for $n \geq 1$

$$\hat{\pi}_n(A) = P(\theta_n \in A \mid X^n, \bar{\theta}^n, \sigma(\xi))$$

and

$$\hat{\pi}_0 = \xi(A).$$

From the proof of Lemma 2, it can be shown that $((\hat{\pi}_n, \bar{\theta}_n, X_n),\, n \in \mathbb{N})$ is a Markov process with the transition operator Π^v given by (2.5). Thus,

$$\Pi^v(\xi, \Delta, x, f) = \mathbb{E}\left[f(\hat{\pi}_1, \bar{\theta}_1, x_1) \mid X_0 = x, \bar{\theta}_0 = \Delta, \sigma(\xi)\right].$$

Since $\mathbb{E}[\hat{\pi}_1 \mid X^1, \bar{\theta}^1] = \pi_1$ it follows by Jensen's inequality, the concavity of $f(\cdot, \Delta, x_1)$ and the equality $\pi_1 = \hat{\pi}_1 = \delta_{\theta_1}$ for $\theta_1 \in E_0^0$ that

$$\begin{aligned}
&\gamma \Pi^v(\mu, \Delta, x, f) + (1-\gamma)\Pi^v(\mu, \Delta, x, f) \\
&\quad = E\left[f(\hat{\pi}_1, \bar{\theta}_1, X_1) \mid X_0 = x, \bar{\theta}_0 = \Delta\right] \\
&\quad \leq \mathbb{E}\left[1_{E_0^0}(\bar{\theta}_1) f(\pi_1, \bar{\theta}_1, X_1) \mid X_0 = x, \bar{\theta}_0 = \Delta\right] \\
&\qquad + \mathbb{E}\left[1_\Delta(\bar{\theta}_1) f(\pi_1, \bar{\theta}_1, x_1) \mid X_0 = x, \bar{\theta}_0 = \Delta\right] \\
&\quad = \Pi^v(\gamma\mu + (1-\gamma)\mu, \Delta, x, f).
\end{aligned}$$

This inequality verifies the concavity. □

Now an ergodic cost functional is introduced. Let $c : E_0 \times E \times U \to \mathbb{R}_+$ be a continuous, bounded function. The family of admissible controls is the family of U-valued, $((X^n \vee \bar{\theta}^n),\, n \in \mathbb{N})$ adapted processes denoted generically as $(v_n,\, n \in \mathbb{N})$. A control is sought from this family that minimizes J given by

$$J_{\xi,\bar{\theta},x}((v_n,\, n \in \mathbb{N})) = \limsup_{N\to\infty} \frac{1}{N} \mathbb{E}_{\xi,\bar{\theta},x} \sum_{i=0}^{N-1} c(\theta_i, X_i, v_i). \tag{2.6}$$

Since a vanishing discount approach is used, initially a discounted cost functional J^β is considered, that is,

$$J^\beta_{\xi,\bar{\theta},x}((v_n,\, n \in \mathbb{N})) = \mathbb{E}_{\xi,\bar{\theta},x} \sum_{i=0}^{\infty} \beta^i c(\theta_i, X_i, v_i) \tag{2.7}$$

with $\beta \in (0,1)$. An analysis of the asymptotic properties of z^β is made, where

$$z^\beta(\xi,\bar{\theta},x) = w^\beta(\xi,\bar{\theta},x) - \inf_{(\xi',\bar{\theta}',x') \in \mathcal{P}(E_0)\times E_0^0 \cup \{\Delta\} \times E} w^\beta(\xi',\bar{\theta}',x') \tag{2.8}$$

and

$$w^\beta(\xi,\bar{\theta},x) = \inf_{(v_n,\, n\in\mathbb{N})} J^\beta_{\xi,\bar{\theta},x}((v_n,\, n \in \mathbb{N})). \tag{2.9}$$

The following Feller property of the transition operator Π^v is used subsequently. If $f : \mathcal{P}(E_0) \times E_0^0 \cup \{\Delta\} \times E \to \mathbb{R}$ is bounded and continuous then

$$\Pi^v f : \mathcal{P}(E_0) \times E_0^0 \cup \{\Delta\} \times E \times U \to \mathbb{R} \tag{2.10}$$

is also continuous. One can easily notice that continuity and boundedness of r as well as continuity and boundedness of the transition density of P_0 with respect to a probability measure (which in fact is assumed starting from section 4) is a sufficient condition for the Feller property of Π^v.

3 Some Properties of the Discounted Value Function

Initially some important properties are verified for the discounted value function.

Theorem 1 *Assume that the transition operator Π^v is Feller. Then*

$$w^\beta : \mathcal{P}(E_0^1) \times E_0^0 \cup \{\Delta\} \times E \to \mathbb{R}$$

is bounded and continuous and it is the unique solution of the following Bellman (dynamic programming) equation

$$w^\beta(\xi,\bar\theta,x)=\inf_{v\in U}[\bar c(\mu,\bar\theta,x,v)+\beta\Pi^v(\mu,\bar\theta,x,w^\beta)] \tag{3.1}$$

where

$$\bar c(\mu,\bar\theta,x,v)=1_{E_0^0}(\bar\theta)c(\bar\theta,x,v)+1_\Delta(\bar\theta)\int_{E_0^1}c(\theta,x,v)\mu(d\theta).$$

Furthermore, the sequence $(w_n^\beta,\ n\in\mathbb{N})$ *where*

$$\begin{aligned} w_n^\beta(\mu,\bar\theta,x) &= T_\beta^n O(\mu,\bar\theta,x) \\ O(\mu,\bar\theta,x) &\equiv 0 \end{aligned} \tag{3.2}$$

and

$$T_\beta^n f(\mu,\bar\theta,x)=\inf_{v\in U}[\bar c(\mu,\bar\theta,x,v)+\beta\Pi^v(\mu,\bar\theta,x,f)] \tag{3.3}$$

is nondecreasing for each $x\in E$ *and*

$$w_n^\beta(\cdot,\Delta,x):\mathcal{P}(E_0^1)\to\mathbb{R}$$

is concave. The sequence $(w_n^\beta,\ n\in\mathbb{N})$ *converges uniformly to* w^β*, that is,*

$$w^\beta(\mu,\bar\theta,x)=\lim_{n\to\infty}w_n^\beta(\mu,\bar\theta,x) \tag{3.4}$$

and for each $x\in E$*, the function*

$$w^\beta(\cdot,\Delta,x):\mathcal{P}(E_0^1)\to\mathbb{R} \tag{3.5}$$

is also concave.

Proof. From the Feller property, it follows that T_β is a contraction in $C(\mathcal{P}(E_0^1)\times E_0^0\cup\{\Delta\}\times E)$. Thus, there is a unique, bounded, continuous solution to the equation (3.1) and a sequence $(w_n^\beta,\ n\in\mathbb{N})$ that converges to the unique solution. Since $\bar c\geq 0$, it follows that $T_\beta O\geq 0$ and by induction it follows that the sequence $(w_n^\beta,\ n\in\mathbb{N})$ is nondecreasing. Moreover, $T_\beta O(\cdot,\Delta,x)$ is concave and by induction from Lemma 3 and the form (3.3) of T_β, the functions $w_n^\beta(\cdot,\Delta,x)$, for each $x\in E$ are also concave. Thus

$w^\beta(\cdot,\Delta,x)$, as a limit of concave function $(w_n^\beta(\cdot,\Delta,x),\, n\in\mathbb{N})$, is concave. By a standard dynamic programming method ([11]) it follows that

$$w^\beta(\mu,\bar\theta,x)=\inf_{(v_n,\, n\in\mathbb{N})} J_{\mu,\bar\theta,x}((v_n,\, n\in\mathbb{N})). \qquad \square$$

Remark 1 Note that if $\bar\theta\in E_0^0$ then the value function $w^\beta(\mu,\bar\theta,x)$ does not depend on $\mu\in\mathcal{P}(E_0^1)$.

Let z_n^β for $n\in\mathbb{N}$ be given by

$$z_n^\beta(\mu,\bar\theta,x)=w_n^\beta(\mu,\bar\theta,x)-\inf_{(\mu',\bar\theta',x')\in\mathcal{P}(E)\times E_0^0\cup\{\Delta\}\times E} w_n^\beta(\mu',\bar\theta',x'). \tag{3.6}$$

By Theorem 1, the sequence $(z_n^\beta,\, n\in\mathbb{N})$ converges uniformly to $z^\beta=w^\beta-\inf w^\beta$. It is shown subsequently that with some additional assumptions $(z_n^\beta,\, n\in\mathbb{N})$ is uniformly bounded in $n\in\mathbb{N}$ and $\beta\in(0,1)$. Assume that the following inequalities are satisfied

$$\lambda_1=\inf_{x,x',z\in E}\inf_{\theta,\theta'\in E_0}\inf_{v,v'\in U}\frac{r(x',\theta',v',z)}{r(x,\theta,v,z)}>0 \tag{3.7}$$

and

$$\lambda_2=\inf_{\theta,\theta'\in E_0}\inf_{v,v'\in U}\inf_{\substack{A\in\mathcal{B}(E_0)\\ P_0^v(\theta,A)>0}}\frac{P_0^{v'}(\theta',A)}{P_0^v(\theta,A)}>0 \tag{3.8}$$

where the equality $0/0=0$ is defined. Clearly $\lambda_1,\lambda_2\le 1$. The constants λ_1 and λ_2 can be used to bound z^β.

Remark 2 The assumptions (3.7) and (3.8) imply that the transition probabilities $P_0^v(\theta,\cdot)$ and $P^{\theta,v}(x,\cdot)$ for different initial states and controls are absolutely continuous with densities that are uniformly bounded from above and uniformly bounded away from zero. For the example considered in section 1 the assumption (3.8) is clearly satisfied, when we assume that the transition matrix of (θ_n) has positive entries. The assumption (3.7) is satisfied, when (w_n) have positive densities with the tails of the same order as (w_n) shifted and multiplied by a constant. Since for the purpose of adaptive control we shall need E to be compact, we assume that $E=[-M,M]^d$ with large M and the value of $S_i(n+1)$ we obtain from (1.1) by the projection on $[-M, M]$. In such case (3.7) is satisfied whenever the density of (w_n) is positive and continuous.

Proposition 1 *If* (3.7) *and* (3.8) *are satisfied and* Π^v *satisfies the Feller property, then*

$$\sup_{\mu \in \mathcal{P}(E_0^1)} \sup_{\bar{\theta} \in E_0^0 \cup \{\Delta\}} \sup_{x \in E} z^\beta(\mu, \bar{\theta}, x) \leq \frac{\|c\|}{\lambda_1^3 \lambda_2^3} \tag{3.9}$$

where

$$\|c\| = \sup_{\theta \in E_0} \sup_{x \in E} \sup_{v \in U} c(\theta, x, v).$$

Proof. For $(\mu, \bar{\theta}, x), (\mu', \bar{\theta}', x') \in \mathcal{P}(E_0^1) \times E_0^0 \cup \{\Delta\} \times E$ there are $v, v' \in U$ such that

$$\begin{aligned}
&w_{n+1}^\beta(\mu, \bar{\theta}, x) - w_{n+1}^\beta(\mu', \bar{\theta}', x') \\
&\quad = \bar{c}(\mu, \bar{\theta}, x, v) - \bar{c}(\mu', \bar{\theta}', x', v') + \beta \left[\Pi^v(\mu, \bar{\theta}, x, w_n^\beta) - \Pi^{v'}(\mu', \bar{\theta}', x', w_n^\beta) \right] \\
&\quad = \bar{c}(\mu, \bar{\theta}, x, v) - \bar{c}(\mu', \bar{\theta}', x', v') + \beta \left[\Pi^v(\mu, \bar{\theta}, x, z_n^\beta) - \Pi^{v'}(\mu', \bar{\theta}', x', z_n^\beta) \right].
\end{aligned} \tag{3.10}$$

By the inequalities (3.7) and (3.8), it follows that

$$\int_{E_0} P_0^{v'}(\theta', A \cap E_0^1) r(x', \theta', v', z) \mu'(d\theta') \geq \lambda_1 \lambda_2 P_0^v(\theta, A \cap E_0^1) r(x, \theta, v, z)$$

for each $\theta \in E$, $A \in \mathcal{B}(E_0)$ and

$$\int_{E_0} P_0^{v'}(\theta', A \cap E_0^1) r(x', \theta', v', z) \mu'(d\theta') \geq \lambda_1 \lambda_2 \int_{E_0} P_0^v(\theta, A \cap E_0^1) r(x, \theta, v, z) \mu(d\theta).$$

Similarly,

$$\lambda_1 \lambda_2 \int_{E_0} P_0^{v'}(\theta', E_0^1) r(x', \theta', v', z) \mu'(d\theta') \leq \int_{E_0} P_0^v(\theta, E_0^1) r(x, \theta, v, z) \mu(d\theta).$$

Thus

$$M^{v'}(x', z, \mu')(A) \geq \lambda_1^2 \lambda_2^2 M^v(x, z, \mu)(A). \tag{3.11}$$

If $\lambda_1 = \lambda_2 = 1$, then the functions w_n^β and w^β are constants and $z^\beta \equiv 0$ so it is assumed that $\lambda_1\lambda_2 < 1$.

Let m be given as

$$
\begin{aligned}
&m(x, x', z, v, v', \mu, \mu')(A) \\
&\quad = \frac{1}{1-\lambda_1^2\lambda_2^2}\left[M^{v'}(x', z, \mu')(A) - \lambda_1^2\lambda_2^2 M^v(x, z, \mu)(A)\right].
\end{aligned}
$$

Clearly $m \in \mathcal{P}(E_0)$ and

$$
M^{v'}(x', z, \mu') = \lambda_1^2\lambda_2^2 M^v(x, z, \mu) - (1-\lambda_1^2\lambda_2^2) m(x, x', z, v, v', \mu, \mu').
$$

By the concavity and nonnegativity of z_n^β (cf. (3.6)), it follows that

$$
\begin{aligned}
&z_n^\beta\left(M^{v'}(x', z, \mu'), \Delta, z\right) \\
&\quad \geq \lambda_1^2\lambda_2^2 z_n^\beta\left(M^v(x, z, \mu), \Delta, z\right) \\
&\qquad - (1-\lambda_1^2\lambda_2^2) z_n^\beta(m(x, x', z, v, v', \mu, \mu'), \Delta, z) \\
&\quad \geq \lambda_1^2\lambda_2^2 z_n^\beta\left(M^v(x, z, \mu), \Delta, z\right). \qquad (3.12)
\end{aligned}
$$

Similarly,

$$
z_n^\beta\left(M^{v'}(x', z, \mu'), \Delta, z\right) \geq \lambda_1^2\lambda_2^2 z_n^\beta\left(M^v(x, z, \delta_{\bar\theta}), \Delta, z\right) \qquad (3.13)
$$

$$
z_n^\beta\left(M^v(x', z, \delta_{\bar\theta}), \Delta, z\right) \geq \lambda_1^2\lambda_2^2 z_n^\beta\left(M^v(x, z, \mu), \Delta, z\right) \qquad (3.14)
$$

and

$$
z_n^\beta\left(M^v(x', z, \delta_{\bar\theta}), \Delta, z\right) \geq \lambda_1^2\lambda_2^2 z_n^\beta\left(M^v(x, z, \delta_{\bar\theta}), \Delta, z\right). \qquad (3.15)
$$

Therefore, for $\bar\theta = \Delta$, $\bar\theta' = \Delta$ it follows that

$$
\begin{aligned}
&\Pi^{v'}\left(\mu', \bar\theta', x', z_n^\beta\right) \\
&\quad = \int_{E_0^1}\int_E\int_{E_0^0} z_n^\beta(\delta_{\theta'}, \theta', z) P_0^{v'}(\theta, d\theta') r(x', \theta, v', z)\eta(dz)\mu'(d\theta) \\
&\qquad + \int_{E_0^1}\int_E z_n^\beta\left(M^{v'}(x', z, \mu'), \Delta, z\right) P^v(\theta, E_0^1) r(x', \theta, v', z)\eta(dz)\mu'(d\theta) \\
&\quad \geq \lambda_1\lambda_2 \int_{E_0^1}\int_E\int_{E_0^0} z_n^\beta(\delta_{\theta'}, \theta', z) P_0^v(\theta, d\theta') r(x, \theta, v, z)\eta(dz)\mu(d\theta)
\end{aligned}
$$

$$
\begin{aligned}
&+ \lambda_1^3\lambda_2^3 \int_{E_0^1}\int_E z_n^\beta \left(M^v(x,z,\mu),\Delta,z\right) P^v(\theta,E_0^1) r(x,\theta,v,z)\eta(dz)\mu'(d\theta)\\
&\geq \lambda_1^3\lambda_2^3 \Pi^v(\mu,\bar\theta,x,z_n^\beta).
\end{aligned}
$$

If $\bar\theta = \Delta$ and $\bar\theta' \in E_0^0$ then it follows from (3.14) that

$$
\begin{aligned}
&\Pi^{v'}\left(\mu',\bar\theta',x',z_n^\beta\right)\\
&\quad= \int_E\int_{E_0^0} z_n^\beta(\delta_\theta,\theta,z) P_0^{v'}(\bar\theta',d\theta) r(x',\theta',v',z)\eta(dz)\\
&\qquad+ \int_E z_n^\beta\left(M^{v'}(x',z,\delta_{\bar\theta'}),\Delta,z\right) P_0^{v'}(\bar\theta',E_0^1) r(x',\bar\theta',v',z)\eta(dz)\\
&\quad\geq \lambda_1\lambda_2 \int_{E_0^1}\int_E\int_{E_0^0} z_n^\beta(\delta_{\theta'},\theta',z) P_0^v(\theta,d\theta') r(x,\theta,v,z)\eta(dz)\mu(d\theta)\\
&\qquad+ \lambda_1^3\lambda_2^3 \int_{E_0^1}\int_E z_n^\beta\left(M^v(x,z,\mu),\Delta,z\right) P_0^v(\theta,E_0^1) r(x,\theta,v,z)\eta(dz)\mu(d\theta)\\
&\quad\geq \lambda_1^3\lambda_2^3 \Pi^v(\mu,\bar\theta,x,z_n^\beta).
\end{aligned}
$$

For the cases $\bar\theta \in E_0^0$, $\bar\theta' = \Delta$ or $\bar\theta \in E_0^0$, $\bar\theta' \in E_0^0$ using (3.13) and (3.15) respectively, the same inequalities are obtained. Thus for each $\bar\theta, \bar\theta' \in E_0^0 \cup \{\Delta\}$, $x, x' \in E$, $v, v' \in U$, $\mu, \mu' \in \mathcal{P}(E_0^1)$, the following inequality is satisfied

$$
\Pi^{v'}(\mu',\bar\theta',x',z_n^\beta) \geq \lambda_1^3\lambda_2^3 \Pi^v(\mu,\bar\theta,x,z_n^\beta). \tag{3.16}
$$

From (3.10) it follows that

$$
w_{n+1}^\beta(\mu,\bar\theta,x) - w_{n+1}^\beta(\mu',\bar\theta',x') \leq \|c\| + \beta(1-\lambda_1^3\lambda_2^3)\Pi^v(\mu,\bar\theta,x,z_n^\beta)
$$

so

$$
z_{n+1}^\beta(\mu,\bar\theta,x) \leq \|c\| + (1-\lambda_1^3\lambda_2^3)\Pi^v(\mu,\bar\theta,x,z_n^\beta) \tag{3.17}
$$

Since

$$
\|z_1^\beta\| = \sup_{\mu\in\mathcal{P}(E_0^0)}\ \sup_{\bar\theta\in E_0^1\cup\{\Delta\}}\ \sup_{x\in E} |z_1^\beta(\mu,\bar\theta,z)| \leq \|c\| \leq \frac{\|c\|}{\lambda_1^3\lambda_2^3}
$$

it follows by induction and (3.17) that

$$
\|z_n^\beta\| \leq \frac{\|c\|}{\lambda_1^3\lambda_2^3}
$$

from which (3.9) follows. $\square$

Let λ be given by

$$\lambda(x, x', \mu, \mu') = \inf_{y \in E} \inf_{v \in U} \inf_{A \in \mathcal{B}(E_0)} \frac{\displaystyle\int_{E_0} P_0^v(\theta, A) r(x', \theta, v, y) \mu'(d\theta)}{\displaystyle\int_{E_0} P_0^v(\theta, A) r(x, \theta, v, y) \mu'(d\theta)}. \tag{3.18}$$

If (3.17) and (3.18) are assumed, then it follows that $\lambda > 0$.

Proposition 2 *If* (3.7) *and* (3.8) *are satisfied and* Π^v *satisfies the Feller property, then for* $\mu, \mu' \in \mathcal{P}(E_0^0)$ *and* $x, x' \in E$

$$\begin{aligned} z^\beta(\mu, \Delta, x) &- z^\beta(\mu', \Delta, x') \\ &\leq \sup_{v \in U} \left[\int_{E_0^0} c(\theta, x, v) \mu(d\theta) - \int_{E_0^0} c(\theta, x', v) \mu'(d\theta) \right] \\ &\quad + (1 - \lambda^2(x, x', \mu, \mu') \lambda(x', x, \mu', \mu)) \frac{\|c\|}{\lambda_1^3 \lambda_2^3} \end{aligned} \tag{3.19}$$

and for $\bar{\theta}, \bar{\theta}' \in E_0^1$

$$\begin{aligned} z^\beta(\mu, \bar{\theta}, x) - z^\beta(\mu', \bar{\theta}', z') &\leq \sup_{v \in U} \left[c(\bar{\theta}, x, v) - c(\bar{\theta}', x', v) \right] \\ &\quad + (1 - \lambda^2(x, x', \delta_{\bar{\theta}}, \delta_{\bar{\theta}'}) \lambda(x', x, \delta_{\bar{\theta}'}, \delta_{\bar{\theta}})) \frac{\|c\|}{\lambda_1^3 \lambda_2^3} \end{aligned} \tag{3.20}$$

Proof. Similar to the methods of the proof of Proposition 1 (cf. (3.11)) it follows that

$$M^v(x', z, \mu')(A) \geq \lambda(x, x', \mu, \mu') \lambda(x', x, \mu', \mu) M^v(x, z, \mu)(A).$$

If $\lambda(x, x', \mu, \mu') = \lambda(x', x, \mu', \mu) = 1$ then $z^\beta(\mu, \Delta, x) = z^\beta(\mu', \Delta, x')$ and (3.19) is satisfied. Assume that $\lambda(x, x', \mu, \mu') \lambda(x', x, \mu', \mu) < 1$. Let m be defined as

$$m(x, x', z, v, \mu, \mu')(A) =$$

$$= \frac{1}{1-\lambda(x,x',\mu,\mu')\lambda(x',x,\mu',\mu)} \times \left[M^v(x',z,\mu')(A) - \lambda(x,x',\mu,\mu')\lambda(x',x,\mu',\mu)M^v(x,z,\mu)(A)\right].$$

By the nonnegativity and concavity of $z^\beta(\cdot,\Delta,x)$ it follows that

$$\begin{aligned}
& z^\beta(M^v(x',z,\mu'),\Delta,z) \\
&\quad \geq \lambda(x,x',\mu,\mu')\lambda(x',x,\mu',\mu)z^\beta(M^v(x,z,\mu),\Delta,z) \\
&\qquad + (1-\lambda(x,x',\mu,\mu')\lambda(x',x,\mu',\mu))z^\beta(m(x,x',z,v,\mu,\mu'),\Delta,z) \\
&\quad \geq \lambda(x,x',\mu,\mu')\lambda(x',x,\mu',\mu)z^\beta(M^v(x,z,\mu),\Delta,z).
\end{aligned}$$

Thus,

$$\begin{aligned}
& \Pi^v\left(\mu',\Delta,x',z^\beta\right) \\
&\quad = \int_{E_0^1}\int_E\int_{E_0^0} z^\beta(\delta_{\theta'},\theta',z)P_0^v(\theta,d\theta')r(x',\theta,v,z)\eta(dz)\mu'(d\theta) \\
&\qquad + \int_{E_0^1}\int_E z^\beta\left(M^v(x',z,\mu'),\Delta,z\right)P_0^v(\theta,E_0^1)r(x',\theta,v,z)\eta(dz)\mu'(d\theta) \\
&\quad \geq \lambda(x,x',\mu,\mu') \\
&\qquad \times \int_{E_0^1}\int_E\int_{E_0^0} z^\beta(\delta_{\theta'},\theta',z)P_0^v(\theta,d\theta')r(x,\theta,v,z)\eta(dz)\mu(d\theta) \\
&\qquad + \lambda^2(x,x',\mu,\mu')\lambda(x',x,\mu',\mu) \\
&\qquad \times \int_{E_0^1}\int_E z^\beta\left(M^v(x,z,\mu),\Delta,z\right)P_0^v(\theta,E_0^1)r(x,\theta,v,z)\eta(dz)\mu(d\theta) \\
&\quad \geq \lambda^2(x,x',\mu,\mu')\lambda(x',x,\mu',\mu)\Pi^v(\mu,\Delta,x,z^\beta).
\end{aligned} \tag{3.21}$$

From (3.1)

$$\begin{aligned}
& z^\beta(\mu,\bar\theta,x) - \beta \inf_{(\mu',\bar\theta',x')\in\mathcal{P}(E_0^1)\times E_0^0\cup\{\Delta\}\times E} z^\beta(\mu',\bar\theta',x') \\
&\qquad = \inf_{v\in U}\left[\bar c(\mu,\bar\theta,x,v) + \beta\Pi^v(\mu,\bar\theta,x,z^\beta)\right]
\end{aligned} \tag{3.22}$$

so

$$z^\beta(\mu,\Delta,x) - z^\beta(\mu',\Delta,x')$$

$$\leq \sup_{v\in U}\left[\int_{E_0^0} c(\theta,x,v)\mu(d\theta) - \int_{E_0^0} c(\theta,x',v)\mu'(d\theta)\right.$$
$$\left. + \beta\left[\Pi^v(\mu,\Delta,x,z^\beta) - \Pi^v(\mu',\Delta,x',z^\beta)\right]\right]$$
$$\leq \sup_{v\in U}\left[\int_{E_0^0} c(\theta,x,v)\mu(d\theta) - \int_{E_0^0} c(\theta,x',v)\mu'(d\theta)\right.$$
$$\left. + \beta\left(1-\lambda^2(x,x',\mu,\mu')\lambda(x',x,\mu',\mu)\right)\Pi^v(\mu,\Delta,x,z^\beta)\right].$$

Using (3.9), it follows that (3.19) is satisfied.

If $\bar\theta,\bar\theta'\in E_0^1$, then analogous to the above it follows that

$$M^v(x',z,\delta_{\bar\theta'})(A) \geq \lambda(x,x',\delta_{\bar\theta},\delta_{\bar\theta'})\lambda(x',x,\delta_{\bar\theta'},\delta_{\bar\theta})M^v(x,z,\delta_{\bar\theta})(A)$$

for each $A\in\mathcal{B}(E_0^1)$ and by analogy to (3.21)

$$\Pi^v(\mu',\delta_{\bar\theta'},x',z^\beta) \geq \lambda^2(x,x',\delta_{\bar\theta},\delta_{\bar\theta'})\lambda(x',x,\delta_{\bar\theta'},\delta_{\bar\theta})\Pi^v(\mu,\delta_{\bar\theta'},x,z^\beta)$$

from which (3.20) follows. □

The following theorem is the main result of this section.

Theorem 2 *If* (3.7) *and* (3.8) *are satisfied, the transition operator* Π^v *satisfies the Feller property and the function*

$$\lambda : E\times E\times\mathcal{P}(E_0^1)\times\mathcal{P}(E_0^1)\to\mathbb{R} \tag{3.23}$$

given by (3.18) *is continuous, then there are a function* $w\in C(\mathcal{P}(E_0^1)\times E_0^0\cup\{\Delta\}\times E)$ *and a constant* γ *such that*

$$w(\mu,\bar\theta,x)+\gamma = \inf_{v\in U}\left[\bar c(\mu,\bar\theta,x,v)+\Pi^v(\mu,\bar\theta,x,w)\right]. \tag{3.24}$$

Furthermore,

$$\gamma = \inf_{(v_n,\,n\in\mathbb{N})} J_{\xi,\bar\theta,x}((v_n,\,n\in\mathbb{N})) \tag{3.25}$$

and the control $v_n = u(\pi_n,\bar\theta_n,X_n)$ *for* $n\in\mathbb{N}$ *is optimal, where* u *is a Borel measurable for which the infimum* (3.24) *is achieved.*

Proof. By Propositions 1 and 2, the family of functions $(z^\beta,\ \beta \in (0,1))$ are bounded and continuous uniformly in $\beta \in (0,1)$. By Theorem 8.33 of [16] for a suitable subsequence $(z^{\beta_n},\ n \in \mathbb{N})$ with $\beta_n \to 1$ as $n \to \infty$, there is a $w \in C(\mathcal{P}(E_0^1) \times E_0^0 \cup \{\Delta\} \times E)$ such that $z^{\beta_n} \to w$ as $n \to \infty$ uniformly on compact subsets. Since

$$0 \leq (1-\beta) \inf_{(\mu',\theta',x') \in \mathcal{P}(E_0^1) \times E_0^0 \cup \{\Delta\} \times E} z^\beta(\mu',\theta',x') \leq \|c\|$$

there is a constant γ such that for a suitable subsequence $(z^{\beta_{n_k}},\ k \in \mathbb{N})$ it follows that

$$\lim_{k \to \infty} (1-\beta_{n_k}) \inf_{(\mu',\theta',x') \in \mathcal{P}(E_0^1) \times E_0^0 \cup \{\Delta\} \times E} z^{\beta_{n_k}}(\mu',\bar{\theta}',x') = \gamma.$$

Letting $\beta_{n_k} \to 1$ as $k \to \infty$ in (3.22) verifies the equality (3.24). Since the map $z^\beta(\cdot,\Delta,x) : \mathcal{P}(E_0^1) \to \mathbb{R}$ is concave it follows that the map $w(\cdot,\Delta,x) : \mathcal{P}(E_0^1) \to \mathbb{R}$ is also concave. The remaining part of the proof is standard for the average cost per unit time control problems (cf. [8]). $\square$

Remark 3 The continuity of λ given by (3.18) is satisfied if, assuming (3.7) and (3.8), the density r and the density $P_0^v(\theta,\cdot)$, with respect to a fixed measure $P_0^{v'}(\theta',\cdot)$, are continuous functions of their arguments.

Remark 4 Because of the assumptions (3.7), (3.8) and continuity of (3.18) the applicability of Theorem 1 is practically limited to compact spaces E_0 and E. There is a possibility to extend Theorem 1 to locally compact spaces E_0 and E using splitting arguments for Markov processes satisfying a minorization property with some suitable, uniform in the control, Lapunov type conditions. Such an approach is based on an adaptation of the recent results by V.Borkar (see [3] and [4]). It is necessary to consider a more general class of admissible controls. Consider a new probability measure P^0 such that the restrictions $P^0_{|n}$ and $P_{|n}$ of P^0 and P to $\mathcal{F}_n$ satisfy for $n = 1, 2, \ldots$

$$P_{|n}(d\omega) = \prod_{k=0}^{n-1} \frac{r(X_k,\theta_k,v_k,X_{k+1})}{r(X_k,\hat{\theta},v_k,X_{k+1})} P^0_{|n}(d\omega)$$

with $\hat{\theta}$ being a fixed element of E_0. As in [3] write that $\theta_{n+1} = F(\theta_n, \xi_n, v_n)$ with a suitably defined function F and a sequence of i.i.d. random variables (ξ_n). A more general class of controls is considered (the so called wide sense admissible controls) which consists of U valued random variables (v_n) such that the family $\{v_m, X_m, m \leq n\}$ is under P^0, independent of $\{\theta_0, \xi_m, X_m, m > n\}$. Instead of (3.7), (3.8) and the continuity of (3.18) it is assumed that there exist functions $V, h \in C(E_0 \times E)$ satisfying $h \geq 1$ with limits at infinity equal to infinity, a constant $C > 0$ and a ball $B \subset E_0 \times E$ such that for any wide sense admissible control (v_n)

$$E\left[V(\theta_{n+1}, X_{n+1}) | \mathcal{F}_n\right] - V(\theta_n, X_n) \leq -h(\theta_n, X_n) + C I_B(\theta_n, X_n)$$

and $\limsup_{n\to\infty} \frac{E[V(\theta_n, X_n)]}{n} = 0$. Then under continuity and positivity of the transition densities of the processes (θ_n) and (X_n) (in analogy to [4]) the existence of continuous (nonnecassary bounded) function w and constant γ are obtained for which (3.24) and (3.25) are satisfied.

To construct a selfoptimizing adaptive strategy a number of continuity results and a finite class of nearly optimal controls require compactness of E_0 and E. Consequently the proof of Theorem 1 in this paper is based on concavity arguments of [18] and [19], which are more convenient to show the above mentioned properties.

4 Parameter Continuity of the Optimal Average Cost per Unit Time

Now it is assumed that the transition probabilities P_0 and P depend on a parameter $\alpha \in \mathcal{A}$ where $\mathcal{A}$ is a compact metric space. The transition operators are $P_0^{v\alpha}(\theta, d\theta') = p_0^{v\alpha}(\theta, \theta')\delta(d\theta')$ and $P_\alpha^{\theta,v}(x, dz) = r_\alpha(x, \theta, v, z)\eta(dz)$ where δ is a probability measure on E_0. It is assumed that

$$p_0 : U \times \mathcal{A} \times E_0 \times E_0 \to \mathbb{R}_+$$

and

$$r : \mathcal{A} \times E \times E_0 \times U \times E \to \mathbb{R}_+$$

are continuous and E_0 and E are compact. By analogy with (3.7) and (3.8) it is assumed that

$$\bar{\lambda}_1 = \inf_{\alpha,\alpha' \in \mathcal{A}} \inf_{x,x',z \in E} \inf_{\theta,\theta' \in E_0} \inf_{v,v' \in U} \frac{r_{\alpha'}(x',\theta',v',z)}{r_\alpha(x,\theta,v,z)} > 0 \tag{4.1}$$

and

$$\bar{\lambda}_2 = \inf_{\alpha,\alpha' \in \mathcal{A}} \inf_{\theta,\theta',\theta'' \in E_0} \inf_{v,v' \in U} \frac{p_0^{v'\alpha'}(\theta',\theta'')}{p_0^{v\alpha}(\theta,\theta'')} > 0. \tag{4.2}$$

Remark 5 Notice first that in the example from section 1 the set of unknown parameters $\mathcal{A}$ consists of the transition matrices for (θ_n). Since we require that the entries of the transition matrix are positive, to have $\mathcal{A}$ compact we shall assume that there is a positive δ such that all entries of the transition matrix are greater or equal to δ. Assuming that the prices of the securities are projected to a large interval $[-M, M]$, under positivity and continuity of the density of w_n, we have (4.1) and (4.2) satisfied.

By analogy to (3.18) let $\bar{\lambda}$ be defined as

$$\bar{\lambda}(\alpha,\alpha',x,x',\mu,\mu') = \inf_{z \in E} \inf_{v \in U} \inf_{\theta' \in E_0} \frac{\displaystyle\int_{E_0} p_0^{v\alpha'}(\theta,\theta') r_\alpha(x',\theta,v,z)\mu'(d\theta)}{\displaystyle\int_{E_0} p_0^{v\alpha}(\theta,\theta') r_\alpha(x,\theta,v,z)\mu(d\theta)}. \tag{4.3}$$

It can be shown that by the assumptions on the densities p and r and (4.1) and (4.2)

$$\bar{\lambda} : \mathcal{A} \times \mathcal{A} \times E \times E \times \mathcal{P}(E_0) \times \mathcal{P}(E_0) \to \mathcal{R}$$

is continuous. Furthermore, if $f : \mathcal{P}(E_0^0) \times E_0^0 \cup \{\Delta\} \times E \to \mathbb{R}$ is bounded and continuous then the transition operator Π has the value

$$\begin{aligned}
&\Pi^{v\alpha}(\mu,\bar{\theta},x,f) \\
&\quad = 1_{E_0^0}(\bar{\theta}) \Bigg[\int_E \int_{E_0^0} f(\delta_\theta,\theta,z) p_0^{v\alpha}(\bar{\theta},\theta)\delta(d\theta) r_\alpha(x,\bar{\theta},v,z)\eta(dz) \\
&\qquad\qquad + \int_E f(M^{v\alpha}(x,z,\delta_\theta),\Delta,z) P_0^{v\alpha}(\bar{\theta},E_0^1) r_\alpha(x,\bar{\theta},v,z)\eta(dz) \Bigg] \\
&\qquad + 1_\Delta(\bar{\theta}) \Bigg[\int_{E_0^1} \int_E \int_{E_0^0} f(\delta_{\theta'},\theta',z) p_0^{v\alpha}(\theta,\theta')\delta(d\theta') r_\alpha(x,\theta,v,z)\eta(dz)\mu(d\theta)
\end{aligned}$$

$$+\int_{E_0^1}\int_E f(M^{v\alpha}(x,z,\mu),\Delta,z)P_0^{v\alpha}(\theta,E_0^1)$$

$$\left.\times\, r_\alpha(x,\theta,v,z)\eta(dz)\mu(d\theta)\right] \tag{4.4}$$

where for each $A \in \mathcal{B}(E_0)$

$$M^{v\alpha}(x,z,\mu)(A) = \frac{\displaystyle\int_{E_0} P_0^{v\alpha}(\theta, A\cap E_0^1) r_\alpha(x,\theta,v,z)\mu(d\theta)}{\displaystyle\int_{E_0} P_0^{v\alpha}(\theta,E_0^1) r_\alpha(x,\theta,v,z)\mu(d\theta)} \tag{4.5}$$

is a continuous function of $(\mu,\bar\theta,x,v,\alpha) \in \mathcal{P}(E_0^1)\times E_0^0\cup\{\Delta\}\times E\times U\times\mathcal{A}$.

The following theorem gives the continuity of the optimal cost with respect to $\alpha \in \mathcal{A}$.

Theorem 3 *If* (4.1) *and* (4.2) *are satisfied, then for each $\alpha \in \mathcal{A}$ there exist a function $w_\alpha \in C(\mathcal{P}(E_0^1)\times E_0^0\cup\{\Delta\}\times E)$ and a constant γ_α such that*

$$w_\alpha(\mu,\bar\theta,x)+\gamma_\alpha = \inf_{v\in U}\left[\bar c(\mu,\bar\theta,x,v)+\Pi^{v\alpha}(\mu,\bar\theta,x,w_\alpha)\right]. \tag{4.6}$$

Furthermore

$$\gamma_\cdot : \mathcal{A}\to\mathbb{R} \tag{4.7}$$

is continuous. If for $\varepsilon>0$ there is a Borel measurable function $u_\alpha^\varepsilon : \mathcal{P}(E)\times E_0^0\cup\{\Delta\}\times E\to U$ such that

$$w_\alpha(\mu,\bar\theta,x)+\gamma_\alpha \ge \bar c(\mu,\bar\theta,x,u_\alpha^\varepsilon(\mu,\bar\theta,x))+\Pi^{u_\alpha^\varepsilon\alpha}(\mu,\bar\theta,x,w_\alpha)-\varepsilon \tag{4.8}$$

for all $(\mu,\bar\theta,x)\in\mathcal{P}(E)\times E_0^0\cup\{\Delta\}\times E$, then the control $v_n = u_\alpha^\varepsilon(\pi_n^{\alpha'},\bar\theta_n,X_n)$ where $\pi_n^{\alpha'} = \bar M^{v_{n-1},\bar\theta_n}(X_{n-1},X_n,\pi_{n-1}^{\alpha'})$ is 2ε optimal for the cost functional $J_{\xi,\bar\theta,x}^{\alpha'}$ given by (2.6) *where the densities of the transition operators for $(\theta_n,\, n\in\mathbb{N})$ and $(X_n,\, n\in\mathbb{N})$ depend on α' for all α' in a neighborhood of α.*

Proof. Initially note that for fixed $\alpha\in\mathcal{A}$ all of the assumptions of Theorem 2 are satisfied. This fact implies the existence of a function $w_\alpha \ge 0$ and a constant γ_α such that (4.6) is satisfied. Furthermore, by Propositions 1 and 2 letting $\beta \uparrow 1$ in (3.3), (3.13) and (3.20) the following inequalities are satisfied

$$\sup_{\alpha\in\mathcal{A}} \sup_{\mu\in\mathcal{P}(E_0^1)} \sup_{\bar\theta\in E_0^0\cup\{\Delta\}} \sup_{x\in E} w_\alpha(\mu,\bar\theta,x) \leq \frac{\|c\|}{\bar\lambda_1^3\bar\lambda_2^3} \tag{4.9}$$

for $\mu,\mu'\in\mathcal{P}(E_0^1)$, $x,x'\in E$

$$\begin{aligned} w_\alpha(\mu,\Delta,x) &- w_\alpha(\mu',\Delta,x') \\ &\leq \sup_{v\in U}\left[\int_{E_0^0} c(\theta,x,v)\mu(d\theta) - \int_{E_0^0} c(\theta,x',v)\mu'(d\theta)\right] \\ &\quad + \left(1-\bar\lambda^2(\alpha,\alpha,x,x',\mu,\mu')\bar\lambda(\alpha,\alpha,x',x,\mu',\mu)\right)\frac{\|c\|}{\bar\lambda_1^3\bar\lambda_2^3} \end{aligned} \tag{4.10}$$

and for $\bar\theta,\bar\theta'\in E_0^1$,

$$\begin{aligned} w_\alpha(\mu,\bar\theta,x) &- w_\alpha(\mu',\bar\theta',x') \\ &\leq \sup_{v\in U}\left[c(\bar\theta,x,v) - c(\bar\theta',x',v)\right] \\ &\quad + \left(1-\bar\lambda^2(\alpha,\alpha,x,x',\delta_{\bar\theta},\delta_{\bar\theta'})\bar\lambda(\alpha,\alpha,x',x,\delta_{\bar\theta'},\delta_{\bar\theta})\right)\frac{\|c\|}{\bar\lambda_1^3\bar\lambda_2^3}. \end{aligned} \tag{4.11}$$

Thus the family of functions $(w_\alpha,\ \alpha\in\mathcal{A})$ is bounded and continuous uniformly in $\alpha\in\mathcal{A}$.

Assume that $(\alpha_n,\ n\in\mathbb{N})$ is a sequence from $\mathcal{A}$ such that $\alpha_n\to\alpha$ and $\gamma_{\alpha_n}\to\bar\gamma_\alpha$ as $n\to\infty$. There is a subsequence $(w_{\alpha_{n_k}},\ k\in\mathbb{N})$ such that this subsequence converges uniformly to $\bar w_\alpha$ so that $(\bar w_\alpha,\bar\gamma_\alpha)$ is a solution to the equation

$$\bar w_\alpha(\mu,\bar\theta,x) + \bar\gamma_\alpha = \inf_{v\in U}\left[\bar c(\mu,\bar\theta,x,v) + \Pi^{v\alpha}(\mu,\bar\theta,x,\bar w_\alpha)\right]$$

for $(\mu,\bar\theta,x)\in\mathcal{P}(E_0^1)\times E_0^0\cup\{\Delta\}\times E$. Thus, by the proof of Theorem 2, $\bar\gamma_\alpha$ is the optimal value of the cost functional $J^\alpha_{\xi,\bar\theta,x}$ and so $\bar\gamma_\alpha=\gamma_\alpha$ and the map $\gamma_\cdot:\mathcal{A}\to\mathbb{R}$ is continuous.

To prove the last part of the theorem, note that

$$M^{v\alpha}(x,z,\mu) \geq \bar\lambda(\alpha,\alpha',x,x,\mu,\mu)\bar\lambda(\alpha',\alpha,x,x,\mu,\mu)M^{v\alpha'}(x,z,\mu)$$

and by the concavity of $w_\alpha(\cdot, \Delta, x)$, which follows from the concavity of z^β as in the proof of Proposition 2, the following inequality is satisfied

$$\Pi^{v\alpha}(\mu, \bar{\theta}, x, w_\alpha) \geq \hat{\lambda}(\alpha, \alpha')[\hat{\lambda}(\alpha', \alpha)]^2 \Pi^{v\alpha'}(\mu, \bar{\theta}, x, w_\alpha)$$

for $(\mu, \bar{\theta}, x) \in \mathcal{P}(E_0^1) \times E_0^0 \cup \{\Delta\} \times E$ with

$$\hat{\lambda}(\alpha, \alpha') = \inf_{x \in E} \inf_{\mu \in \mathcal{P}(E)} \bar{\lambda}(\alpha, \alpha', x, x, \mu, \mu).$$

Clearly the map $\hat{\lambda} : \mathcal{A} \times \mathcal{A} \to \mathbb{R}$ is continuous. Let $\mathcal{O}(\alpha)$ be a neighborhood of α such that for $\varepsilon > 0$ fixed

$$|\gamma_\alpha - \gamma_{\alpha'}| \leq \frac{\varepsilon}{2}$$

and

$$1 - \hat{\lambda}(\alpha, \alpha')[\hat{\lambda}(\alpha', \alpha)]^2 \leq \frac{\varepsilon}{2\|c\|} \bar{\lambda}_1^3 \bar{\lambda}_2^3$$

for each $\alpha' \in \mathcal{O}(\alpha)$.

From (4.8) using (4.9) it follows for $\alpha' \in \mathcal{O}(\alpha)$ that

$$\begin{aligned} w_\alpha(\mu, \bar{\theta}, x) + \gamma_{\alpha'} \geq\ & \bar{c}(\mu, \bar{\theta}, x, u_\alpha^\varepsilon(\mu, \bar{\theta}, x)) - |\gamma_\alpha - \gamma_{\alpha'}| \\ & + \hat{\lambda}(\alpha, \alpha')[\hat{\lambda}(\alpha', \alpha)]^2 \Pi^{u_\alpha^\varepsilon(\mu,\theta,x)\alpha'}(\mu, \bar{\theta}, x, w_\alpha) - \varepsilon \\ \geq\ & \bar{c}(\mu, \bar{\theta}, x, u_\alpha^\varepsilon(\mu, \bar{\theta}, x)) + \Pi^{u_\alpha^\varepsilon(\mu,\theta,x)\alpha'}(\mu, \bar{\theta}, x, w_\alpha) - \frac{3}{2}\varepsilon \\ & - \left(1 - \hat{\lambda}(\alpha, \alpha')[\hat{\lambda}(\alpha', \alpha)]^2\right) \Pi^{u_\alpha^\varepsilon(\mu,\theta,x)\alpha'}(\mu, \bar{\theta}, x, w_\alpha) \\ \geq\ & \bar{c}(\mu, \bar{\theta}, x, u_\alpha^\varepsilon(\mu, \bar{\theta}, x)) + \Pi^{u_\alpha^\varepsilon(\mu,\theta,x)\alpha'}(\mu, \bar{\theta}, x, w_\alpha) - 2\varepsilon. \end{aligned}$$

This inequality verifies the 2ε optimality for the control $u_\alpha^\varepsilon(\pi_n^{\alpha'}, \bar{\theta}_n, x_n)$ for the cost $J_{\xi,\bar{\theta},x}^{\alpha'}$ if $\alpha' \in \mathcal{O}(\alpha)$. □

Remark 6 The second part of Theorem 3 can be described as follows: an almost optimal control u_α for α that satisfies (4.8) is also a good control for $\alpha' \in \mathcal{A}$ in a neighborhood of α.

The results given so far are valid if $E_0^0 = \emptyset$ so there are only partial observations of $(\theta_n,\, n \in \mathbb{N})$ in the usual sense. However, for the subsequent adaptive control results it is assumed that $E_0^0 \neq \emptyset$ and $P_0^{v\alpha}(\theta, E_0^0) > 0$ for each $(\theta, v, \alpha) \in E_0 \times U \times \mathcal{A}$. Let $\mathcal{U} = \{u \mid u : \mathcal{P}(E_0) \times E_0^0 \cup \{\Delta\} \times E \to U$ is Borel measurable$\}$. Let $D = \{(\delta_\theta, \theta) : \theta \in E_0^0\} \times E$. If $u \in \mathcal{U}$ and $v_n = u(\pi_n^\alpha, \bar{\theta}_n, X_n)$ for $n \in \mathbb{N}$ then by (4.2)

$$P\left((\pi_n^\alpha, \bar{\theta}_n, X_n) \in D \text{ i.o.}\right) = 1.$$

Furthermore, if

$$\tau_D^\alpha = \inf\left\{m \geq 0 : (\pi_m^\alpha, \bar{\theta}_m, X_m) \in D\right\}$$

then again using (4.2)

$$\sup_{u \in U} \sup_{\mu \in \mathcal{P}(E_0^1)} \sup_{\theta \in E_0^0 \cup \{\Delta\}} \sup_{x \in E} \sup_{\alpha \in \mathcal{A}} \mathbb{E}_{\mu\theta x}^{u\alpha}[\tau_D^\alpha] < \infty \tag{4.12}$$

Define the sequence of stopping times $(\tau_n,\, n \in \mathbb{N})$ as follows

$$\begin{aligned} \tau_1 &= \tau_D^\alpha \\ \tau_{n+1} &= \tau_n + \tau_D^\alpha \circ \Theta_{\tau_n} \end{aligned} \tag{4.13}$$

for $n = 1, 2, \ldots$ where Θ is the Markov shift operator. A Markov process is induced from $(\tau_n,\, n \in \mathbb{N})$ which is used to define an invariant measure for the Markov process $((\pi_n^\alpha, \bar{\theta}_n, X_n),\, n \in \mathbb{N})$.

Theorem 4 *For each $\alpha \in \mathcal{A}$ and $u \in \mathcal{U}$ there is a unique invariant measure φ_α^u for the embedded Markov process $((\pi_{\tau_n}^\alpha, \bar{\theta}_{\tau_n}, X_{\tau_n}),\, n = 1, 2, \ldots)$. Furthermore, there is also a unique invariant measure Φ_α^u for the Markov process $((\pi_n^\alpha, \bar{\theta}_n, X_n),\, n \in \mathbb{N})$ which can be given by*

$$\Phi_\alpha^u(F) = \frac{\displaystyle\int_D \mathbb{E}_{\mu\bar{\theta}x}^{u\alpha}\left[\sum_{j=0}^{\tau_D^\alpha - 1} F(\pi_j^\alpha, \bar{\theta}_j, X_j)\right] \varphi_\alpha^u(d\mu, d\bar{\theta}, dx)}{\displaystyle\int_D \mathbb{E}_{\mu\bar{\theta}x}^{u\alpha}[\tau_D^\alpha] \varphi_\alpha^u(d\mu, d\bar{\theta}, dx)} \tag{4.14}$$

for each bounded Borel measurable function $F : \mathcal{P}(E_0^0) \times E_0^0 \cup \{\Delta\} \times E \to \mathbb{R}$.

Proof. Note that $\pi_{\tau_n}^\alpha = \delta_{\bar{\theta}_{\tau_n}}$. By (4.1) and (4.2) it follows that for $B \in \mathcal{B}(D)$ and $(\mu, \bar{\theta}, x) \in D$

$$\mathbb{E}_{\mu\bar{\theta}x}^{u\alpha}\left[(\pi_{\tau_1}^\alpha, \bar{\theta}_{\tau_1}, X_{\tau_1}) \in B\right] \geq \mathbb{E}_{\hat{\mu}\Delta\hat{x}}^{u\alpha}\left[(\pi_{\tau_D^\alpha}, \bar{\theta}_{\tau_D^\alpha}, X_{\tau_D^\alpha}) \in B\right] \bar{\lambda}_1 \bar{\lambda}_2$$

where $\hat{\mu} \in \mathcal{P}(E_0^1)$ and $\hat{x} \in E$ are fixed. By [8], the embedded Markov process is uniformly ergodic with a unique invariant measure Φ_α^u. It easily follows that Φ_α^u is invariant for the Markov process $((\pi_n, \bar{\theta}_n, X_n),\, n \in \mathbb{N})$. The uniqueness of the invariant measure Π follows from the fact that no matter where the process commences it enters D by (4.12) and the process has a unique minimal invariant set. $\square$

Corollary 1 *For $v_n = u(\pi_n^\alpha, \bar{\theta}_n, X_n)$ the following equality is satisfied*

$$\begin{aligned} &J^\alpha_{\xi,\bar{\theta},x}((v_n,\, n \in \mathbb{N})) \\ &\quad = \int_{\mathcal{P}(E_0)\times E_0^0\cup\{\Delta\}\times E} \int_{E_0^0} c(\theta, x, u(\mu, \bar{\theta}, x))\mu(d\theta)\Phi_\alpha^u(d\mu, d\bar{\theta}, dx). \end{aligned} \tag{4.15}$$

The following result is used for the construction of almost optimal adaptive controls.

Proposition 3 *Let $\alpha \in \mathcal{A}$ and $u \in \mathcal{U}$ be fixed and consider the control $v_n^{\alpha'} = u(\pi_n^{\alpha'}, \bar{\theta}_n, X_n)$ for $n \in \mathbb{N}$ and $\alpha' \in \mathcal{A}$ where the process $(\pi_n^{\alpha'},\, n \in \mathbb{N})$ is constructed using the recursive formula* (2.2) *with θ_n, X_n having transition probabilities depending on α. Then the quadruple $((\pi_n^\alpha, \pi_n^{\alpha'}, \bar{\theta}_n, X_n),\, n \in \mathbb{N})$ is a Markov process with the unique invariant measure $\Psi^u_{\alpha\alpha'}$ that converges to the measure $\Psi^u_{\alpha\alpha}$ in the weak* topology if $u \in \mathcal{U}_c = \{u \in \mathcal{U} : u \text{ is continuous}\}$ and $\alpha' \to \alpha$ where*

$$\Psi^u_{\alpha\alpha}(F) = \int_{\mathcal{P}(E_0)} \int_{E_0^0\cup\{\Delta\}} \int_E F(\mu, \mu, \bar{\theta}, x)\Phi_\alpha^u(d\mu, d\bar{\theta}, dx) \tag{4.16}$$

for each bounded, Borel measurable function $F : \mathcal{P}(E_0)\times\mathcal{P}(E_0)\times E_0^0\cup\{\Delta\}\times E \to \mathbb{R}$ and Φ_α^u is given by (4.14).

Furthermore, for $u \in \mathcal{U}_c$,

$$\lim_{\alpha'\to\alpha} J^\alpha_{\xi,\bar{\theta},x}((v_n^{\alpha'},\, n \in \mathbb{N})) = J^\alpha_{\xi,\bar{\theta},x}((v_n,\, n \in \mathbb{N})) \tag{4.17}$$

where $v_n = u(\pi_n^\alpha, \bar{\theta}_n, X_n)$.

Proof. By standard methods, it follows that the quadruple $((\pi_n^\alpha, \pi_n^{\alpha'}, \bar{\theta}_n, X_n),$ $n \in \mathbb{N})$ is a Markov process with transition operator $\bar{\Pi}^{u\alpha}_{\alpha'}$. Let $\bar{D} = \{(\delta_\theta, \delta_\theta, \theta) :$

$\theta \in E_0^0\} \times E$. It easily follows that the process $((\pi_n^\alpha, \pi_n^{\alpha'}, \bar\theta_n, X_n),\ n \in \mathbb{N})$ visits $\bar D$ infinitely often. Define a sequence of random times as follows

$$\bar\tau_D^{\alpha\alpha'} = \inf\{i \geq 0 : (\pi_i^\alpha, \pi_i^{\alpha'}, \bar\theta_i, X_i) \in D\}$$
$$\bar\tau_1 = \bar\tau_D^{\alpha\alpha'}$$

and for $n \geq 1$

$$\bar\tau_{n+1} = \bar\tau_n + \bar\tau_1 \circ \Theta_{\bar\tau_n}.$$

As in the proof of Theorem 4, it is shown that $((\pi_{\bar\tau_n}^\alpha, \pi_{\bar\tau_n}^{\alpha'}, \bar\theta_{\bar\tau_n}, X_{\bar\tau_n}),\ n \in \mathbb{N})$ is uniformly ergodic so $((\pi_n^\alpha, \pi_n^{\alpha'}, \bar\theta_n, X_n),\ n \in \mathbb{N})$ has a unique invariant measure $\Psi_{\alpha\alpha'}^u$. Let $(\alpha_n',\ n \in \mathbb{N})$ be a sequence in $\mathcal{A}$ that converges to α. Since the sequence of measures $(\Psi_{\alpha\alpha_n'}^u,\ n \in \mathcal{N})$ is tight, which follows from the continuity of the transition densities p_0 and r_α and the control function u as well as (4.1) and (4.2), there is a measure Ψ such that the subsequence $(\Psi_{\alpha\alpha_{n_k}'}^u,\ k \in \mathcal{N})$ converges in the weak* topology to Ψ. By the continuity of the transition densities for each $F \in C(\mathcal{P}(E) \times \mathcal{P}(E) \times E_0^0 \cup \{\Delta\} \times E)$ and $(\alpha_n',\ n \in \mathbb{N})$ a sequence converging to α

$$\lim_{n \to \infty} \bar\Pi_{\alpha_n'}^{u\alpha} F = \bar\Pi_\alpha^{u\alpha} F$$

uniformly on compact subsets of $\mathcal{P}(E) \times \mathcal{P}(E) \times E_0^0 \cup \{\Delta\} \times E$.

By the Feller property of $\Pi_\alpha^{u\alpha}$ which follows by the continuity of $u \in \mathcal{U}_c$ the following inequality is satisfied

$$\begin{aligned} |\Psi(\bar\Pi_\alpha^{u\alpha} F) - \Psi(F)| \leq{}& |\Psi(\bar\Pi_\alpha^{u\alpha} F) - \Psi_{\alpha\alpha_{n_k}'}^u(\bar\Pi_\alpha^{u\alpha} F)| \\ &+ |\Psi_{\alpha\alpha_{n_k}'}^u(\bar\Pi_\alpha^{u\alpha} F) - \Psi_{\alpha\alpha_{n_k}'}^u(\bar\Pi_{\alpha_{n_k}'}^{u\alpha} F)|. \end{aligned}$$

Since the right hand side of the inequality tends to 0 as $k \to \infty$, it follows that Ψ is an invariant measure for $\bar\Pi_\alpha^{u\alpha}$. Since $\Psi_{\alpha\alpha}^u$ is the unique invariant measure for $\bar\Pi_\alpha^{u\alpha}$ it follows that $\Psi = \Psi_{\alpha\alpha}^u$ and the sequence $(\Psi_{\alpha\alpha_n'}^u,\ n \in \mathbb{N})$ converges in the weak* topology to $\Psi_{\alpha\alpha}^u$. By the form of the transition operator $\bar\Pi_\alpha^{u\alpha}$, it is clear that (4.16) is satisfied with Φ_α^u the unique invariant measure for $((\pi_n^\alpha, \bar\theta_n, X_n),\ n \in \mathbb{N})$ with the control $v_n = u(\pi_n^\alpha, \bar\theta_n, X_n)$. The convergence of the optimal costs (4.17) follows from (4.15) and (4.16) and the weak* convergence of $(\Psi_{\alpha\alpha_n'}^u,\ n \in \mathbb{N})$ to $\Psi_{\alpha\alpha}^u$. $\square$

5 Bayesian Adaptive Control and Identification

Now it is assumed that the unknown parameter α, that occurs in the transition operators P_0^α and P^α, is a random variable with the known probability law $\alpha_0 \in \mathcal{P}(\mathcal{A})$. The measure valued process $(\hat{\alpha}_n,\ n \in \mathbb{N})$ given by

$$\begin{aligned}\hat{\alpha}_0(B) &= \alpha_0(B) \\ \hat{\alpha}_n(B) &= P(\alpha^0 \in B \mid X^n, \bar{\theta}^n)\end{aligned} \tag{5.1}$$

for $B \in \mathcal{B}(\mathcal{A})$ and α^0 is a random variable with probability law α_0. The random variable $\hat{\alpha}_n$ describes the information of α^0 at time n.

The following lemma is easily verified.

Lemma 4 *For $n \in \mathbb{N}$ and $B \in \mathcal{B}(\mathcal{A})$*

$$\hat{\alpha}_{n+1}(B) = G^{v_n \theta_{n+1}}(X_n, X_{n+1}, (\pi_n^\alpha,\ \alpha \in \mathcal{A}), \hat{\alpha}_n)(B) \quad a.s.\ P \tag{5.2}$$

where

$$G^{v\Delta}(x, z, (\mu^\alpha,\ \alpha \in \mathcal{A}), \hat{\alpha})(B) = \frac{\displaystyle\int_B \int_{E_0} P_0^{v\alpha}(\theta, E_0^1) r_\alpha(x, \theta, v, z) \mu^\alpha(d\theta) \hat{\alpha}(d\alpha)}{\displaystyle\int_{\mathcal{A}} \int_{E_0} P^{v\alpha}(\theta, E_0^1) r_\alpha(x, \theta, v, z) \mu^\alpha(d\theta) \hat{\alpha}(d\alpha)} \tag{5.3}$$

and for $\bar{\theta} \in E_0^0$

$$G^{v\bar{\theta}}(x, z, (\mu^\alpha,\ \alpha \in \mathcal{A}), \hat{\alpha})(B) = \frac{\displaystyle\int_B \int_{E_0} p_0^{v\alpha}(\theta, \bar{\theta}) r_\alpha(x, \theta, v, z) \mu^\alpha(d\theta) \hat{\alpha}(d\alpha)}{\displaystyle\int_{\mathcal{A}} \int_{E_0} p^{v\alpha}(\theta, \bar{\theta}) r_\alpha(x, \theta, v, z) \mu^\alpha(d\theta) \hat{\alpha}(d\alpha)}. \tag{5.4}$$

For subsequent considerations, the following identifiability property is used.

(I) If for $\alpha, \alpha' \in \mathcal{A}$ the following equalities are satisfied

$$\begin{aligned} r^\alpha(x, \theta, v, z) &= r^{\alpha'}(x, \theta, v, z) \\ p_0^{v\alpha}(\theta, \theta') &= p_0^{v\alpha'}(\theta, \theta') \end{aligned}$$

and

$$P_0^{v\alpha}(\theta, E_0^1) = P_0^{v\alpha'}(\theta, E_0^1)$$

for all $x, z \in E$, $\theta \in E_0^0$, $\theta' \in E_0$, and $v \in U_f$ where U_f is a fixed finite nonempty subset of U, then $\alpha = \alpha'$.

Remark 7 In the case of the example from section 1 we clearly have that transition densities of $S_i(n)$ do not depend on the parameter α, which is the transition matrix of (θ_n). The identifiability property in this case means that knowing the transition probabilities within class E_0^0, which can be estimated, we should be allowed to identify the whole transition matrix. This is the case when we know that transition probabilities are certain functions of the transition probabilities within class E_0^0 e.g. we know that certain transition probabilites are equal, the other are two times smaller and so on. In the case, when we know entries of the transition matrix concerning the unobserved states and don't know other entries, the identifiability property is clearly satisfied.

Note that by the properties of the conditional expectation and the martingale convergence theorem, $\hat{\alpha}_n \Rightarrow \hat{\alpha}_\infty$ a.s. P where $\Rightarrow$ denotes the weak convergence of probability measures on $\mathcal{A}$ and for $B \in \mathcal{B}(\mathcal{A})$

$$\hat{\alpha}_\infty(B) = P(\alpha^0 \in B \mid X^\infty, \bar{\theta}_\infty) \tag{5.5}$$

and $X^\infty = \sigma(X_0, X_1, \ldots)$ and $\bar{\theta}^\infty = \sigma(\bar{\theta}_0, \bar{\theta}_1, \ldots)$. The (random) parameter α^0 is identified if $\hat{\alpha}_\infty = \delta_{\alpha^0}$ a.s. P.

Given a control $(v_n,\, n \in \mathbb{N})$, consider the sequence of random times $(\tau_n,\, n \in \mathbb{N})$ given in (4.13). Let $(\sigma_n,\, n \in \mathbb{N})$ be a sequence of random times such that $\sigma_{n+1} > \sigma_n$ a.s. for each $n \in \mathbb{N}$, $\sigma_n = \tau_{T_n}$ where T_n is an $\mathbb{N}$-valued random variable and

$$\lim_{t \to \infty} \frac{1}{t} \mathbb{E}_{\xi,\theta,x} \left[\sum_{j=0}^{\infty} 1_{\{\sigma_j \le t\}} \right] = 0. \tag{5.6}$$

At the random times $(\sigma_n,\, n \in \mathbb{N})$, the control $(v_n,\, n \in \mathbb{N})$ is modified to use a successive selection of the controls in the finite set U_f or a random selection with equal probabilities of these controls. The control strategy $(v_n,\, n \in \mathbb{N})$ with these modifications is called a *control with forcing*.

Proposition 4 *If* (I) *is satisfied, then for each control strategy with forcing* $\hat{\alpha}_\infty = \delta_{\alpha^0}$ *a.e.* P*, that is, the unknown random parameter is identified.*

Proof. This proof is an adaptation of the proof of Proposition 2 [7] as well as the proof of Proposition 1 [20]. By forcing and by (4.1) and (4.2), any point $(v, x, z, \bar{\theta}, (\delta_{\bar{\theta}'})_{\alpha \in \mathcal{A}})$, where $v \in U_f$, $x, z \in E$, $\bar{\theta} \in E_0^0 \cup \{\Delta\}$, $\bar{\theta}' \in E_0^0$, is frequent for the sequence $((v_n, X_n, X_{n+1}, \bar{\theta}_{n+1}, (\pi_n^\alpha)_{\alpha \in \mathcal{A}},\, n \in \mathbb{N})$ (that is, for each $\varepsilon > 0$ there is a subsequence $((v_{n_k}, X_{n_k}, X_{n_k+1}),\, k \in \mathbb{N})$ that is in the ball of radius ε with center (v, x, z), $\pi_{n_k}^\alpha = \delta_{\bar{\theta}''}$ with $\bar{\theta}'' \in E_0^0$ in an ε neighborhood of $\bar{\theta}'$, and for $\bar{\theta} = \Delta$, $\bar{\theta}_{n_k+1} = \Delta$ while for $\bar{\theta} \in E_0^0$, $\bar{\theta}_{n_k+1}$ is in the ε-ball centered at $\bar{\theta}$. Using the continuity of the transition densities (cf. [7], [18]) for α, α' in the support of $\hat{\alpha}_\infty$ the following equalities are satisfied

$$p_0^{v\alpha}(\bar{\theta}', \bar{\theta}) r_\alpha(x, \bar{\theta}', v, z) = p_0^{v\alpha'}(\bar{\theta}', \bar{\theta}) r_{\alpha'}(x, \bar{\theta}', v, z) \tag{5.7}$$

and

$$P_0^{v\alpha}(\bar{\theta}', E_0^1) r_\alpha(x, \bar{\theta}', v, z) = P_0^{v\alpha'}(\bar{\theta}', E_0^1) r_{\alpha'}(x, \bar{\theta}', v, z) \tag{5.8}$$

for each $v \in U_f$, $x, z \in E$, $\bar{\theta}' \in E_0^0$ and $\bar{\theta} \in E_0^0$. Integrating (5.9) and (5.10) with respect to z and using the measure η it follows that

$$p_0^{v\alpha}(\bar{\theta}', \bar{\theta}) = p_0^{v\alpha'}(\bar{\theta}', \bar{\theta}) \tag{5.9}$$

and

$$P_0^{v\alpha}(\bar{\theta}', E_0^1) = P_0^{v\alpha'}(\bar{\theta}', E_0^1). \tag{5.10}$$

Thus from (5.9), (5.10), (5.7), and (5.8) it follows that

$$r_\alpha(x, \bar{\theta}', v, z) = r_{\alpha'}(x, \bar{\theta}', v, z)$$

and by assumption (I), $\alpha = \alpha'$. This means that $\hat{\alpha}_\infty$ is a measure concentrated at the random point K. For each $B \in \mathcal{B}(\mathcal{A})$ there is the equality

$$\mathbb{E}[1_B(\alpha^0) 1_{B^c}(K)] = \mathbb{E}\left[1_{B^c}(K) \mathbb{E}[1_B(\alpha^0) \mid X^\infty, \bar{\theta}^\infty]\right] = \mathbb{E}[1_{B^c}(K) 1_B(K)] = 0.$$

Thus α^0 is concentrated (a.e. P) at the random point K. $\square$

6 Almost Self-Optimal Adaptive Controls

In this section it is assumed that for a fixed $\varepsilon > 0$ and $\alpha \in \mathcal{A}$ there is a continuous function $u_\alpha : \mathcal{P}(E_0') \times E_0^0 \cup \{\Delta\} \times E \to U$ such that for each $\mu, \bar{\theta}, x) \in \mathcal{P}(E_0') \times E_0^0 \cup \{\Delta\} \times E$ the following inequality is satisfied

$$w_\alpha(\mu, \bar{\theta}, x) + \gamma_\alpha \geq \bar{c}(\mu, \bar{\theta}, x, u_\alpha(\mu, \bar{\theta}, x)) + \Pi^{u_\alpha(\mu,\bar{\theta},x)\alpha}(\mu, \bar{\theta}, w_\alpha) - \varepsilon.$$

Remark 8 It is clear that there is a Borel measurable function u_α satisfying the above inequality even with $\varepsilon = 0$. We require above more - the function u_α should be continuous. In particular cases, if dimension of the $\mathcal{P}(E_0')$ is not too large we can construct continuous u_α by smoothing piecewise constant controls.

By Theorem 3 and Proposition 3, given $\varepsilon > 0$ there are a finite family of continuous functions $(u_{\alpha^j}, j = 1, \ldots, k)$ and a finite partition $(A_j, j = 1, \ldots, k)$ of $\mathcal{A}$ such that the control $u_{\alpha^j}(\pi_n^{\alpha^j}, \bar{\theta}_n, X_n)$ is 2ε optimal for $\alpha \in A_j$. Furthermore, the partition can be chosen so that the boundary of each A_i, $i = 1, \ldots, k$ has α_0 measure zero.

The model with the unknown random parameter α^0 is controlled at each time n by a randomized control $v_n^\alpha = u_{\alpha^i}(\pi_n^{\alpha^i}, \bar{\theta}_n, X_n)$ if $\alpha \in A_i$ with the measure $\hat{\alpha}_n$. This control is almost optimal.

Theorem 5 *Let $\varepsilon > 0$ be fixed. If the randomized control v_n^α with measure $\hat{\alpha}_n$ for $n \in \mathbb{N}$ is used and the identifiability condition* (I) *is satisfied, then this control is 2ε optimal.*

Proof. Some arguments of the proof of Theorem 1 [18] are used. The quadruple $((\hat{\alpha}_n, (\pi_n^\alpha)_{\alpha \in \mathcal{A}}, \bar{\theta}_n, X_n), n \in \mathbb{N})$ is a Markov process with transition operator $\hat{\Pi}^v$ which is defined by

$$\begin{aligned}
&\hat{\Pi}^v F(\hat{\alpha}, (\mu^\alpha)_{\alpha \in \mathcal{A}}, \Delta, x) \\
&\quad = \int_A \int_{E_0} \int_E F(G^{v\Delta}(x, z, (\mu^\alpha)_{\alpha \in \mathcal{A}}, \hat{\alpha}), (M^{v\alpha}(x, z, \mu^\alpha))_{\alpha \in \mathcal{A}}, \Delta, z) \\
&\qquad\qquad \times P_0^{v\alpha}(\theta, E_0^1) r_\alpha(x, \theta, v, z) \eta(dz) \mu^\alpha(d\theta) \hat{\alpha}(d\alpha)
\end{aligned}$$

$$+\int_A\int_{E_0}\int_E\int_{E_0^0} F(G^{v\bar\theta'}(x,z,(\mu^\alpha)_{\alpha\in\mathcal{A}},\hat\alpha),(\delta_{\bar\theta'})_{\alpha\in\mathcal{A}},\bar\theta',z)$$
$$\times P_0^{v\alpha}(\theta,d\bar\theta')r_\alpha(x,\theta,v,z)\eta(dz)\mu^\alpha(d\theta)\hat\alpha(d\alpha) \quad (6.1)$$

and for $\bar\theta\in E_0^0$

$$\begin{aligned}
&\hat\Pi^v F(\hat\alpha,(\mu^\alpha)_{\alpha\in\mathcal{A}},\bar\theta,x)\\
&\quad=\int_A\int_E F(G^{v\Delta}(x,z,(\mu^\alpha)_{\alpha\in\mathcal{A}},\hat\alpha),(M^{v\alpha}(x,z,\mu^\alpha))_{\alpha\in\mathcal{A}},\Delta,z)\\
&\qquad\qquad\times P_0^{v\alpha}(\bar\theta,E_0^1)r_\alpha(x,\bar\theta,v,z)\eta(dz)\hat\alpha(d\alpha)\\
&\qquad+\int_A\int_E F(G^{v\bar\theta'}(x,z,(\mu^\alpha)_{\alpha\in\mathcal{A}},\hat\alpha),(\delta_{\bar\theta'})_{\alpha\in\mathcal{A}},\bar\theta',z)\\
&\qquad\qquad\times P_0^{v\alpha}(\theta,d\bar\theta')r_\alpha(x,\bar\theta,v,z)\eta(dz)\mu^\alpha(d\theta)\hat\alpha(d\alpha)
\end{aligned}\quad (6.2)$$

for a bounded Borel measurable function $F:\mathcal{P}(\mathcal{A})\times\prod_{\alpha\in\mathcal{A}}\mathcal{P}(E_0)\times E_0^0\cup\{\Delta\}\times E\to\mathbb{R}$. Since the points of discontinuity of the controlled transition operator $\hat\Pi$, for F continuous, are boundary points of the partition (A_i, $i=1,\dots,n$), a subsequence formed from Césaro averages of the iterations of $\hat\Pi$ starting with the initial measure $\alpha=\alpha_0$ converges in the weak* topology to an invariant measure Ψ for the quadruple $((\hat\alpha_n,(\pi_n^\alpha)_{\alpha\in\mathcal{A}},\bar\theta_n,X_n),\ n\in\mathbb{N})$.

If $F\in C(\mathcal{P}(\mathcal{A}))$ then by the identifiability assumption for each $n\in\mathbb{N}$

$$\begin{aligned}
&\Psi\left(F,\prod_{\alpha\in\mathcal{A}}\mathcal{P}(E_0)\times E_0^0\cup\{\Delta\}\times E\right)\\
&\quad=\int_{\mathcal{P}(\mathcal{A})}\int_{\prod_{\alpha\in\mathcal{A}}\mathcal{P}(E_0)}\int_{E_0^0\cup\{\Delta\}}\int_E \mathbb{E}\left[F(\hat\alpha_n)\mid\hat\alpha,(\mu_\alpha)_{\alpha\in\mathcal{A}},\bar\theta,x\right]\\
&\qquad\qquad\times\Psi(d\hat\alpha,d(\mu_\alpha)_{\alpha\in\mathcal{A}},d\bar\theta,d\alpha)\\
&\quad\to\int_{\mathcal{P}(\mathcal{A})}\int_{\mathcal{A}} F(\delta_\alpha)\hat\alpha(d\alpha)\Psi\left(d\hat\alpha,\prod_{\alpha\in\mathcal{A}}\mathcal{P}(E_0)\times E_0^0\cup\{\Delta\}\times E\right)
\end{aligned}$$

as $n\to\infty$. Since the barycenter of $\Psi(\cdot,\prod_{\alpha\in\mathcal{A}}\mathcal{P}(E_0)\times E_0^0\cup\{\Delta\}\times E)$ is α_0, it follows that

$$\Psi\left(F,\prod_{\alpha\in\mathcal{A}}\mathcal{P}(E_0)\times E_0^0\cup\{\Delta\}\times E\right)=\int F(\delta_\alpha)\alpha_0(d\alpha). \quad (6.3)$$

Let $p(\cdot,\cdot)$ be a regular conditional probability of Ψ with respect to the σ-algebra $\mathcal{B}(\mathcal{P}(\mathcal{A}))\otimes\{\phi,\prod_{\alpha\in\mathcal{A}}\mathcal{P}(E_0)\times E_0^0\cup\{\Delta\}\times E\}$. For a bounded Borel measurable function $F:\mathcal{P}(\mathcal{A})\times\prod_{\alpha\in\mathcal{A}}\mathcal{P}(E_0)\times E_0^0\cup\{\Delta\}\times E\to\mathbb{R}$, there is the equality

$$\Psi(F)=\int_A\int_{\prod_{\alpha\in\mathcal{A}}\mathcal{P}(E_0)\times E_0^0\cup\{\Delta\}\times E}F(\delta_\alpha,\omega_2)p(\delta_\alpha,d\omega_2)\alpha_0(d\alpha). \tag{6.4}$$

In particular,

$$\Psi(F)=\Psi(\hat{\Pi}F)=\int_A\int_{\prod_{\alpha\in\mathcal{A}}\mathcal{P}(E_0)\times E_0^0\cup\{\Delta\}\times E}\Pi^{v^\alpha}F(\delta_\alpha,\omega_2)p(\delta_\alpha,d\omega_2)\alpha_0(d\alpha), \tag{6.5}$$

where $v^\alpha=u_{\alpha^i}(\mu^{\alpha^i},\bar{\theta},x)$ for $\alpha\in A_i$ and $\omega_2=((\mu^\alpha)_{\alpha\in\mathcal{A}},\bar{\theta},x)$. Choosing

$$F\left(\hat{\pi},(\mu^\alpha)_{\alpha\in\mathcal{A}},\bar{\theta},x\right)=F_1(\hat{\pi})F_2\left(\hat{\pi},(\mu^\alpha)_{\alpha\in\mathcal{A}},\bar{\theta},x\right)$$

by (6.4) and (6.5) and since F_1 is arbitrary the following equality is satisfied

$$\begin{aligned}\int_{\prod_{\alpha\in\mathcal{A}}\mathcal{P}(E_0)\times E_0^0\cup\{\Delta\}\times E}&F_2(\delta_\alpha,\omega_2)p(\delta_\alpha,d\omega_2)\\&=\int_{\prod_{\alpha\in\mathcal{A}}\mathcal{P}(E_0)\times E_0^0\cup\{\Delta\}\times E}\Pi^{v^\alpha}F_2(\delta_\alpha,\omega_2)p(\delta_\alpha,d\omega_2)\end{aligned} \tag{6.6}$$

for α_0 almost all α. Moreover, it can be shown that (6.6) is satisfied for each F_2 for α_0 almost all α. Now let $F_2(\delta_\alpha,(\mu^\alpha)_{\alpha\in\mathcal{A}},\bar{\theta},x)=f_2(\mu^\alpha,\mu^{\alpha^i},\bar{\theta},x)$ if $\alpha\in A^i$. By (6.6) it follows that

$$p\left(\delta_\alpha,d\mu^\alpha,d\mu^{\alpha^i},\prod_{\substack{\alpha'\in\mathcal{A}\\\alpha'\neq\alpha,\alpha^i}}P(E_0),d\bar{\theta},dx\right)$$

is an invariant measure for $((\pi_n^\alpha,\pi_n^{\alpha^i},\bar{\theta}_n,X_n),\,n\in\mathbb{N})$ which by Proposition 3 coincides with $\Psi^{u_{\alpha^i}}_{\alpha\alpha^i}$.

Thus, for $\hat{F}(\delta_\alpha,(\mu^\alpha)_{\alpha\in\mathcal{A}},\bar{\theta},x)=\bar{c}(\mu_\alpha,\bar{\theta},x,u_{\alpha^i}(\mu^{\alpha^i},\bar{\theta},x))$ the following equality is satisfied

$$\Psi(\hat{F})=\sum_{i=1}^k\int_{A_i}\int_{\mathcal{P}(E_0)\times\mathcal{P}(E_0)}\int_{E_0^0\cup\{\Delta\}}\int_E\bar{c}(\mu_\alpha,\bar{\theta},x,u_{\alpha^i}(\mu^{\alpha^i},\bar{\theta},x))$$

$$\times \Psi^{u_{\alpha^i}}_{\alpha\alpha^i}(d\mu_\alpha, d\mu_{\alpha^i}, d\bar{\theta}, dx)\alpha_0(d\alpha) \tag{6.7}$$

and by 2ε optimality of the control $u_{\alpha^i}(\pi_n^{\alpha^i}, \bar{\theta}_n, X_n)$, $n \in \mathbb{N})$ for $\alpha \in A_i$ it follows that

$$\Psi(\hat{F}) \leq \int_{\mathcal{A}} \gamma_\alpha \alpha_0(d\alpha) - 2\varepsilon$$

which implies that the randomized control $(v_n^\alpha,\ n \in \mathbb{N})$ with the measure $\hat{\alpha}_n$ at time n is 2ε optimal because the value of (6.7) is the same for any weak* limit of Cesaro averages of the iterations of $\hat{\Pi}$. □

Remark 9 To construct nearly optimal adaptive control we required two important factors: complete observation of (θ_n) in an ergodic set E_0^0 and then identifiability property (I). The first factor guarantees nice ergodic properties of (π_n). The other alIows to identify α^0. In the case of complete observation one could characterize the support of the measure $\hat{\alpha}_\infty$ following [5]. Unfortunately, with partial observation the description (5.2) of $\hat{\alpha}_{n+1}$ involves family $\{\pi_n^\alpha,\ \alpha \in \mathcal{A}\}$, so the methods of [5] cannot be used.

References

[1] T.R. Bielecki, S.R. Pliska, Risk Sensitive Dynamic Asset Management, *JAMO* 39 (1999), 337-360.

[2] T.R. Bielecki, M. Rutkowski, *Credit Risk: Modelling, Valuyation and Hedging.* Springer, Berlin, Heidelberg 2002.

[3] V.S. Borkar, Dynamic programming for ergodic control with partial observation, *Stoch. Proc. Appl.* 103 (2003), 293-310.

[4] V.S. Borkar and A. Budhiraja, Further Remark on Dynamic Programming for Partially Observed Markov Processes, preprint.

[5] V.S. Borkar and S.M. Mundra, Bayesian Parameter Estimation and Adaptive Control of Markov Processes with Time - Averaged Cost, *Applicationes Mathematicae* 25,3 (1998), 339-358.

[6] M.J. Brennan, E.S. Schwarz and R. Lagnado, Stategic Asset Allocation, *Journal of Economic Dynamics and Control* 21 (1997), 1377-1403.

[7] G. B. Di Masi and L. Stettner, Bayesian ergodic control of discrete time Markov processes, *Stochastics Stochastics Rep.* 54 (1995), 301–316.

[8] J. L. Doob, *Stochastic Processes.* J. Wiley, New York, 1953.

[9] T. E. Duncan, B. Pasik-Duncan, and L. Stettner, Adaptive control of a partially observed discrete time Markov process, *Appl. Math. Optim.* 37 (1998), 269–293.

[10] R. J. Elliott, L Aggoun, and J. B. Moore, *Hidden Markov Models: Estimation and Control.* Springer, New York, 1996.

[11] D. Hernandez-Lerma and J. B. Lasserre, *Discrete Time Markov Control Processes: Basic Optimality Criteria.* Springer, New York, 1996.

[12] T. Kazangey, D.D. Sworder, *Effective federal policies for regulating residential housing*, Proc. Summer Comp. Simulation Conf., Los Angeles, 1120-1128, 1971.

[13] R. Korn, *Optimal Portfolios. Stochastic Models for Optimal Investment and Risk Management in Continuous Time.* World Scientific, Singapore 1997.

[14] M. Mariton, *Jump Linear Systems in Automatic Control.* Marcel Dekker, New York and Bassel, 1990.

[15] R.C. Montgomery, Reliability considerations in placement of control systems components, *Proc. AIAA Guidance and Control Conference*, Gatlinburg, 1983.

[16] H. L. Royden, *Real Analysis.* MacMillan, New York, 1968.

[17] W. J. Runggaldier and L. Stettner, *Approximations of Discrete Time Partially Observed Control Problems*, Giardini, Pisa, 1994.

[18] S.P. Sethi and Q. Zhang, *Hierarchical Decision Making in Stochastic Manufacturing Systems*, Birkhäuser, Boston, Basel, Berlin 1994.

[19] L. Stettner, Ergodic control of partially observed Markov processes with equivalent transition probabilities, *Applicationes Math.* 22 (1993), 25–38.

[20] L. Stettner, Bayesian adaptive control of discrete time partially observed Markov processes, *Proc. Workshop Stochastic Theory and Control*, Lec. Notes Control Info. Sci. 280 (2002), 435-446, Springer.

[21] A.S. Willsky and B.C. Levy, Stochastic stability research for complex power systems, *Lab. Inf. Decision Systems*, Mass. Inst. Technology, report no. ET-76-C-01-2295, 1979.