ON OPTIMAL STOCHASTIC CONTROL PROBLEM OF LARGE SYSTEMS

J.P. QUADRAT
Domaine de Voluceau - B.P. 105
78153 LE CHESNAY Cédex

I - INTRODUCTION.

We discuss three different approaches, leading to numerical methods, for the solution of optimal stochastic control problem of large dimension :

- Optimization in the class of local feedbacks,
- Monte Carlo and stochastic gradient techniques,
- Perturbation methods in the small intensity noise case.

We consider the stochastic control problem of diffusion processes in the complete observation case

$$(1)\qquad \begin{cases} dX_t = b\,(X_t,U_t)dt + dW_t\,, & X_t \in \mathbb{R}^n,\ U_t \in \mathbb{R}^m \\ V(o,y) = \underset{u}{\mathrm{Min}}\ \mathbb{E}\,\{\int_o^{+\infty} e^{-\lambda t}\, C(X_t,U_t)dt\ |X(o) = y\} & \end{cases}$$

The solution of the Hamilton Jacobi equation

$$(2)\qquad \underset{u}{\mathrm{Min}}\,\{b(x,u)\ \mathrm{grad}\ V + C(x,u)\} + \Delta V - \lambda V = o$$

gives the optimal cost and the optimal strategies of (1).

The numerical solution of (2) is almost impossible in the general situation when n is large. The difficulty is not a problem of numerical analysis but an irreducible difficulty. To see that consider the simpler problem

$$(3)\qquad \begin{cases} \Delta V - \lambda V = C & x \in \mathcal{O} = [0,1]^n \\ V_{\partial\mathcal{O}} = 0 & \end{cases}$$

where $\partial\mathcal{O}$ denotes the boundary of $\mathcal{O}$.

For such problem it is easy to show that the number of eigen vectors associated to an eigen value smaller than a fixed value, increases exponentially with the dimension. But we need to have a good representation of the eigen vector associated to eigen value of small modulus, in any good finite dimensional approximation of (2). And thus, whatever could be the approximation, the obtention of a given precision will be obtain at a cost which increases exponentially with the dimension.

In the three following point of view we avoid this difficulty but we have a lost of optimality.

II - OPTIMIZATION IN THE CLASS OF LOCAL FEEDBACKS.

In this paragraph we give the optimality conditions in the class of local feedbacks, and show that it is more difficult to solve these conditions than to compute the solution of the Hamilton-Jacobi equation. Than we study two particular cases :

- the case of the uncoupled dynamics,
- the case of systems having the product form property.

In these cases only it is possible to compute the optimal local feedbacks for large systems. Finally we discuss briefly the decoupling point of view.

II-1. The general situation.

Given I the indexes of the subsystems $I = \{1,2,..,k\}$ n_i,[resp-m_i] denotes the number of the states [resp. the controls] of the subsystem $i \in I$. The local feedback S_i is a mapping of $\mathbb{R}^+ \times \mathbb{R}^{n_i}$ in $\mathcal{V}_i \subset \mathbb{R}^{m_i}$ the set of the admissible values of the control i. $\mathcal{S}_L$ denotes the class of local feedbacks $\mathcal{S}_L = \{S = (S_1,\ldots,S_I)\}$. Given the drift term of the system :

$$b : \mathbb{R}^+ \times \mathbb{R}^n \times \mathcal{V} \to \mathbb{R}^n$$
$$t \quad x \quad u \quad b(t,x,u)$$

with $n = \sum_{i\in I} n_i$, $\mathcal{V} = \prod_{i\in I} \mathcal{V}_i$,

- the diffusion term :

$$\sigma : \mathbb{R}^+ \times \mathbb{R}^n \to M_n$$
$$\quad t \quad x \qquad \sigma(t,x)$$

with M_n the set of matrices (n,n) and $a = \frac{1}{2}\sigma\sigma^*$ where $*$ denotes the transposition

- the instantaneous cost :

$$c = \mathbb{R}^+ \times \mathbb{R}^n \times \mathcal{V} \to \mathbb{R}^+$$
$$\quad t \quad x \quad u \qquad c(t,x,u)$$

than boS [resp coS] denotes the functions $\mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^n$

[resp $\mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^+$] $b(t,x,S(t,x))$ [resp $c(t,x,S(t,x))$]

Than if X^S denotes the diffusion (boS,a) (drift boS, and diffusion term σ) and P^S_μ its measure defined on $\Omega = C(\mathbb{R}^+, \mathbb{R}^n)$ with μ the law of the initial condition we want to solve

$$\underset{S\in\mathcal{S}_L}{\text{Min}} \ \mathbb{E}_{P^S_\mu} \int_o^T CoS(t,\omega_t)dt$$

where $\omega \in \Omega$, T denotes the time horizon. We have here a team of I players working to optimize a unique criterium.

A simple way to obtain the optimality conditions is to consider another formulation of this problem : the control of the Fokker Planck equation that is :

$$\underset{S\in\mathcal{S}_L}{\text{Min}} \ J^S = \int_Q CoS(t,x)p^S(t,x)dt\,dx$$

with p solution of

$$\mathcal{L}^*_S\, p^S = 0$$

$$p^S(o,.) = \mu$$

with $Q = [0,T] \times \mathcal{O}$ and $\mathcal{O} = \mathbb{R}^n$

$$\mathcal{L}_S = \frac{\partial}{\partial t} + \sum_j b_j oS \frac{\partial}{\partial x_j} + \sum_{i,j} a_{ij} \frac{\partial^2}{\partial x_i \partial x_j}$$

μ the law of the initial condition.

Than we have :

Theorem 1

A N.S.C. for $J^R \geq J^S$, R $S \in \mathcal{S}_L$, is that :

(1) $H(t,R,p^R,V^S) \geq H(t,S,p^R,V^S)$ pp in t

with

$$(2) \quad \begin{cases} H(t,R,p,V) = \int_{\mathcal{O}} [CoR(t,x) + \sum_i b_i oR(t,x) \frac{\partial V}{\partial x_i}(t,x)]\, p(t,x)dx \\ \mathcal{L}_R^* p^R = 0 \quad p^R(o,.) = \mu \; ; \; \mathcal{L}_S V^S + CoS = 0, \quad V^S(T,.) = 0 \end{cases}$$

Remark 1. From this theorem the Pontriagyn can be obtained, that is a necessary condition of optimality of the strategy S is that : p,V,S satisfies

$$H(t,S,p^S,V^S) = \underset{R \in \mathcal{S}_L}{\text{Min}}\ H(t,R,p^S,V^S) \; ;$$

$$(3) \quad \begin{cases} \mathcal{L}_S^* p^S = 0 \quad , \quad p(o,.) = \mu \quad ; \\ \mathcal{L}_S V^S + CoS = 0 \quad , \quad V^S(T,.) = 0. \end{cases}$$

A proof is given in J.L. Lions [1].

Remark 2. This theorem give an algorithm to improve a given strategy R that is :

Step 1 : compute p^R
Step 2 : solve backward simultaneously

$$(4) \quad \begin{cases} \mathcal{L}_S V^S + CoS = 0 \qquad V^S(T,.) = 0 \\ S \in \text{Arg}\ \underset{Z}{\text{Min}}\ H(t,Z,p^R,V^S) \end{cases}$$

By this way we obtain a better strategy S.

A fixed point of the application $R \to S$ will satisfy the conditions (3).

We see that one iteration (4) of this algorithm is more expensive than the computation cost of the solution of the H.J.B. equation.

II-2. Uncoupled dynamic systems.

This is the particular case where b_i is a function of x_i and u_i, $\forall i \in I$

$$b_i : \mathbb{R}^+ \times \mathbb{R}^{n_i} \times \mathcal{V}_i \rightarrow \mathbb{R}^{n_i}$$
$$\quad t \quad x_i \quad u_i \qquad b_i(t,x_i,u_i)$$

and the noises are not coupled between the subsystems that is :

$$\sigma_i : \mathbb{R}^+ \times \mathbb{R}^{n_i} \rightarrow M_{n_i}$$
$$\quad t \quad x_i \qquad \sigma_i(t,x_i)$$

In this situation we have

$$p^R = \prod_{i\in I} p_i^{Ri}$$

with p_i^{Ri} solution of

(5) $$\mathcal{L}^*_{i,R_i} p_i^{Ri} = 0 \qquad p_i^{Ri}(o,.) = \mu_i \quad \text{with } \mu = \prod_{i\in I} \mu_i$$

and

$$\mathcal{L}_{i,R_i} = \frac{\partial}{\partial t} + \sum_{k\in I_i} b_k o R_i(t,X)\frac{\partial}{\partial X_k} + \sum_{k,l\in I_i} a_{kl}\frac{\partial^2}{\partial X_k \partial X_l}$$

with $\quad Ii = \{ \sum_{j<i} n_j < k \leq \sum_{j<i+1} n_j \}.$

Let us denote by

(6) $$C_i^R o R_i : \mathbb{R}^+ \times \mathbb{R}^{ni} \rightarrow \mathbb{R}^+$$
$$\quad t \quad x_i \qquad \int CoR(t,x) \prod_{j\neq i} p_j^{Rj}(t,X_j)dX_j$$

That is the conditional expectation of the instantaneous cost knowing the information only on the local subsystem i.
We have the following sufficient conditions to be optimal player by player :

Theorem 2. A sufficient condition for a strategy S to be optimal player by player is that the following conditions are satisfied :

(7) $\underline{\min_{R_i} [\mathcal{L}_{i,R_i} V_i + C_i^R o R_i] = o, \; i \in I \;;}$

$\underline{\text{with } C_i^R o R_i \text{ defined by (6) and (5)}}$

$\underline{\text{The optimal cost is } \mu_1(V_1)..=\mu_I(V_I) \text{ with } \mu_i(V_i) = \int_{\mathbb{R}^{ni}} \mu_i(dX_i)V_i(o,X_i)}$

Remark 3. The theorem 3 gives an algorithm to compute a feedback optimal player by player

given ε, $\nu \in \mathbb{R}^+$

Step 1) Choose $i \in I$
Solve (7)
if : $\mu_i(V_i) \leq \nu - \varepsilon$ than $\nu := \mu_i(V_i)$
$R_i := \operatorname{Arg} \min_{R_i} \{\mathcal{L}i, R_i V_i + C_i^R o R_i\}$

if not choose another $i \in I$ until

$$\mu_i(V_i) \geq \nu - \varepsilon, \; \forall i \in I.$$

Step 2) When $\mu_i(V_i) \geq \nu - \varepsilon$, $\forall i \in I$, than $\varepsilon := \frac{\varepsilon}{2}$, go to step 1.

By this algorithm we obtain a decreasing sequence $\nu^{(n)}$ which converges to a cost optimal player by player.

For a proof of a discrete version of this algorithm see Quadrat-Viot [2].

Remark 4. The interpretation of $V_i(t,X_i)$ $i \in I$ in terms of the variables of theorem 1 is :

$$V_i(t,x_i) = \int V(t,x) \prod_{j \neq i} p^{Rj}(t,X_j)dX_j$$

Remark 5. In this situation we have to solve a coupled system of P.D.E. but each of them is on a space of small dimension. By this way we can optimize , in the class of local feedback, systems which are not reachable by H.J.B. equation. An application to hydropower systems is given in Delebecque-Quadrat [4].

II-3. Systems having the product form property.

The property that a system has its dynamic uncoupled is very restrictive in this paragraph, we show systems which have their invariant measure uncoupled, they are limit of network of queues of Jackson type. This property can be used to apply to them the results of II-2 for the corresponding ergodic control problem that is:

$$\min_{S} \lim_{T\to\infty} \frac{1}{T} \int_0^T C\circ S(\omega_t)dt$$

Given B a generator of a Markov chain defined on $E = \{1,2,\dots,n\}$, a function $E \times \mathbb{R} \to \mathbb{R}$, $(i,x) \mapsto u_i(x)$ a matrix $\sigma \in M_n$, $A = \frac{1}{2}\sigma\sigma^*$, D a diagonal matrix satisfying :

(8) $$DB^* + BD + 2A = 0$$

Theorem 3

The invariant measure of probability p of the diffusion ($b = Bu$, $a=A$) such that (8) is true has the product form property that is :

(9) $$p(x) = C \prod_{i=1}^{n} p_i(x_i) , \qquad i \in E$$

(10) $$p_i(x_i) = \exp - \frac{1}{d_{ii}} \int_0^{x_i} u_i(s)ds$$

where C is a constant of normalization.

Demonstration : The Fokker-Planck equation can be written :

(11) $$-\mathrm{div}\,[bp] + \mathrm{div}\,[A \,\mathrm{grad}\, p] = 0$$

Let us make the change of variables $p = \exp V$ in (11), we obtain

$$(\mathrm{grad}\, V, b-A \,\mathrm{grad}\, V) + \mathrm{div}\,(b-A\,\mathrm{grad}\, V) = 0$$

Using (10), we have :

(12) $$(D^{-1}u, (B + AD^{-1})u) + \mathrm{tr}\,[(B + AD^{-1})\,\mathrm{grad}\, u] = 0$$

The quadratic part in (u) of (12) is equal to 0 if and only if :

$$D^{-1}B + B^*D^{-1} + 2D^{-1}AD^{-1} = 0$$

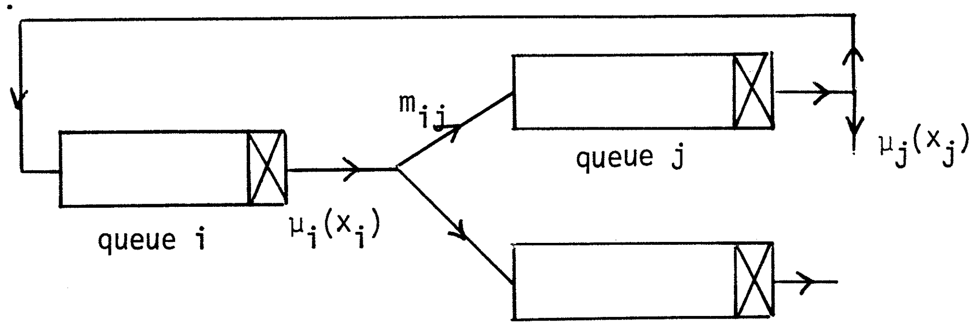
which can be written :

$$BD + DB^* + 2A = 0$$

which is (8).

We have also $\operatorname{tr}[B + AD^{-1}]\operatorname{grad} u = 0$. Indeed grad u is diagonal because u_i is a function of x_i only and the coefficient of $\frac{\partial u_i}{\partial x_i}$ is $b_{ii} + a_{ii} / d_{ii}$ which is equal to zero by (8).

Remark 6. This class of diffusion processes are quite natural if we see them as the limit process when $N \to \infty$, obtained from Jackson network of queues by the scaling $x \to \frac{x}{N}$, $t \to \frac{t}{N^2}$.



where $\mu_i(x_i)$ is the output rate of the queue i, m_{ij} is the probability of a customer leaving the queue i to go to the queue j.

The correlation of the noise given by (8) corresponds to system for which the noise satisfies a conservation law (for example the total number of customer in a closed network of queues).

Remark 7. We can now apply the result of II-2 to compute the optimal local feedback for systems having the product form property and an ergodic criterium. Indeed :

$$\operatorname{Min} \frac{1}{T}\int_0^T C\circ S\ (\omega_t)dt = \int C\circ S\ (x)\ p(x)dx$$

$$p(x) = \prod_{i=1}^{n} p_i(x_i)$$

and p_i satisfies :

$$-\frac{\partial}{\partial x_i}[u_i p_i] + \frac{\partial}{\partial x_i^2}[d_{ii}\ p_i] = 0, \qquad i \in E$$

$$\int p_i(x_i)\ dx_i = 1$$

II-4. Remarks on decoupling.

Another way to use the results of II-2 when the dynamic is coupled is to do a change of feedback let us consider the simpler case

$$\begin{array}{ccccc} b : \mathbb{R}^n & \times & U & \to & \mathbb{R}^n \\ x & & u & & b(x,u) \end{array} \qquad \text{with } u \subset \mathbb{R}^n$$

we use the feedback transformation $v = b(x,u)$ to decouple the drift terms. Now v is the control and we can apply the results of II-2 to compute the best local feedback $v_i = S_i(x_i)$.

Then the solution in u of

(13) $\qquad b(x,u) = S(x)$

gives the best feedback among the class that we can call "local decoupling feedbacks".

One difficulty with this approach is for example the constraints on the control : the image by b of an hypercube is not in general an hypercube and if we take for constraints on the new control $v \in V(x) \subset b(x,\mathcal{U})$ with $V(x)$ an hypercube of $\mathbb{R}^n$, the lost of optimality can become unacceptable.

This approach is well studied for deterministic linear and non linear systems Womham [8], Isidori [17] and in the dynamic programming litterature Larson [11].

III - OPTIMIZATION IN A PARAMETRIZED CLASS OF FEEDBACKS BY MONTE CARLO TECHNIQUES.

We have seen in §2 that we are able to compute the optimal local feedback only in particular cases. Moreover, sometimes the local information is not good ; we can have, a priori, an idea on a better one and would like to use this a priori information to solve a simpler problem than the general one. A way to do that is to parametrize the feedback, optimize the open loop parameter by a Monte Carlo technique. More precisely given the stochastic control problem

(1) $$\begin{cases} dX_t = b(t,x_t,u_t)dt + dw_t & x_t \in \mathbb{R}^n, \quad u_t \in \mathbb{R}^m \\ \text{Min } E \int_0^T C(t,x_t,u_t)dt \end{cases}$$

we make the feedback transformation

(2) $$u(t) = S(t,x_t,v_t) \qquad v_t \in \mathbb{R}^p$$

where $S : \mathbb{R}^+ \times \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^m$ is given.

We use for the approximation of the probability law of the noise P, the distribution

$$\mu = \frac{1}{r} \sum_{j=1}^{r} \delta_{\omega_j}(\omega)$$

where ω_j are trajectories of the noise obtained by random generation perhaps after a discretization time if we want to avoid the difficulties of the non existence of a solution trajectory by trajectory to (1). And now we have to solve :

$$(3)\quad \begin{aligned} & dx_t^j = b(t,x_t^j,S(t,x_t^j,v_t))dt + dw_t^j \\ & \underset{v}{\mathrm{Min}}\ \frac{1}{r} \sum_{j=1}^{r} \int_0^T C(t,x_t^j,S(t,x_t^j,v_t))dt \end{aligned}$$

w_t^j denotes here a particular trajectory of the noise. Thus, at the end, we have to solve a deterministic dynamic control problem. For that, we can use a gradient technique or the Pontryagin principle. For discrete time system the convergence properties of this approach has been studied in Quadrat-Viot [15]. Application to the French hydropower system is currently done at EDF now. Feedbacks on the demand of electricity and the level of water in the local dam are optimized with success by this technique (Lederer - Colleter [2]).

The idea of the stochastic gradient method is the same than the former one but we use a recursive way to optimize. The recursivity being on the index of the trajectory of the noise generated. The problem (1) (2) can be reduced to the problem

$$\underset{v \in V}{\mathrm{Min}}\ \mathbb{E}\, J(v)$$

is a situation where we are able to compute $\frac{\partial J}{\partial v}$ byadjoint state technique here

$$J(v) = \int_0^T C(t,x_t,S(t,x_t,v_t))dt.$$

Moreover, we can consider that after discretization v is finite dimensional. Then the stochastic gradient algorithm is the following recursive way to improve the parameter v

$$v_{r+1} = P_V \{v_r - \rho_r \frac{\partial J}{\partial v}(v_r,\omega_r)\} \qquad \rho_r \in \mathbb{R}^+, \quad \forall r \in \mathbb{N} ,$$

$$\sum_{r \in \mathbb{N}} \rho_r = \infty \quad , \qquad \sum_{r \in \mathbb{N}} \rho_r^2 < \infty$$

ω_r denotes a generated random realization of the stochastic parameter in the definition of J(v), for our problem (1), (2) that is a realization of the Wiener process w_t. P_V denotes the projection on the set V.

In a convex situation which is not in general the case for the problem (1), (2), we can give some global convergence results.

Theorem 1

On the hypotheses

1) $v \to J(\omega,v)$ convex $\forall\omega$;

2) $\omega \to J(\omega,v)$ is L' , $\forall v$

3) $\sup_{\substack{v\in V\\ \omega\in\Omega}} |\frac{\partial J}{\partial v}(v,\omega)| \leq q$;

4) $\mathbf{E}\, J(v) - \bar{J}^* \geq C \; \ell^2(v)$

where $\bar{J}^*$ denotes the optimal cost and $\ell(v)$ denotes the distance of v to the optimal set.

5) V a bounded convex set, we have $\lim_n \mathbf{E}\, \ell^2(v_n) = 0$ and moreover if $\rho_r = \dfrac{1}{cr + \frac{q^2}{y_\circ C}}$

with $y_o = \mathbf{E}\, \ell^2(v_o)$ we have :

$$\mathbf{E}\, \ell^2(v_r) \leq \frac{1}{\frac{c^2}{q^2} r + \frac{1}{y_o}}$$

The proof of this theorem can be found Dodu-Goursat-Hertz-Quadrat-Viot [5] a lot of similar results can be found in Polyak [14] and in the reference of this paper. In Kushner-Clark [12] local convergence are proved in the non convex case.

The following result shows that in some sense the stochastic gradient algorithm is optimal. We suppose that :

6) the noise is finite valued and we denote by $v_\mu = \text{Argmin}\, \mathbf{E}_\mu \; J(v)$ and we suppose that

7) $v \to J(\omega,v)$ is two times differentiable and uniformly convex $\forall\omega \in \Omega$ than we have D.G.H.Q.V. [5].

Theorem 2

On the hypothese 6) and 7) we have :

$$E_\mu \, (\hat{v} - v)^{\otimes 2} \geq \frac{1}{r} H_\mu^{-1} Q_\mu H_\mu^{-1}$$

with

$$H_\mu = \frac{\partial^2}{\partial v^2} \mathbf{E}_\mu \; J(v)$$

$$Q_\mu = \mathbf{E}_\mu \left(\frac{\partial J}{\partial v}(v_\mu)\right)^{\otimes 2}$$

for all unbiased statistic $\hat{v}$ of v_μ defined on $(\Omega,\mu)^{\otimes r}$.

If we remark that $\frac{1}{C}$ is an estimation of H_μ^{-1} and q^2 an estimation of Q_μ we see that in a certain sense the speed of convergence of the stochastic gradient technique is optimal.

We have applied this algorithm to the problem of the optimization of the investment of a transmission network of electricity D.G.H.Q.V. [5]. The comparison with a sophisticated simplex approach shows that the stochastic gradient mathematic is undoubtly better.

IV - PERTURBATION METHODS.

By perturbation methods we can reduce a difficult problem to a simpler one. In this chapter we study the small intensity noise case. In this situation it is possible to build an affine control which leads to ε^4 error with respect to the optimal control, where ε denotes diffusion term.

We consider the following stochastic control problem :

(1)
$$\begin{cases} dX_t = f(x_t,u_t)dt + \varepsilon dw_t \quad , \quad x_t \in \mathbb{R}^n \ , \ u_t \in \mathbb{R}^m \\ V^\varepsilon(o,y) = \underset{u}{\text{Min}}\ \mathbf{E}\ [\int_o^T C(x_t,u_t)dt \mid X(o) = y]. \end{cases}$$

where ε belongs to $\mathbf{R}^+$ and is small.

We denote by

(2) $$H(x,u,p) = p\ f(x,u) + C(x,u)$$

We suppose that :

(3) $u \to f(x,u)$ is linear ;

(4) $(v,\frac{\partial^2 C}{\partial u^2} v) \geq k\ |v|^2$ where k is a positive constant, $\forall x$;

Let us consider the deterministic control problem

(5)
$$dX_t = f(x_t,u_t)dt$$
$$V(o,y) = \underset{u}{\text{Min}}\{\int_o^T C(x_t,u_t)dt \mid X(o) = y\} ;$$

and denote by $u_o(t)$ the optimal open loop deterministic control problem.

The second variation calculus around the optimal trajectory of (6) Cruz [3] gives the quadratic form "osculatrice" of the optimal cost V around the optimal trajectory. This quadratic form is defined by the (n,n) time dependent matrix P solution of the Riccati equation :

(6) $\overset{\circ}{P} + PA + A'P - PSP + Q = 0 \qquad P(T) = 0$

where

(7) $A = f_x - f_u H_{uu}^{-1} H'_{ux}$

(8) $S = f_u H_{uu}^{-1} f'_u$

(9) $Q = H_{xx} - H_{ux} H_{uu}^{-1} H'_{ux}$

are evaluated along the optimal trajectory of (5) on the hypotheses that :

(10) $H_{uu} > 0, \quad H_{xx} - H_{ux} H_{uu}^{-1} H'_{ux} \geq 0$

Let us consider the following affine control :

(11) $u_f(t,X(t)) = u_o(t) + K(t) [X(t) - X_o(t)]$

where $X_o(t)$ denotes the optimal trajectory of the deterministic control problem (5), $X(t)$ the actual trajectory of the diffusion process (1_1) when the control is (11), and $K(t)$ is defined by

(12) $K(t) = H_{uu}^{-1} (H'_{ux} + f'_u P)(t)$

evaluated on the optimal trajectory $X_o(t)$.

We have :

Theorem

On the hypotheses (3), (4), (10) and (f,c) two time differentiable, the affine control build on the deterministic control problem, used in the stochastic control problem leads to a lost of optimality of order $\mathcal{O}(\varepsilon^4)$.

Ideas of the proof : Fleming [6]has shown that the optimal deterministic feedback used in the stochastic control problem leads to an error of $\mathcal{O}(\varepsilon^4)$. But in the estimation of the proof he does not need the optimal deterministic control but a control which gives exact V, $\frac{\partial V}{\partial y}$, $\frac{\partial^2 V}{\partial y^2}$ along the optimal trajectory of the deterministic control problem.

Using for example Cruz [3] we know that the affine feedback (11) has this property and thus the result is proved.

REFERENCES.

[1] BASKET - CHANDY - MUNTZ - PALACIOS : Open Closed and Mixed Network of Queues with Different Class of Customers, JACM 22, pp. 248-260

[2] COLLETER - LEDERER : Rapports internes EDF sur la gestion des réservoirs hydroélectriques français

[3] CRUZ : Feedback Systems, MacGraw Hill, 1972.

[4] DELEBECQUE - QUADRAT : Contribution of Stochastic Control Singular Perturbation Averaging and Team Theories to an Example of Large Scale System : Management of Hydropower Production, IEEE AC, April 1978.

[5] DODU - GOURSAT - HERTZ - QUADRAT - VIOT : "Méthodes de Gradient Stochastique pour l'Optimisation des Investissements dans un Réseau Electrique", EDF Bulletin Série C, n°2, 81, pp. 133-164.

[6] FLEMING : Stochastic Control for Small Noise Intensities, Brown University Report, April 1970.

[7] HOLLAND :-Small Noise Open Loop Control, SIAM J. Control, 12, August 1974.
-Gaussian Open Loop Control Problems, SIAM J. Control, 13, August 1975.

[8] ISIDORI - KRENER - GORI GIORGI - MONACO : Non Linear Decoupling via Feedback a Differential Geometric Approach, IEEE, AC 26, n°2, April 1981.

[9] JACKSON : Jobshop Like Queueing Systems, Management Science, 10, pp. 131-142, 1963.

[10] KELLY : Reversibility and Stochastic Networks, J. Wiley and Sons, 1979.

[11] KORSAK - LARSON : Dynamic Successive Approximation Technique with Convergence Proofs, Automatica, 1969.

[12] KUSHNER - CLARK : Stochastic Approximation Methods for Constrained and Unconstrained Systems, Springer Verlag, 1978.

[13] LIONS : Contrôle Optimal des Systèmes Distribués, Dunod, 1968.

[14] POLYAK : Subgradient Methods : a Survey of Soviet Research in Non Smooth Optimization, eds. C. Lemaréchal & R. Mifflin, Pergamon Press, 1978.

[15] QUADRAT - VIOT : Méthodes de Simulation en Programmation Dynamique Stochastique, RAIRO, April 1973, pp. 3-22.

[16] " " : Product Form and Optimal Local Feedback for Multiindex Markov Chains, Allerton Conference, 1980.

[17] WONHAM : Linear Multivariable Control : A Geometric Approach , New York, Springer Verlag, 1979.