Chapter 1

APPROXIMATIONS OF A CONTROLLED DIFFUSION MODEL FOR RENEWABLE RESOURCE EXPLOITATION

Sara Pasquali
Dipartimento di Matematica - Università degli Studi di Parma
Via M. D'Azeglio, 85 - 43100 Parma - Italy
pasquali@prmat.math.unipr.it

Wolfgang J. Runggaldier
Dipartimento di Matematica Pura ed Applicata - Università degli Studi di Padova
Via Belzoni, 7 - 35131 Padova - Italy
runggal@galileo.math.unipd.it

Abstract We study the problem of a renewable resource exploitation as a problem of optimal stochastic control with the renewable resource being managed for social benefit.

The aim is to maximize finite horizon total discounted utility by controlling per capita consumption and extraction capacity. We suppose the state is not directly observable, so we have an optimal stochastic control problem with partial observations.

The exact solution is difficult to obtain, so we aim at a nearly optimal control determined via an approximation approach involving a discretization procedure in time and space and leading thus to a Markovian decision problem.

1. STATEMENT OF THE PROBLEM

Let $(\Omega, \mathbf{F}, P)$ be a probability space.

Let us denote by $x(t)$ the resource and by $y(t)$ the population at time t. Assume that x and y satisfy the following equations:

$$dx(t) = [g(x(t)) - F(x(t), s(y(t)), D(t))]\, dt + \sigma_1 dw^{(1)}(t) \ , \ x(0) = x_0 \tag{1.1}$$

$$dy(t) = [\theta(t) F(x(t), s(y(t)), D(t)) - (vD(t) + C(t)) s(y(t))] dt$$
$$+\sigma_2 dw^{(2)}(t) \quad , \qquad y(0) = y_0 \tag{1.2}$$

where $g(x(t))$ is the renewal rate of the resource, $F(x(t), s(y(t)), D(t))$ is the total harvest of the resource, $\theta(t)$ is the proportion of harvested resource used by each member of $y(t)$, $D(t)$ is the potential harvest capacity, $vD(t)$ is the harvest cost, $C(t)$ is per-capita consumption and $s(y)$ is a C^∞ transformation, bounded and Lipschitz in y. Moreover, σ_1 and σ_2 are two positive constants with σ_2 small, while $w^{(1)}(t)$ and $w^{(2)}(t)$ are independent Wiener processes. We observe that in equations (1.1) and (1.2) there is an additive noise term because both resource and population can be subject to random variations. We also remark that in equation (1.2) we have a small noise term because random fluctuations in population are smaller than those in the resource; in fact sharp variations in population are due only to catastrophe. Model (1.1), (1.2) is a stochastic extension of a deterministic one studied by Regev, Gutierrez, Schreiber and Zilbermann in [2].

Furthermore, θ satisfies the following equation

$$d\theta(t) = \frac{(\theta(t) - \lambda)(1 - \lambda - \theta(t))}{1 - 2\lambda} \left(\mu + \frac{1 - 2\theta(t)}{2(1 - 2\lambda)} \right) dt \tag{1.3}$$

$$+ \frac{(\theta(t) - \lambda)(1 - \lambda - \theta(t))}{1 - 2\lambda} dw^{(3)}(t) \quad , \qquad \theta(0) = \theta_0$$

where μ is a constant parameter:

$$d\mu(t) = 0 \quad , \qquad \mu(0) = \mu_0 \tag{1.4}$$

and $w^{(3)}$ is a Wiener process independent of $w^{(1)}$ and $w^{(2)}$. Drift and diffusion coefficients of equation (1.3) are such that $\theta \in (\lambda, 1 - \lambda)$ with λ small (this is possible by modeling θ by a C^∞ transformation from $\mathbb{R}$ to $(\lambda, 1 - \lambda)$ - see, for example, [4]) because θ represents a percentage. In order to have θ in a compact set, we restrict the interval $(\lambda, 1 - \lambda)$ by considering the interval $[\lambda + \epsilon, 1 - \lambda - \epsilon]$ with $\epsilon \to 0$. In what follows we shall denote, improperly, this closed set by $[\lambda, 1 - \lambda]$.

ASSUMPTIONS.

A1) g is bounded and Lipschitz in x.
A2) F has the form

$$F(x(t), s(y(t)), D(t)) = D(t) s(y(t)) h\left(\frac{\alpha x(t)}{D(t) s(y(t))} \right)$$

where h is the proportion of the potential demand for resources actually required, Lipschitz in (x, y) uniformly in D, and α (positive constant) is the technology parameter of resource harvesting (it represents the resource harvesting efficiency).

Let $\xi(t) := \frac{\alpha x(t)}{D(t)s(y(t))}$ and assume that the function $h(\xi)$ satisfies the following properties

- $h(\xi)$ is concave
- $\lim_{\xi\to\infty} \xi h'(\xi) = 0$
- $h'(0) = \lim_{\xi\to\infty} h(\xi) = 1$

A3) $d \leq D(t) \leq \overline{D} \ \ \forall t \in [0, \infty)$ for $\overline{D} > 0$ and for some $0 < d < \overline{D}$.

We consider the extended state (see [6]) $X(t) = \left(X^{(1)}(t), \mu(t)\right)$, with

$$X^{(1)}(t) = (x(t), y(t), \theta(t))$$

The components of the vector $X(t)$ satisfy the equations (1.1), (1.2), (1.3), (1.4). As a consequence, $X^{(1)}(t)$ satisfies an equation of the form

$$dX^{(1)}(t) = f^{(1)}(X(t), u(t))\,dt + \sigma(X(t))\,dW(t) \qquad X^{(1)}(0) = X_0^{(1)} \tag{1.5}$$

with obvious meaning of the symbols and with $u(t) = (C(t), D(t))$ being the control at time t.

Equations (1.1), (1.2), (1.3) and (1.4) represent the dynamics for a stochastic control problem in which we want to find the supremum, over the controls $C(t)$ and $D(t)$, of the reward

$$J(C(\cdot), D(\cdot)) = E\int_0^T e^{-\delta t} s(y(t))\,U(C(t))\,dt \tag{1.6}$$

where $U(C(t))$ is the per-capita utility function which is supposed to be Lipschitz in C and $\delta > 0$ is the discount factor.

The state is not fully observable, in particular for the first component (representing the renewable resource) we do not have precise information about the quantity of available resource due both to errors in measurement and to incorrect information supplied by exploiters of the resource. We, thus, introduce the following assumption:

A4) We suppose to have information on the state only through the observation process $\eta(t)$ satisfying the equation

$$d\eta(t) = x(t)\,dt + \epsilon\, d\widetilde{w}(t) \qquad \eta(0) = \widetilde{\eta}_0 \qquad \epsilon > 0 \tag{1.7}$$

where $\widetilde{w}(t)$ is a Wiener process independent of $w^{(1)}(t)$, $w^{(2)}(t)$, $w^{(3)}(t)$ and ϵ is a small positive constant.

Control space.

Recall that, by **A3)**, $d \leq D(t) \leq \overline{D}$. As regards $C(t)$, one can find by simple calculation that $C(t) \leq D(t)$, so we can take for $C(t)$ the upper bound of $D(t)$. The control space is thus

$$\mathbf{V} = \left[0, \overline{D}\right] \times \left[d, \overline{D}\right] \tag{1.8}$$

which is a compact metric space.

State space.

The processes $x(t)$ and $y(t)$ are modeled as solutions of stochastic differential equations, so their trajectories are continuous functions a.s., with values in $\mathrm{I\!R}$; $\theta(t)$ takes values in the interval $[\lambda, 1-\lambda]$ and μ_0 is uniformly distributed over an interval $[a, b]$. Consequently the state space is

$$\mathbf{X} = \mathrm{I\!R} \times \mathrm{I\!R} \times [\lambda, 1-\lambda] \times [a, b] \tag{1.9}$$

Since x and y represent quantities that must be positive, we are interested only in solutions of (1.1) and (1.2) satisfying the constraints

$$x(t) \geq 0 \quad \forall t \quad , \quad y(t) \geq \Delta \quad \forall t \tag{1.10}$$

namely such that the renewable resource does not become negative (otherwise we have exhaustion of the resource) and the population does not go below level Δ (otherwise we have extinction of the species). Moreover we want to have bounds

$$x(t) \leq \overline{X} \quad \forall t \quad , \quad y(t) \leq \overline{Y} \quad \forall t \tag{1.11}$$

that is, resource must not exceed a level $\overline{X}$ and population a level $\overline{Y}$, for $\overline{X}$ and $\overline{Y}$ sufficiently large. Since we are over a finite interval, both resource and population cannot explode. We observe that the upper limitations are introduced only in order to have x and y in a compact space, but it is possible to choose $\overline{X}$ and $\overline{Y}$ so large that in practice resource and population never reach these values.

Therefore we are interested in the trajectories contained in the space

$$\widehat{\mathbf{X}} = \left[0, \overline{X}\right] \times \left[\Delta, \overline{Y}\right] \times [\lambda, 1-\lambda] \times [a, b] \tag{1.12}$$

instead of in $\mathbf{X}$. We consider the following part of the boundary of $\widehat{\mathbf{X}}$

$$\partial\widehat{\mathbf{X}} := \left(\left\{0, \overline{X}\right\} \times \left[\Delta, \overline{Y}\right] \times [\lambda, 1-\lambda] \times [a, b]\right)$$

$$\cup \left(\left[0,\overline{X}\right] \times \left\{\Delta,\overline{Y}\right\} \times [\lambda, 1-\lambda] \times [a,b]\right)$$

which we shall improperly call boundary of $\widehat{\mathbf{X}}$.

The aim is to maximize the total utility over the set $\widehat{\mathbf{X}}$ since outside this set the problem has no economic meaning. In order to achieve this goal, we introduce a new control component, the stopping time τ (with respect to the σ-algebra generated by the observations). We want the optimal τ to be near the hitting time of the boundary of $\widehat{\mathbf{X}}$ and near the final time T otherwise we would have too quickly exhaustion of the resource or population extinction.

To this end we consider a penalization function $\overline{Q}\left(\tau, X\left(\tau\right)\right)$ which depends both on the stopping time τ and on the state evaluated at the stopping time. We choose $\overline{Q}$ in such a way that it penalizes both a τ far from the hitting time of the set $\widehat{\mathbf{X}}$ and also from the final time T. Furthermore, we choose $\overline{Q}$ bounded by $\tilde{Q}$ and Lipschitz in X.

Therefore, the new functional is:

$$J\left(u,\tau\right) = E\left\{\int_0^{T\wedge\tau} e^{-\delta t} L\left(X\left(t\right), u\left(t\right)\right) dt - \overline{Q}\left(\tau, X\left(\tau\right)\right)\right\} \tag{1.13}$$

where

$$L\left(X\left(t\right), u\left(t\right)\right) := s\left(y\left(t\right)\right) U\left(C\left(t\right)\right)$$

with $L : \mathbf{X} \times \mathbf{V} \to \mathbb{R}$ continuous in the state and the control and, for U bounded, L is bounded over the interval $[0,T]$ because $s\left(y\right)$ is bounded (we shall denote by $\widetilde{L}$ its bound). Moreover, since $s\left(y\right)$ is Lipschitz, we have that L is Lipschitz in X. Owing to the fact that U is Lipschitz in C, it follows that L is Lipschitz also in the control.

The aim is to find

$$\sup_{u(\cdot)\in\overline{\mathbf{V}},\tau\in[0,T]} J\left(u,\tau\right)$$

where $\overline{\mathbf{V}}$ is the set of the admissible controls $u\left(\cdot\right)$ (that is the set of the controls taking values in $\mathbf{V}$ and adapted to $\mathbf{F}_t^{\eta} = \sigma\left\{\eta\left(s\right), s \leq t\right\}$) and the state equation is given by (1.4) - (1.5) where the function $f^{(1)} : \mathbf{X} \times \mathbf{V} \to \mathbb{R}^3$ is a continuous function of the state. It is Borel bounded and Lipschitz in X uniformly in u; the function $\sigma : \mathbf{X} \to \mathcal{M}_d\left(3 \times 3\right)$, where $\mathcal{M}_d\left(3 \times 3\right)$ denotes the space of the diagonal 3×3 matrices with values in $\mathbb{R}$, is Borel bounded and Lipschitz in X.

The exact solution is difficult to obtain, so we aim at a nearly optimal control determined via an approximation approach involving a discretization procedure.

2. DISCRETIZATION IN TIME

In this section we approximate the continuous problem by a time discretized stochastic control problem, that is, by a Markov decision problem.

2.1 DISCRETIZATION OF THE STATE AND OF THE CONTROL

For each fixed N, we consider the subset $\mathbf{V}_N \subset \overline{\mathbf{V}}$ of step controls corresponding to the deterministic splitting

$$0 = t_0 < t_1 < ..... < t_N = T$$

of the time interval with

$$|t_i - t_{i-1}| = \frac{T}{N} \qquad \forall i = 1, ..., N$$

Therefore we have

$$u(t) = u_n \qquad \text{for } t \in [t_n, t_{n+1}[$$

with $u_n \in \mathbf{V}$ and $\mathbf{F}_{t_n}^{\eta}$-measurable ($\mathbf{F}_{t_n}^{\eta} = \sigma(\eta(t), t \leq t_n)$ where $\eta(t)$ is the observation process defined in (1.7)). We discretize also the τ-space and, instead of $[0,T]$, we consider the discrete set of time points $\{t_0, t_1, ..., t_N\}$, so that $\tau \in \left\{n\frac{T}{N}; n = 0, ..., N\right\}$.

Proposition 1 *Given $\epsilon > 0$, for N sufficiently large*

$$\left| \sup_{\tau \in \left\{n\frac{T}{N},\, n=0,1,...,N\right\}} \sup_{u \in \mathbf{V}_N} J(u,\tau) - \sup_{\tau \in [0,T]} \sup_{u \in \overline{\mathbf{V}}} J(u,\tau) \right| \leq \epsilon \qquad \textit{(1.14)}$$

Proof. The left hand side of (1.14) is equivalent to

$$\left| \sup_{\tau \in \left\{n\frac{T}{N},\, n=0,1,...,N\right\}} \sup_{u \in \mathbf{V}_N} J(u,\tau) - \sup_{\tau \in R_T} \sup_{u \in \mathbf{V}_S} J(u,\tau) \right| \qquad (1.15)$$

where $R_T = [0,T] \cap \mathbb{Q}$ and $\mathbf{V}_S = \bigcup_N \mathbf{V}_N$ (see [4]).
Formula (1.15) can be bounded above by

$$\sup_{\tau \in \left\{n\frac{T}{N},\, n=0,1,...,N\right\}} \left| \sup_{u \in \mathbf{V}_N} J(u,\tau) - \sup_{u \in \mathbf{V}_S} J(u,\tau) \right|$$

$$+ \sup_{u \in \mathbf{V}_S} \left| \sup_{\tau \in \left\{n\frac{T}{N},\, n=0,1,...,N\right\}} J(u,\tau) - \sup_{\tau \in R_T} J(u,\tau) \right|$$

It can be shown that, given $\epsilon > 0$, for each τ there exists $\overline{N}_\tau$ such that for $N \geq \overline{N}_\tau$

$$\left| \sup_{u \in \mathbf{V}_N} J(u,\tau) - \sup_{u \in \mathbf{V}_S} J(u,\tau) \right| \leq \frac{\epsilon}{2}$$

Let $\overline{N} = \max_{\tau \in \left\{n\frac{T}{N},\, n=0,\ldots,N\right\}} \overline{N}_\tau$, then for $N \geq \overline{N}$

$$\sup_{\tau \in \left\{n\frac{T}{N},\, n=0,1,\ldots,N\right\}} \left| \sup_{u \in \mathbf{V}_N} J(u,\tau) - \sup_{u \in \mathbf{V}_S} J(u,\tau) \right| \leq \frac{\epsilon}{2}$$

It can also be proved that, given $\epsilon > 0$, there exists $\overline{\overline{N}}$ such that for $N \geq \overline{\overline{N}}$

$$\left| \sup_{\tau \in \left\{n\frac{T}{N},\, n=0,1,\ldots,N\right\}} J(u,\tau) - \sup_{\tau \in R_T} J(u,\tau) \right| \leq \frac{\epsilon}{2}$$

uniformly in u. Thus, for $N \geq \max\left\{\overline{N}, \overline{\overline{N}}\right\}$, (1.14) holds. $\triangle$

Proposition 1 states that we can restrict ourselves to consider only controls $u \in \mathbf{V}_N$ and controls $\tau \in \left\{n\frac{T}{N}; n = 0,..,N\right\}$, for N sufficiently large. Therefore, we have a simpler problem than the original one. The optimal step control of the corresponding time-discretized stochastic control problem will be shown to be an ϵ-optimal control for the original problem.

Remark 1 >*From a practical point of view it is not restrictive to consider only step controls because the controls are actually applied in discrete time. In fact, whoever controls, decides how much resource can be harvested and how much resource can be consumed within a given time interval, for example once each month or each week, that is in discrete time.*

Corresponding to the splitting of the time interval into subintervals of the same width $\frac{T}{N}$, for each $N \in \mathbb{N}$, we consider a time discretized state process:

$$X^N(t) = \begin{cases} X_j^N & \text{for } t \in \left[j\frac{T}{N}, (j+1)\frac{T}{N}\right[\\ X_N^N & \text{for } t = T \end{cases} \tag{1.16}$$

where X_j^N $(j = 0, ..., N-1)$ and X_N^N depend on j and on T, respectively, and are obtained from an Euler discretization of (1.5) and (1.4). What one usually does in these cases is a Girsanov change of measure in order to transform the original problem into a problem in which state and observations are independent. This allows one to work in a product space in which the distribution of the state is furthermore the same as in the original space (see, e.g., [4] and [6]).

We shall denote by P^0 the measure under which $\eta(t)$ is a Wiener process independent of $X(t)$ and $X^N(t)$, and by P^N the measure under which $\eta(t)$

has the same form as under P, but as a function of the discretized state. More precisely,

$$d\eta\left(t\right)=x^{N}\left(t\right)dt+\epsilon dw^{N}\left(t\right)$$

with w^N a P^N-Wiener process and x^N the first component of X^N (see (1.7)).

It can be shown that the process $X^N(t)$ converges to the continuous state process $X(t)$ in fourth mean both in the original measure P and in the transformed measure P^0 (see [4] or [6]).

2.2 DISCRETIZATION OF THE REWARD FUNCTIONAL

For each $N \in \mathbb{N}$ we define the following functional:

$$J^{N}\left(u,\tau\right):=E^{0}\left\{z^{N}\left(T\right)\left[\int_{0}^{T\wedge\tau}e^{-\delta t}L\left(X^{N}\left(t\right),u\left(t\right)\right)dt\right.\right.$$
$$\left.\left.-\overline{Q}\left(\tau,X^{N}\left(\tau\right)\right)\right]\right\} \tag{1.17}$$

where $z^N(T)=\frac{dP^N}{dP^0}$, $u \in \mathbf{V}_N$, that is $u=u_j$ in the interval $\left[j\frac{T}{N},(j+1)\frac{T}{N}\right[$ with u_j measurable with respect to the σ-algebra generated by the increments of $\eta(t)$ up to time $j-1$ and $\tau \in \left\{n\frac{T}{N}; n=0,1,...,N\right\}$ is a stopping time with respect to this same σ-algebra. It is important to notice that here we use the σ-algebra $\mathbf{F}_{j-1}^{\Delta\eta}$ generated by the increments $\eta_{j-1}:=\eta\left(j\frac{T}{N}\right)-\eta\left((j-1)\frac{T}{N}\right)$, while in the continuous time case we used the σ-algebra $\mathbf{F}_t^{\eta}$ generated by the continuous process $\eta(t)$. In [4] we show that, for the time discretized problem, there is no difference in taking controls adapted to $\mathbf{F}_t^{\eta}$ or to $\mathbf{F}_{j-1}^{\Delta\eta}$ since this does not modify the solution (see also [1]).

Proposition 2

$$\left|J\left(u,\tau\right)-J^{N}\left(u,\tau\right)\right|\leq\widetilde{K}\left(\frac{T}{N}\right)^{\frac{1}{2}} \tag{1.18}$$

uniformly in the control (u,τ), with $\widetilde{K}$ constant.

Proof. From the definitions of J and J^N, applying the change of measure and Jensen's inequality, we obtain

$$\left|J\left(u,\tau\right)-J^{N}\left(u,\tau\right)\right|$$

$$\leq E^{0}\left\{z\left(T\right)\int_{0}^{T\wedge\tau}e^{-\delta t}\left|L\left(X\left(t\right),u\left(t\right)\right)-L\left(X^{N}\left(t\right),u\left(t\right)\right)\right|dt\right.$$

$$+\left|z(T)-z^N(T)\right|\left[\int_0^{T\wedge\tau}e^{-\delta t}\left|L\left(X^N(t),u(t)\right)\right|dt+\left|\overline{Q}\left(\tau,X^N(\tau)\right)\right|\right]$$

$$+z(T)\left|\overline{Q}\left(\tau,X^N(\tau)\right)-\overline{Q}(\tau,X(\tau))\right|\Big\}$$

where $z(T)=\frac{dP}{dP^0}$. By Tonelli's theorem and Hölder's inequality, recalling that L and Q are bounded by $\widetilde{L}$ and $\widetilde{Q}$ respectively and Lipschitz, it follows that

$$\left|J(u,\tau)-J^N(u,\tau)\right|\le L'\int_0^T e^{-\delta t}\left\{E\left[\left\|X(t)-X^N(t)\right\|^4\right]\right\}^{\frac{1}{4}}dt+\Big(\widetilde{Q}$$

$$+\widetilde{L}\frac{1-e^{-\delta T}}{\delta}\Big)\cdot\left[E^0\left|z(T)-z^N(T)\right|^2\right]^{\frac{1}{2}}+Q'\left[E\left\|X(t)-X^N(t)\right\|^4\right]^{\frac{1}{4}}$$

where L' and Q' are the Lipschitz constants relative to L and Q, respectively.

As mentioned, the sequence $X^N(t)$ converges to $X(t)$ in fourth mean under both P and P^0; furthermore the sequence $z^N(T)$ converges to $z(T)$ in mean square under P^0, in particular we have $E\left\|X(t)-X^N(t)\right\|^4\le K_1\left(\frac{T}{N}\right)^2$ where K_1 is a constant (and the same upper bound is valid for the expectation E^0) and $E^0\left\|z(T)-z^N(T)\right\|^2\le K_2\frac{T}{N}$ with K_2 constant (see [1, 4]), so we immediately obtain the thesis. △

Thanks to proposition 2, we can approximate the initial reward J by the discretized reward J^N, for N sufficiently large. In this way, the discretized problem consists in maximizing J^N by choosing the controls $u\in\mathbf{V}_N$ and $\tau\in\left\{n\frac{T}{N}:n=0,...,N\right\}$. Due to the uniformity, in (u,τ), of the bound in (1.18), these controls are nearly optimal for the original problem.

We have now a partially observable discrete time stochastic control problem over a finite horizon with discounted reward or, equivalently, a partially observed Markov decision problem. We shall write the so-called separated problem associated to it, namely a corresponding problem with fully observed state, and then apply the DP algorithm to this problem after discretizing also in space.

2.3 THE SEPARATED PROBLEM

Let

$$P\left(X_{n+1},X_n,u_n,\eta_n\right):=P\left(X_{n+1}\left|X_n,u_n\right.\right)\exp\left\{\frac{1}{\epsilon^2}x_n\eta_n-\frac{1}{2\epsilon^2}\left(x_n\right)^2\frac{T}{N}\right\} \tag{1.19}$$

where

$$P\left(X_{n+1}\left|X_n,u_n\right.\right)=\Phi_1\left(x_{n+1};X_n,u_n\right)\Phi_2\left(y_{n+1};X_n,u_n\right)$$
$$\cdot\Phi_3\left(\theta_{n+1};X_n,u_n\right)$$

is the transition kernel obtainable as product of the conditional distributions of each component of $X^{(1)}$. The factors Φ_1, Φ_2 and Φ_3 are obtained from the discretization of the state equation (1.5) and, using an Euler discretization of the state equation, Φ_1 and Φ_2 are normal and Φ_3 can be transformed into a normal distribution by a change of variable (we have to consider the same C^∞ transformation mentioned in section 1 - see [4]).

Let $\eta^n=(\eta_0,...,\eta_n)$ be the observation vector up to step n and $u^n=(u_0,...,u_n)$ the sequence of controls up to step n. Consider the process $\{q_n(\cdot)\}$ defined recursively as follows

$$q_0\left(X_0\right)=p_0\left(X_0\right)$$

$$q_{n+1}\left(X_{n+1};\eta^n,u^n\right)=\int_{\mathbf{X}}P\left(X_{n+1},X_n,u_n,\eta_n\right)q_n\left(X_n;\eta^{n-1},u^{n-1}\right)dX_n \tag{1.20}$$

for $n=0,...,N-1$, where $p_0\left(X_0\right)$ is the initial distribution of the process $\{X_n\}$.

Assuming that the initial values of resource, population and percentage of the resource that are being used are known and that the parameter μ is uniformly distributed over the interval $[a,b]$, the initial distribution of the state is

$$p_0\left(X\right)=\delta\left(x-x_0\right)\delta\left(y-y_0\right)\delta\left(\theta-\theta_0\right)\frac{1}{b-a} \tag{1.21}$$

We observe that (1.20) corresponds to a recursive Bayes' formula for computing an unnormalized conditional distribution $q_{n+1}\left(X_{n+1};\eta^n,u^n\right)$ of X_{n+1} given the observations η^n and the controls u^n.

Proposition 3 *Any function $\Phi\left(X_{j-1},u_{j-1},\tau\right)$, of one of the following types*

1. $\Phi'\left(X_{j+1},u_{j+1}\right)I_{\left]j\frac{T}{N},T\right]}\left(\tau\right)$
2. $\Phi''\left(X_{j+1},u_{j+1}\right)I_{\left[j\frac{T}{N},(j+1)\frac{T}{N}\right[}\left(\tau\right)$

satisfies the property:

$$E^0\left[\Phi\left(X_{j+1},u_{j+1},\tau\right)z^N\left(T\right)\Big|\eta^j,u^j\right]=\int_{\mathbf{X}}\Phi\left(X,u_{j+1},\tau\right)$$

$$\cdot q_{j+1}\left(X;\eta^j,u^j\right)dX \tag{1.22}$$

where the control u_{j+1} is adapted to the σ-algebra generated by the increments η_n of the observations up to time j and the control τ is a stopping time with respect to this same σ-algebra. We recall that $z^N\left(T\right)=\frac{dP^N}{dP^0}$.

For the proof see [1].

Going back to the functional $J^N\left(u\right)$ in (1.17), recalling (1.16) and applying proposition 3, we have:

$$J^N\left(u,\tau\right)=E^0\left\{\sum_{j=0}^{N-1}\int_{\mathbf{X}}\left[\frac{\exp\left\{-\delta\left(j\frac{T}{N}\right)\right\}-\exp\left\{-\delta\left(\left(j+1\right)\frac{T}{N}\right)\right\}}{\delta}\right.\right.$$

$$\left.\cdot L\left(X,u_j\right)I_{\left]j\frac{T}{N},T\right]}\left(\tau\right)-\overline{Q}\left(j\frac{T}{N},X\right)I_{\left[j\frac{T}{N},\left(j+1\right)\frac{T}{N}\right[}\left(\tau\right)\right] \tag{1.23}$$

$$\left.\cdot q_j\left(X;\eta^{j-1},u^{j-1}\right)dX-\int_{\mathbf{X}}\overline{Q}\left(T,X\right)I_{\{T\}}\left(\tau\right)q_N\left(X;\eta^{N-1},u^{N-1}\right)dX\right\}$$

Now we have a completely observable stochastic control problem where the new state is the unnormalized conditional density $q_n\left(X_n;\eta^{n-1},u^{n-1}\right)$. The aim is to find the supremum of (1.23) subject to the state equation (1.20).

3. FURTHER DISCRETIZATIONS

At this point we are not yet able to apply the dynamic programming algorithm to the separated problem because the state is infinite-dimensional and takes a continuum of possible values. Therefore we discretize the state space $\widehat{\mathbf{X}}$.

3.1 STATE DISCRETIZATION

First we note that the state X can leave the set $\widehat{\mathbf{X}}$ so we have to define the problem on all of $\mathbf{X}$, but one can show that there exists a finite band $\overline{B}$ such that, with probability close to 1, X never leaves $\widehat{\mathbf{X}}\cup\overline{B}$. Consequently, for each positive integer m, we consider a partition $\{B_h\}_{h=1}^{\overline{M}}$ (with $\overline{M}=2^{4m}$) of the state space $\widehat{\mathbf{X}}\cup\overline{B}$ such that each B_h has width going to zero as we refine the partition and, from each of these subsets B_h, we select a representative element. The set of these representative elements forms the discretized state space.

Define the process

$$q_0^{(m)}(X_0) = p_0(X_0)$$

$$q_{n+1}^{(m)}(X_{n+1};\eta^n,u^n) = \int_{\mathbf{X}} P^{(m)}(X_{n+1},X_n,u_n,\eta_n) \tag{1.24}$$

$$\cdot q_n^{(m)}\left(X_n;\eta^{n-1},u^{n-1}\right)dX_n$$

for $n = 0, ..., N-1$, which has the same form as (1.20), but with P substituted by $P^{(m)}$, that is the analog of (1.19) discretized in the state in such a way that

- $E^0\left\|P(X_{n+1},X_n,u_n,\eta_n) - P^{(m)}(X_{n+1},X_n,u_n,\eta_n)\right\| \le \overline{H}_m$ where $\|\cdot\|$ is a norm in L^1 with respect to X_{n+1} and the expectation is with respect to η_n.
- $lim_{m\to\infty}\overline{H}_m = 0$

With these assumptions we can show the following result:

Proposition 4 *For N sufficiently large, for each $n = 0, 1, ..., N$, we have*

$$E^0\left[\left\|q_n\left(X_n;\eta^{n-1},u^{n-1}\right) - q_n^{(m)}\left(X_n;\eta^{n-1},u^{n-1}\right)\right\|\right] \le K_1\overline{H}_m \tag{1.25}$$

uniformly in the control u, where K_1 is a constant, the norm (which is a norm in L^1) is with respect to X_n, the expectation is with respect to η^{n-1}.

Proof. We proceed by induction. For $n = 0$, (1.25) is immediately verified. Suppose that (1.25) is true for n and consider $n+1$:

$$E^0\left[\left\|q_{n+1}(X_{n+1};\eta^n,u^n) - q_{n+1}^{(m)}(X_{n+1};\eta^n,u^n)\right\|\right]$$

$$\le E^0\int_{\mathbf{X}}\int_{\mathbf{X}} P^{(m)}(X_{n+1},X_n,u_n,\eta_n)\left|q_n\left(X_n;\eta^{n-1},u^{n-1}\right)\right. \tag{1.26}$$

$$\left.-q_n^{(m)}\left(X_n;\eta^{n-1},u^{n-1}\right)\right|dX_n dX_{n+1}$$

$$+E^0\int_{\mathbf{X}}\int_{\mathbf{X}}\left|P(X_{n+1},X_n,u_n,\eta_n) - P^{(m)}(X_{n+1},X_n,u_n,\eta_n)\right| \tag{1.27}$$

$$q_n\left(X_n;\eta^{n-1},u^{n-1}\right)dX_n dX_{n+1}$$

Applying Tonelli's theorem and observing that

$$E^0\left\|P^{(m)}(X_{n+1},X_n,u_n,\eta_n)\right\| \le H_1$$

where H_1 is a constant, (1.26) can be bounded above by

$$H_1 E^0 \left\| q_n \left(X_n; \eta^{n-1}, u^{n-1} \right) - q_n^{(m)} \left(X_n; \eta^{n-1}, u^{n-1} \right) \right\| \leq H_2 \overline{H}_m \quad (1.28)$$

where we have applied the induction hypotesis. H_2 is a constant.

By Tonelli's theorem we find for (1.27) the following upper bound

$$\int_{\mathbf{X}} E^0 \left[q_n \left(X_n; \eta^{n-1}, u^{n-1} \right) E^0 \left\| P \left(X_{n+1}, X_n, u_n, \eta_n \right) \right. \right.$$

$$\left. \left. - P^{(m)} \left(X_{n+1}, X_n, u_n, \eta_n \right) \right\| \right] dX_n$$

and it can be shown that this quantity is bounded above by

$$H_3 \overline{H}_m \quad (1.29)$$

where H_3 is a constant. From (1.28) and (1.29) we obtain the thesis. △

We denote by $J^{N,m}(u, \tau)$ the functional corresponding to (1.23) for the discretized state, that is obtained by substituting q_j with $q_j^{(m)}$ in (1.23).

Proposition 4 allows us to show the following

Proposition 5 *For N sufficiently large we have*

$$\left| J^N (u, \tau) - J^{N,m} (u, \tau) \right| \leq N \left(\frac{1}{\delta} \tilde{L} + \widetilde{Q} \right) K_1 \overline{H}_m$$

uniformly in $u \in \mathbf{V}_N$ and $\tau \in \left\{ n \frac{T}{N}; n = 0, 1, ..., N \right\}$ and where K_1 is a constant.

Proof.

$$\left| J^N (u, \tau) - J^{N,m} (u, \tau) \right| \leq \sum_{n=1}^{N} \left(\frac{\widetilde{L}}{\delta} + \widetilde{Q} \right)$$

$$\cdot E^0 \left\| q_n \left(X; \eta^{n-1}, u^{n-1} \right) - q_n^{(m)} \left(X; \eta^{n-1}, u^{n-1} \right) \right\|$$

Applying the previous proposition we immediately obtain the thesis. △

It follows that for N and m sufficiently large we can approximate J^N by $J^{N,m}$ and look for an optimal control for the discretized problem. Again, by the uniformity of the bound in (u, τ), this latter control is nearly optimal for the original problem.

3.2 THE DISCRETIZED PROBLEM IN ALTERNATIVE FORM

Just as q_n, also $q_n^{(m)}$ is infinite-dimensional but, if the function $P^{(m)}$ can be expressed as product of a function of the state at time $n+1$ and a function of the state and observation at time n, then the functions $q_n^{(m)}$ can be written in terms of finite-dimensional statistics of $\left(\eta^{n-1}, u^{n-1}\right)$.

In fact, if we define

$$d_1^{(h)}(\eta_0) := \exp\left(\frac{1}{\epsilon^2} x_0 \eta_0 - \frac{1}{2\epsilon^2} x_0^2 \frac{T}{N}\right) I_{\pi_x(B_h)}(x_0)\, I_{\pi_y(B_h)}(y_0)\, I_{\pi_\theta(B_h)}(\theta_0) \tag{1.30}$$

and

$$d_{n+1}^{(h)}\left(\eta^n, u^{n-1}\right) := \sum_{\substack{h' = 1 \\ \pi_\mu(B_{h'}) \equiv \pi_\mu(B_h)}}^{\overline{M}} d_n^{(h')}\left(\eta^{n-1}, u^{n-2}\right) \varphi_{h,h'}(\eta_n, u_{n-1}) \tag{1.31}$$

where

$$\varphi_{h,h'}(\eta, u) := \int_{\pi_x(B_h)\times\pi_y(B_h)\times\pi_\theta(B_h)} \exp\left(\frac{1}{\epsilon^2} x\eta - \frac{1}{2\epsilon^2} x^2 \frac{T}{N}\right) \cdot \Psi_{h'}((x,y,\theta), u)\, dx dy d\theta$$

with

$$\Psi_h((x_{n+1}, y_{n+1}, \theta_{n+1}), u_n) := \Phi_1(x_{n+1}; X_h, u_n) \cdot \Phi_2(y_{n+1}; X_h, u_n)\, \Phi_3(\theta_{n+1}; X_h) \tag{1.32}$$

and X_h is the representative point of the set B_h, then it follows

Proposition 6 *For each $n = 1, ..., N$ we have*

$$q_n^{(m)}\left(X_n; \eta^{n-1}, u^{n-1}\right) = \sum_{h=1}^{\overline{M}} d_n^{(h)}\left(\eta^{n-1}, u^{n-2}\right) \Psi_h((x_n, y_n, \theta_n), u_{n-1}) \cdot \frac{1}{b-a} I_{\pi_\mu(B_h)}(\mu_n) \tag{1.33}$$

where $d_1^{(h)}\left(\eta^0, u^{-1}\right) = d_1^{(h)}(\eta_0)$ and $\pi_\mu(B_h)$ is the projection of the set B_h over the component μ (see [4]).

The objective functional $J^{N,m}$ can be expressed as

$$E^0\left\{\frac{1-\exp\left[-\delta\frac{T}{N}\right]}{\delta}s\left(y_0\right)U\left(C_0\right)I_{]0,T]}\left(\tau\right)-I_{\left[0,\frac{T}{N}\right[}\overline{Q}\left(0,X_0\right)\right.$$

$$+\sum_{n=1}^{N-1}\sum_{h=1}^{\overline{M}}d_n^{(h)}\left(\eta^{n-1},u^{n-2}\right)\left[\frac{\exp\left[-\delta n\frac{T}{N}\right]-\exp\left[-\delta\left(n+1\right)\frac{T}{N}\right]}{\delta}\right.$$

$$I_{\left]n\frac{T}{N},T\right]}\left(\tau\right)U\left(C_n\right)\int_{\mathbb{R}}s\left(y\right)\Phi_2\left(y;X_h,u_{n-1}\right)dy$$

$$\int_{\pi_\mu(B_h)}\frac{1}{b-a}d\mu-I_{\left[n\frac{T}{N},(n+1)\frac{T}{N}\right[}\left(\tau\right)\int_{\mathbf{X}}\overline{Q}\left(n\frac{T}{N},X\right)\Psi_h\left(\left(x,y,\theta\right),u_{n-1}\right) \tag{1.34}$$

$$\left.\frac{1}{b-a}I_{\pi_\mu(B_h)}\left(\mu\right)dX\right]-\sum_{h=1}^{\overline{M}}d_N^{(h)}\left(\eta^{N-1},u^{N-2}\right)$$

$$\left.\int_{\mathbf{X}}\overline{Q}\left(T,X\right)I_{\{T\}}\left(\tau\right)\Psi_h\left(\left(x,y,\theta\right),u_{N-1}\right)\frac{1}{b-a}I_{\pi_\mu(B_h)}\left(\mu\right)dX\right\}$$

where the expectation is with respect to the sequence $\{\eta_n\}$. We note that, under the measure P^0, $\{\eta_n\}$ is a sequence of i.i.d. random variables, normally distributed with mean 0 and variance $\epsilon^2\frac{T}{N}$.

The process $\left\{d_n^{(h)}\left(\eta^{n-1},u^{n-2}\right)\right\}$ is now finite-dimensional, but it still takes an infinite number of possible values since η_n and u_n do. It is then necessary to make some further approximations.

3.3 ϵ-OPTIMALITY

Controls take values in a compact space $\mathbf{V}$ and we can assume that also the observations take values in a compact space $\mathbf{S}$ since the expectation in (1.34) is finite and if we restrict the values of η_n to a sufficiently large compact set $\mathbf{S}$, then, due to the boundedness of the costs, the corresponding change in the value of $J^{N,m}$ is, uniformly in u, negligible (see [6]).

Take finite partitions $\{\mathbf{V}_k\}_{k=1}^K$ and $\{\mathbf{S}_z\}_{z=1}^Z$ of the compact sets $\mathbf{V}$ and $\mathbf{S}$, respectively, and choose a representative element for each of these sets:

$$v_k\in\mathbf{V}_k\qquad k=1,...,K;\qquad s_z\in\mathbf{S}_z\qquad z=1,...,Z$$

We denote by $\mathbf{V}_N^K$ the set of the control sequences taking as possible values the representative elements v_k and by $\mathbf{V}_N^{Kes}$ the set of the controls obtained from the controls of $\mathbf{V}_N^K$ by a step interpolation relative to the partition of the time interval into subintervals of the same width $\frac{T}{N}$.

Let $J^{N,m,Z}$ be the functional obtained from $J^{N,m}$ by substituting the observations and the controls with their discretized values.

>From propositions 1, 2, 5 and the continuity, in u and y, of the functions in (1.34), we have the following two theorems (see [4]):

Theorem 1 *For $\epsilon > 0$ fixed, for N, m, Z and K sufficiently large*

$$\left| \sup_{\tau \in \left\{ n\frac{T}{N}; n=0,...,N \right\}} \sup_{u \in \mathbf{V}_N^K} J^{N,m,Z}(u,\tau) - \sup_{\tau \in [0,T]} \sup_{u \in \overline{\mathbf{V}}} J(u,\tau) \right| \leq \epsilon$$

Theorem 2 *For each control $(u,\tau) \in \mathbf{V}_N \times \left\{ n\frac{T}{N}; n = 0, ..., N \right\}$, for N, m, Z, K sufficiently large, we have*

$$\left| J^{N,m,Z}(u,\tau) - J(u,\tau) \right| \leq \epsilon$$

Consequently

Corollary 1 *An optimal control for the discretized control problem, extended in the sense seen above (that is by a step interpolation), is ϵ-optimal for the original problem.*

>From here follows immediately that it sufficies to find an optimal control for the problem discretized in time and space and the control obtained from the latter by a step interpolation (which is a control in $V_N^{K_{es}}$) will be nearly optimal for the original continuous time problem.

3.4 DP ALGORITHM

Since the functional $J^{N,m,Z}$ has the same form as (1.34) but with the observations substituted by their discretized values, and since, furthermore, the optimal value of the functional is calculated for controls $u \in \mathbf{V}_N^K$ with (1.30) and (1.31) as the state equations, we can write the DP algorithm to obtain an optimal control for this problem, in the following way:

- $$\widehat{J}_N\left(\overline{\eta}^{N-1}, \overline{u}^{N-1}\right) = -\sum_{h=1}^{\overline{M}} d_N^{(h)}\left(\overline{\eta}^{N-1}, \overline{u}^{N-2}\right) \int_{\mathbf{X}} \overline{Q}(T,X)\, I_{\{T\}}(\tau)$$
$$\Psi_h\left((x,y,\theta), \overline{u}_{N-1}\right) \frac{1}{b-a} I_{\pi_\mu(B_h)}(\mu)\, dX \qquad (1.35)$$

■

$$\widehat{J}_n\left(\overline{\eta}^{n-1},\overline{u}^{n-1}\right)=\max\left\{\max_{\overline{u}_n}\left[\sum_{h=1}^{\overline{M}}d_n^{(h)}\left(\overline{\eta}^{n-1},\overline{u}^{n-2}\right)U\left(\overline{C}_n\right)\right.\right.$$
$$\cdot\frac{\exp\left[-\delta n\frac{T}{N}\right]-\exp\left[-\delta\left(n+1\right)\frac{T}{N}\right]}{\delta}\int_{\mathbb{R}}s\left(y\right)\Phi_2\left(y;X_h,\overline{u}_{n-1}\right)dy$$
$$\left.\int_{\pi_\mu(B_h)}\frac{1}{b-a}d\mu+\sum_{z=1}^{Z}P^0\left(\eta\in S_z\right)\widehat{J}_{n+1}\left(\left(\overline{\eta}^{n-1},s_z\right),\left(\overline{u}^{n-1},\overline{u}_n\right)\right)\right];$$
$$-\sum_{h=1}^{\overline{M}-1}d_n^{(h)}\left(\overline{\eta}^{n-1},\overline{u}^{n-2}\right)\left[\int_{\mathbf{X}}\overline{Q}\left(n\frac{T}{N}X\right)\Psi_h\left(\left(x,y,\theta\right),\overline{u}_{n-1}\right)\right.$$
$$\left.\left.\cdot\frac{1}{b-a}I_{\pi_\mu(B_h)}\left(\mu\right)dX\right]\right\}\tag{1.36}$$

for $n=N-1,...,1.$

■

$$\widehat{J}_0=\max\left\{\max_{\overline{u}_0}\left[\frac{1-e^{-\delta\frac{T}{N}}}{\delta}s\left(y_0\right)U\left(\overline{C}_0\right)+\sum_{z=1}^{Z}P^0\left(\eta\in S_z\right)\right.\right.$$
$$\left.\left.\widehat{J}_1\left(s_z,\overline{u}_0\right)\right];-\overline{Q}\left(0,X_0\right)\right\}\tag{1.37}$$

By $\overline{\eta}_n$ we denote the discretized observations and by $\overline{u}_n$ the discretized controls. We have used the fact that $\overline{\eta}_n$ can take only a finite number Z of possible values. We observe that the dependence from the current control is in the term $U\left(\overline{C}_n\right)$.

Remark that at each step it is necessary to calculate two rewards: one in the case in which we stop at that moment and the other in the case in which we decide to go on. We then choose the maximum between the two rewards in order to know the optimal stopping time τ for each pair of observations and controls.

The DP algorithm provides us with the sequence of optimal controls (that is the optimal strategy) to be applied at times $\left\{n\frac{T}{N};n=0,1,...,N-1\right\}$ as function of the current statistic $\left(d_n^{(h)}\left(\overline{\eta}^{n-1},\overline{u}^{n-2}\right),\overline{u}_{n-1}\right)$. Extending this sequence in the sense described previously we obtain the optimal strategy for the initial problem.

This DP algorithm can be implemented in order to determine the optimal strategy. In [4] one can find an application relative to an Euler discretization of the state equation along with the numerical aspects and various numerical results.

References

[1] A. Bensoussan, W. Runggaldier, An approximation method for stochastic control problems with partial observation of the state - a method for constructing ϵ-optimal controls, *Acta Appl. Math.*, **10** (1987), 145-170.

[2] A.P. Gutierrez, U. Regev, S.J. Schreiber, D. Zilbermann, Biological and Economic Foundation of Renewable Resource Exploitation, *Ecological Economics*, **26**(3) (1998), 227-242.

[3] H.J. Kushner, *Probability methods for approximations in stochastic control and for elliptic equations*, Academic Press, New York,1977.

[4] S.Pasquali, *A stochastic control problem for renewable resource exploitation*, Ph.D. Thesis, University of Padova, 1999.

[5] W.J. Runggaldier, L. Stettner, *Approximations of Discrete Time Partially Observed Control Problems*, Giardini Editiori e Stampatori in Pisa, 1994.

[6] W.J. Runggaldier, O. Zane, Approximations for discrete-time adaptive control: construction of ϵ-optimal controls, *Math. Control, Signals and Systems*, **4** (1991), 269-291.