STOCHASTIC MINIMUM PRINCIPLE FOR PARTIALLY OBSERVED SYSTEMS SUBJECT TO CONTINUOUS AND JUMP DIFFUSION PROCESSES AND DRIVEN BY RELAXED CONTROLS

N.U. Ahmed[†] and C.D. Charalambous[‡]

University of Ottawa[†] , Ottawa, Canada
University of Cyprus[‡], Nicosia, Cyprus.

ABSTRACT

In this paper we consider non convex control problems of stochastic differential equations driven by relaxed controls. We present existence of optimal controls and then develop necessary conditions of optimality. We cover both continuous diffusion and Jump processes.

Key Words stochastic differential equations, continuous Diffusion, Jump processes, Relaxed controls, Existence of optimal controls, necessary conditions of optimality.

2000 AMS Subject Classification 49J55, 49K45,93E20.

1 Introduction

The basic idea of the deterministic minimum principle introduced by Pontryagin and his colleagues in the 1950's is to derive a set of necessary and sufficient conditions that must be satisfied by any control which yields an optimal cost or pay-off. It consists of a system of forward-backward differential equations (e.g., state and adjoint equations) and the extremum of a Hamiltonian functional. Since then the theory has been extensively developed in many directions, such as, optimal control theory for finite dimensional deterministic systems with regular controls [1, 2], where one can find a broad and deep generalization of the classical Pontryagin minimum (equivalently maximum) principle for deterministic systems. The minimum principle is also extended to infinite dimensional systems, see [3, 4, 5, 6] and the references cited therein.

The stochastic minimum principle is another important extension of the Pontryagin minimum principle for systems subject to probabilistic randomness. In the stochastic case, there are basically different approaches based on the assumptions employed to derive the stochastic minimum principle. Specifically, [7] utilizes spike variations and Neustadt's variational principle, [8] utilizes Girsanov's measure transformation for non degenerate controlled diffusion processes, while [9] utilizes the martingale representation to derive the adjoint equation. The martingale representation approach is further developed in [10, 11]. Further results utilizing the martingale representation approach are established in [12] for control dependent diffusion processes utilizing second-order variations leading to a minimum principle which differs

from the deterministic case in the sense that the effect of control dependent diffusion terms are fully explored. Subsequent extensions are given in [13] for stochastic systems with random coefficients, in [14] utilizing stochastic flows to derive results similar to [12], and in [15] establishing relationships between stochastic minimum principle and dynamic programming. The martingale approach to stochastic minimum principle sparked the interest in studying backward and forward stochastic differential equations. An excellent account on the stochastic minimum principle is found in [16] which also includes an anthology of references. Extensions of the stochastic maximum principle for relaxed controls using the topology of weak convergence are found in [17, 18, 19, 20, 21], where relations to strict controls are also investigated. Recent developments and extensions are found in [22, 23, 24] and references therein. The area of mathematical finance, specifically portfolio optimization, has utilized the stochastic minimum principle extensively to derive optimal strategies.

In general, the stochastic minimum principle is specific to the information structures available to the control. Specifically, in applications of control theory, there are many problems in physical sciences and engineering, where systems are modeled by stochastic differential equations driven by controls which are also stochastic processes with specific information structure, such as, full information or partial information. Mathematically, information structures are modeled via the minimal sigma algebra generated by the available information process, and it is this process that the controller uses to generate control actions. For full-information problems in which the information structure is Markovian, one often employs Bellman's principle of optimality to construct, what is known as, the HJB (Hamilton-Jacob-Bellman) equation, a nonlinear PDE defined on the state space of the system under investigation. This equation describes the evolution of the value function which is used to construct the state feedback control law provided this function is at least once differentiable with respect to the state variable. This however requires solving the HJB equation which may have a viscosity solution but not sufficiently smooth [16]. For non-Markovian controlled diffusion systems with general information structures the HJB equation does not apply. For information structures which correspond to full information or partial information the stochastic minimum principle is often employed [11, 25, 6], although the partial information case is mathematically more demanding. However, this line of research is feasible provided existence of optimal controls is guaranteed. For non convex control problems, it is well known that the problem may have no optimal solution if the admissible controls are merely measurable functions with values in the set U which is non convex. Nevertheless, this problem can be partially overcome by introducing the relaxed controls and then approximating the relaxed controls by the standard regular controls.

In this article we consider stochastic control systems with information structures corresponding to full information and partial information, which are driven by relaxed controls. Specifically, controls which are conditional probability distributions, measurable with respect to full or partial information. We treat stochastic differential equations driven by both Brownian motion and Levy process or Poisson jump process. We show existence of optimal policies among the class of relaxed controls under gen-

eral conditions, with respect to an established topology of weak* convergence. Then we proceed with the derivation of stochastic minimum principle, for both the full information and the partial information cases. The Hamiltonian system of equations is derived in a systematic manner utilizing the semi martingale representation theorem and the Riesz representation theorem, leading very naturally to the existence of the adjoint processes satisfying a Backward stochastic differential equation in an appropriate space. We also discuss the realizability of relaxed controls by regular controls using the Krein-Millman theorem. The methodology we consider is applied to stochastic differential equations driven by both Brownian motion and Poisson jump process. The basic procedure follows the one introduced in [5, 2] for deterministic systems, augmented by the martingale representation approach to stochastic control. The material presented for full information compliment the previous work on relaxed controls found in [17, 18, 19, 20, 21], where the authors utilize alternative methods to derive related results.

The rest of the paper is organized as follows. In section 2 we present some typical notations and formulate the optimal control problem considered in this paper. In section 3, we consider the question of existence of optimal relaxed controls. Section 4 contains an interesting fundamental result characterizing semi martingales. Here we construct a Hilbert space characterizing the space of semi martingales (starting from zero). This is used later in the development of necessary conditions. Section 5 is devoted to the development of necessary conditions of optimality. In section 6 we extend the previous results to cover stochastic systems driven by jump processes. In section 7, we specialize to regular controls and obtain the usual necessary conditions of optimality. In section 8 we address the question of realizability of relaxed controls by regular controls. The paper is concluded with some comments on possible extensions of our results.

2   Formulation of Stochastic Relaxed Control Problem

In this section we introduce the mathematical model for the stochastic control system and the pay-off functional as a measure of performance. The distinction between full and partial information structures are also presented.

Let $(\Omega, \mathcal{F}, \mathcal{F}_{t\geq 0}, P)$ denote a complete filtered probability space where $\{\mathcal{F}_t, t \geq 0\}$ is an increasing family of subsigma algebras of the σ-algebra $\mathcal{F}$. For any random variable z, $\mathcal{E}(z) \equiv \int_{\Omega} z(\omega) P(d\omega)$ denotes the expected value (average) of the random variable z. Let $\{W(t), t \geq 0\}$ denote the R^m-valued standard Brownian motion with $P\{W(0) = 0\} = 1$ defined on the filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_{t\geq 0}, P)$. Let $\mathcal{G}_t \subset \mathcal{F}_t$ denote a family of sub-sigma algebras of the σ-algebra $\mathcal{F}_t, t \geq 0$.

Let $I = [0, T]$ be any finite interval, U any closed bounded subset of R^d and $\mathcal{M}(U)$ the space of regular bounded signed Borel measures on $\mathcal{B}(U)$, the Borel subsets of U and $\mathcal{M}_1(U) \subset \mathcal{M}(U)$ the space of regular probability measures. Controls based on partial information (respectively full information) will be described through the topological dual of the Banach space $L_1^a(I, C(U))$, the L_1-space of $\mathcal{G}_t$ (respectively $\mathcal{F}_t$) adapted $C(U)$ valued functions. The dual of this space is given by $L_\infty^a(I, \mathcal{M}(U))$

which, for partial information, consists of weak star measurable $\mathcal{G}_t$ adapted $\mathcal{M}(U)$ valued functions (signed measures), while for full information it consists of $\mathcal{F}_t$ adapted functions defined similarly. For controls based on partial information (respectively full information) we are interested in the subspace $L^a_\infty(I, \mathcal{M}_1(U)) \subset L^a_\infty(I, \mathcal{M}(U))$ of probability measure valued $\mathcal{G}_t$ (respectively $\mathcal{F}_t$) adapted functions. Let $\mathcal{U}_{ad} \equiv L^a_\infty(I, \mathcal{M}_1(U))$ denote the class of admissible controls, called the relaxed controls, where the distinction between full information and partial information is only specified in terms of the σ-algebras $\mathcal{F}_t$ and $\mathcal{G}_t$, respectively.

We consider the following stochastic system in R^n governed by the Ito differential equation which is driven by relaxed control,

$$dx(t) = \left(\int_U b(t, x(t), \xi)u_t(d\xi)\right) dt + \left(\int_U \sigma(t, x(t), \xi)u_t(d\xi)\right) dW(t), x(0) = x_0, t \in I,$$

where $b : I \times R^n \times U \longrightarrow R^n$ denotes the drift and $\sigma : I \times R^n \times U \longrightarrow \mathcal{L}(R^m, R^n)$ the diffusion parameters. For simplicity of notation we prefer to write the above controlled dynamic system in the form

$$dx(t) = b(t, x(t), u_t) \, dt + \sigma(t, x(t), u_t) \, dW(t), x(0) = x_0, t \in I, \tag{1}$$

for any $u \in \mathcal{U}_{ad}$. The cost functional is given by

$$J(u) \equiv \mathcal{E}\left\{\int_0^T \ell(t, x(t), u_t)dt + \Phi(x(T))\right\}.$$

The problem is to find a control $u^o \in \mathcal{U}_{ad}$ such that $J(u^o) \leq J(u)$ for all $u \in \mathcal{U}_{ad}$. We consider the question of existence of optimal controls and characterization of such controls in the form of necessary conditions of optimality (Pontryagin minimum principle). For necessary conditions of optimality we follow the procedure developed in [2], pp.271-293 which we extend from deterministic to stochastic systems.

3 Existence of Optimal Relaxed Controls

Consider the system (1) with b and σ denoting the infinitesimal generators representing the drift and diffusion given by the Borel measurable maps:

$$b : I \times R^n \times R^d \longrightarrow R^n, \sigma : I \times R^n \times R^d \longrightarrow \mathcal{L}(R^m, R^n).$$

We assume that they satisfy the following basic properties: there exists a $K \in L_2^+(I)$(nonnegative functions belonging to $L_2(I)$), such that

(A1) : $|b(t, x, \xi) - b(t, y, \xi)|_{R^n} \leq K(t)|x - y|_{R^n}$ uniformly in $\xi \in U$
(A2) : $|b(t, x, \xi)|_{R^n} \leq K(t)(1 + |x|_{R^n})$ uniformly in $\xi \in U$
(A3) : $|\sigma(t, x, \xi) - \sigma(t, y, \xi)|_{\mathcal{L}(R^m, R^n)} \leq K(t)|x - y|_{R^n}$ uniformly in $\xi \in U$
(A4) : $|\sigma(t, x, \xi)|_{\mathcal{L}(R^m, R^n)} \leq K(t)(1 + |x|_{R^n})$ uniformly in $\xi \in U$
(A5) : $b(t, x, \cdot), \sigma(t, x, \cdot)$ are continuous in $\xi \in U$ uniformly in $t \in [0, T], x \in R^n$.

For admissible controls, we choose the set of relaxed controls given by $\mathcal{U}_{ad} \equiv L_\infty^a(I, \mathcal{M}_1(U))$ which are stochastic processes, adapted to a given sigma algebra (to be specified later), and taking values in the space of probability measures $M_1(U)$. This is endowed with the weak star topology also called vague topology. A sequence $u^n \in \mathcal{U}_{ad}$ is said to converge vaguely to u^o, written $u^n \overset{v}{\longrightarrow} u^o$, iff for every $\varphi \in L_1^a(I, C(U))$

$$\mathcal{E} \int_{I \times U} \varphi(t, \xi) u_t^n(d\xi) dt \to \mathcal{E} \int_{I \times U} \varphi(t, \xi) u_t^o(d\xi) dt \quad \text{as} \quad n \to \infty.$$

With respect to this vague (weak star) topology, $\mathcal{U}_{ad}$ is compact and from here on we assume that $\mathcal{U}_{ad}$ has been endowed with this vague topology.

Let $B_\infty^a(I, L_2(\Omega, R^n))$ denote the space of $\mathcal{F}_t$-adapted R^n valued second order random processes endowed with the norm topology $\| \cdot \|$ given by

$$\| x \|^2 \equiv \sup\{\mathcal{E}|x(t)|_{R^n}^2, t \in I\}.$$

With this preparation, we can now present the following lemma proving existence of solutions and their continuous dependence on controls.

Lemma 3.1 Consider the controlled stochastic differential equation (1) and suppose the assumptions (A1)-(A5) hold. Then for any $\mathcal{F}_0$-measurable initial state x_0 having finite second moment, and any $u \in \mathcal{U}_{ad}$, the system (1) has a unique solution $x \in B_\infty^a(I, L_2(\Omega, R^n))$ having continuous modification. In other words, $x \in C(I, R^n)$ P-a.s. Further, the solution is continuously dependent on the control in the sense that as $u^n \overset{v}{\longrightarrow} u^o$ in $\mathcal{U}_{ad}$, the corresponding solutions $x^n \overset{s}{\longrightarrow} x^o$ in $B_\infty^a(I, L_2(\Omega, R^n))$.

Proof. The proof for the first part of the lemma is classical and hence we present only an outline. It is based on the Banach fixed point theorem applied to the operator F on the Banach space $B_\infty^a(I, L_2(\Omega, R^n))$ where

$$(Fx)(t) \equiv x_0 + \int_0^t b(s, x(s), u_s) \, ds + \int_0^t \sigma(s, x(s), u_s) \, dW(s), t \in I \equiv [0, T]. \quad (2)$$

Under the assumptions (A1)-(A4), it is easy to verify using classical martingale inequality that $F : B_\infty^a(I, L_2(\Omega, R^n))$ to itself. Then using the metric d given by $d = d_T$ where

$$d_t^2(x, y) \equiv \sup\{\mathcal{E}|x(s) - y(s)|_{R^n}^2, 0 \le s \le t\}$$

for $t \in I$, one can verify that the $n - th$ iterate of F denoted by $F^n \equiv FoF \cdots oF$ (n times) is a contraction. Then by Banach fixed point theorem F^n has a unique fixed point in $B_\infty^a(I, L_2(\Omega, R^n))$ and hence F itself has one and the same fixed point [1]. The continuity of the sample paths however follows from classical Borel-Canteli lemma. Now consider the second part asserting the continuity of the control to solution map $u \longrightarrow x$. For this one proceeds as follows. Suppose the assumption (A5) holds and let $\{u^n, u^o\}$ be any sequence of controls from $\mathcal{U}_{ad}$ and $\{x^n, x^o\}$ denote the corresponding sequence of solutions of the system (1). Let $u^n \overset{v}{\longrightarrow} u^o$. We must show that $x^n \overset{s}{\longrightarrow} x^o$ in $B_\infty^a(I, L_2(\Omega, R^n))$. We present only a hint. Using the definition of solution, it is

easy to verify that

$$x^n(t) - x^o(t) = \int_0^t [b(s, x^n(s), u_s^n) - b(s, x^o(s), u_s^n)]ds$$

$$+ \int_0^t [\sigma(s, x^n(s), u_s^n) - \sigma(s, x^o(s), u_s^n)]dW(s) + e_{1,n}(t) + e_{2,n}(t), t \in I$$

.
$$(3)$$

where

$$e_{1,n}(t) = \int_0^t [b(s, x^o(s), u_s^n) - b(s, x^o(s), u_s^o)]ds$$

$$e_{2,n}(t) = \int_0^t [\sigma(s, x^o(s), u_s^n) - \sigma(s, x^o(s), u_s^o)]dW(s).$$

Using the standard martingale inequality it follows from this that there exist constants $C_1, C_2 > 0$ such that

$$E|x^n(t) - x^o(t)|^2 \leq C_1 \int_0^t K^2(s)E|x^n(s) - x^o(s)|^2 + C_2\big(\mathcal{E}|e_{1,n}|^2 + \mathcal{E}|e_{2,n}|^2\big). \quad (4)$$

Clearly,

$$\mathcal{E}|e_{1,n}|^2 \leq T\mathcal{E} \int_0^t |b(s, x^o(s), u_s^n) - b(s, x^o(s), u_s^o)|_{R^n}^2 ds$$

and

$$\mathcal{E}|e_{2,n}|^2 \leq 4\mathcal{E} \int_0^T |\sigma(s, x^o(s), u_s^n) - \sigma(s, x^o(s), u_s^o)|_{\mathcal{L}(R^m, R^n)}^2 ds.$$

Now note that by virtue of vague convergence of u^n to u^o, the integrands of the above inequalities converge to zero for almost all $s \in I$, P-a.s and it follows from (A2) and (A4) that they are dominated by integrable functions. So by Lebesgue dominated convergence theorem the integrals $\{e_{1,n}, e_{2,n}\}$ converge to zero uniformly on I. The assertion then follows from Gronwall inequality, applied to the inequality (4). This completes the outline. ●

Optimal Control Problem. Consider the controlled system (1) and the cost functional given by

$$J(u) \equiv \mathcal{E}\Big\{ \int_0^T \ell(t, x(t), u_t)dt + \Phi(x(T)) \Big\} \quad (5)$$

where ℓ and Φ are suitable functions which are measures of mismatch between the desired flow and the flow that results from the choice of the control u. The problem, as stated in section 2, is to find a control from the class of admissible (relaxed) controls $\mathcal{U}_{ad}$ that minimizes the functional (5). We present the following existence result.

Theorem 3.2 Consider the control problem as stated above. Suppose the assumptions of Lemma 3.1 hold, and further suppose $\ell : I \times R^n \times U \longrightarrow (-\infty, +\infty]$ and $\Phi : R^n \longrightarrow (-\infty, +\infty]$ are Borel measurable maps satisfying the following conditions:

(a1): $x \longrightarrow \ell(t, x, \xi)$ is continuous on R^n for each $t \in I$, uniformly with respect to $\xi \in U$.

(a2): $\exists \; h \in L_1^+(I)$ such that $|\ell(t, x, \xi)| \leq h(t)(1 + |x|^2_{R^n})$

(a3): $x \longrightarrow \Phi(x)$ is lower semi continuous on R^n and $\exists \; c_0, c_1 \geq 0$ such that $|\Phi(x)| \leq c_0 + c_1 |x|^2_{R^n}$.

Then, there exists an optimal control $u \in \mathcal{U}_{ad}$ at which J attains its minimum.

Proof. Since $\mathcal{U}_{ad}$ is compact in the vague topology, it suffices to prove that J is lower semi continuous with respect to this topology. Suppose $u^n \xrightarrow{v} u^o$ in $\mathcal{U}_{ad}$ and let $\{x^n, x^o\} \subset B^a_\infty(I, L_2(\Omega, R^n))$ denote the solutions of equation (1) corresponding to the sequence of controls $\{u^n, u^o\} \subset \mathcal{U}_{ad}$. Then by Lemma 3.1, along a subsequence if necessary, $x^n \xrightarrow{s} x^o$ in $B_\infty(I, L_2(\Omega, R^n))$. First note that, in view of the strong convergence, along a subsequence if necessary, $x^n(T) \to x^o(T)$ P-a.s. Thus it follows from assumption (a3) and Fatou's Lemma that

$$\mathcal{E}\{\Phi(x^o(T))\} \leq \liminf_n \mathcal{E}\{\Phi(x^n(T))\}. \tag{6}$$

Considering the running cost, it is easy to see that

$$\mathcal{E} \int_I \ell(t, x^o(t), u^o_t) \; dt = \mathcal{E} \int_I \ell(t, x^o(t), u^o_t - u^n_t) \; dt$$
$$+ \mathcal{E} \int_I (\ell(t, x^o(t), u^n_t) - \ell(t, x^n(t), u^n_t)) \; dt + \mathcal{E} \int_I \ell(t, x^n(t), u^n_t) \; dt. \tag{7}$$

By virtue of vague convergence of u^n to u^o, it is evident that for every $\varepsilon > 0$ there exists an integer $n_{1,\varepsilon}$ sufficiently large, such that the absolute value of the first term on the right hand side of equation (7) is less than $\varepsilon/2$ for all $n \geq n_{1,\varepsilon}$. By virtue of assumption (a1)-(a2), in particular the continuity of ℓ in x uniformly in U, it is easy to verify that there exists an integer $n_{2,\varepsilon}$ such that for all $n \geq n_{2,\varepsilon}$, the absolute value of the second term on the right hand side is less than $\varepsilon/2$. By combining these facts we obtain the following inequality

$$\mathcal{E} \int_I \ell(t, x^o(t), u^o_t) \; dt \leq \varepsilon + \int_I \ell(t, x^n(t), u^n_t) dt$$

for all $n \geq n_{1,\varepsilon} \bigvee n_{2,\varepsilon}$. Since $\varepsilon > 0$ is otherwise arbitrary, it follows from the above inequality that

$$\mathcal{E} \int_I \ell(t, x^o(t), u^o_t) \; dt \leq \liminf_n \mathcal{E} \int_I \ell(t, x^n(t), u^n_t) \; dt. \tag{8}$$

Combining (6) and (8) we arrive at the conclusion that $J(u^o) \leq \liminf_n J(u^n)$ thereby proving lower semi continuity of J in the vague topology. Since $\mathcal{U}_{ad}$ is compact in this vague topology, J attains its minimum on it. This proves the existence of an optimal control. ●

Note that the existence is proved under general conditions, irrespectively of whether the information structure to the control is full or partial.

4 Construction of a Hilbert Space of Semi Martingales

In the preceding section we have presented a result on existence of optimal controls. In the following section we consider the problem of characterizing optimal controls in the form of necessary conditions of optimality. For this we shall utilize martingale approach hence we need to consider certain fundamental properties of semi martingales. These properties are studied in this section. Before we consider such properties, we wish to provide the technical reasons for their study. Consider the system (1) with the cost functional (5) and the admissible controls $\mathcal{U}_{ad} \equiv L_\infty^a(I, \mathcal{M}_1(U))$ as described above. Recall that these are either $\mathcal{F}_t$ or $\mathcal{G}_t$-adapted probability measure valued random processes, depending on whether the information structure used to construct the controls is full or partial. For the necessary conditions of optimality we need stronger regularity properties for the drift and diffusion parameters $\{b, \sigma\}$ as well as the cost integrands $\{\ell, \Phi\}$. They are presented as follows:

(NC1): The triple $\{b, \sigma, \ell\}$ are measurable in $t \in I$, and the quadruple $\{b, \sigma, \ell, \Phi\}$ are once continuously differentiable with respect to the state variable $x \in R^n$. The first spatial derivatives of $\{b, \sigma\}$ are bounded uniformly on $I \times R^n \times U$.

Considering the Gateaux derivative of σ with respect to the state variable at the point $(t, z, \nu) \in I \times R^n \times M_1(U)$ in the direction $\eta \in R^n$ we have

$$\lim_{\varepsilon \to 0}(1/\varepsilon)(\sigma(t, z + \varepsilon\eta, \nu) - \sigma(t, z, \nu)) \equiv \sigma_x(t, z, \nu; \eta).$$

Note that $\eta \longrightarrow \sigma_x(t, z, \nu; \eta)$ is linear and it follows from the assumption (NC1) that there exists a finite positive number β such that

$$|\sigma_x(t, z, \nu; \eta)|_{\mathcal{L}(R^m, R^n)} \leq \beta|\eta|_{R^n}.$$

In order to present the necessary conditions of optimality we need the so called variational equation. Suppose $u^o \in \mathcal{U}_{ad}$ denote the optimal control and $u \in \mathcal{U}_{ad}$ any other control. Since $\mathcal{U}_{ad}$ is convex, for any $\varepsilon \in [0, 1]$, the control

$$u^\varepsilon \equiv u^o + \varepsilon(u - u^o) \in \mathcal{U}_{ad}.$$

Let $x^\varepsilon, x^o \in B_\infty^a(I, L_2(\Omega, R^n))$ denote the solutions of the system equation (1) corresponding to the controls u^ε and u^o respectively. Consider the limit

$$y \equiv \lim_{\varepsilon \downarrow 0}(1/\varepsilon)(x^\varepsilon - x^o).$$

We have the following result characterizing the process y.

Lemma 4.1 The process y is an element of the Banach space $B_\infty^a(I, L_2(\Omega, R^n))$ and it is the unique solution of the variational SDE

$$\begin{aligned} dy(t) &= b_x(t, x^o(t), u_t^o) \; y(t) \; dt + \sigma_x(t, x^o(t), u_t^o; y(t)) \; dW(t) \\ &\qquad + b(t, x^o(t), u_t - u_t^o) \; dt + \sigma(t, x^o(t), u_t - u_t^o) \; dW(t), \quad (9) \\ y(0) &= 0, \end{aligned}$$

having a continuous modification.

Proof. This is a linear SDE and so one can have a closed form solution. Indeed, considering the homogenous part given by

$$dz(t) = b_x(t, x^o(t), u_t^o) \, z(t) \, dt + \sigma_x(t, x^o(t), u_t^o; z(t)) \, dW(t), z(s) = \zeta, 0 \leq s \leq t < \infty,$$

it follows from the assumption (NC1) and Lemma 3.1 that it has a unique solution z given by

$$z(t) = \Psi(t, s)\zeta, t \geq s,$$

where $\Psi(t, s), 0 \leq s \leq t < \infty$ is the random ($\mathcal{F}_t$ measurable) transition operator for the homogenous system. Since the spatial derivatives of b and σ are uniformly bounded, the transition operator $\Psi(t, s), 0 \leq s \leq t \leq T$ is uniformly P almost surely bounded (with values in the space of $n \times n$ matrices). Considering the non homogenous system (9), the solution is then given by

$$y(t) = \int_0^t \Psi(t, s)d\eta(s) \tag{10}$$

where η is the semi martingale given by

$$d\eta(t) = b(t, x^o(t), u_t - u_t^o) \, dt + \sigma(t, x^o(t), u_t - u_t^o) \, dW(t), \eta(0) = 0. \tag{11}$$

Note that η is a continuous square integrable $\mathcal{F}_t$ semi martingale. This proves the existence, uniqueness and regularity property of the solutions of system (9). This is one approach. An alternate approach is the same as that of Lemma 3.1. Here one notes that the drift and the diffusion terms of equation (9) satisfy the basic assumptions of Lemma 3.1. So the existence of a solution follows from the Banach fixed point theorem as in lemma 3.1. The fact that it has continuous modification follows directly from the representation (10) and the continuity of the semi martingale η. ●

Later in the sequel we need certain important and interesting properties of semi martingales. Let $L_2^a(I, R^n) \subset L_2(I \times \Omega, R^n)$ denote the space of $\mathcal{F}_t$-adapted random processes $\{v(t), t \in I\}$ such that

$$\mathcal{E} \int_I |v(t)|_{R^n}^2 dt < \infty.$$

Similarly, let $L_2^a(I, \mathcal{L}(R^m, R^n)) \subset L_2(I \times \Omega, \mathcal{L}(R^m, R^n))$ denote the space of $\mathcal{F}_t$-adapted $n \times m$ matrix valued random processes $\{\Sigma(t), t \in I\}$ such that

$$\mathcal{E} \int_I |\Sigma(t)|_{\mathcal{L}(R^m, R^n)}^2 dt = \mathcal{E} \int_I tr(\Sigma^*(t)\Sigma(t))dt < \infty.$$

Since I is a finite interval , it is clear that $B_\infty^a(I, L_2(\Omega, R^n)) \subset L_2^a(I, R^n)$.

Definition 4.2 An R^n-valued random process $\{m(t), t \in I\}$ is said to be a square integrable continuous $\mathcal{F}_t$-semi martingale iff it is representable in the form

$$m(t) = m(0) + \int_0^t v(s)ds + \int_0^t \Sigma(s)dW(s), t \in I, \tag{12}$$

for some $v \in L_2^a(I, R^n)$ and $\Sigma \in L_2^a(I, \mathcal{L}(R^m, R^n))$ and for some R^n-valued $\mathcal{F}_0$ measurable random variable $m(0)$ having finite second moment.

We introduce the following class of $\mathcal{F}_t$-semi martingales:

$$\mathcal{SM}_0^2 \equiv \left\{ m : m(t) = \int_0^t v(s)ds + \int_0^t \Sigma(s)dW(s), t \in I, \right.$$

$$\left. for \ v \in L_2^a(I, R^n) \ \ and \ \Sigma \in L_2^a(I, \mathcal{L}(R^m, R^n)) \right\}. \qquad (13)$$

Now we present a fundamental result which has the potential of many other applications.

Theorem 4.3 The class $\mathcal{SM}_0^2$ is a real linear vector space and it is a Hilbert space with respect to the norm topology $\| m \|_{SM_0^2}$ arising from

$$\| m \|_{\mathcal{SM}_0^2}^2 \equiv \mathcal{E} \int_I |v(t)|_{R^n}^2 dt + \mathcal{E} \int_I tr(\Sigma^*(t)\Sigma(t))dt.$$

Further, the space $\mathcal{SM}_0^2$ is isometrically isomorphic to $L_2^a(I, R^n) \times L_2^a(I, \mathcal{L}(R^m, R^n))$, written as $\mathcal{SM}_0^2 \cong L_2^a(I, R^n) \times L_2^a(I, \mathcal{L}(R^m, R^n))$.

Proof Note that each $m \in \mathcal{SM}_0^2$ corresponds to a pair

$$(v, \Sigma) \in L_2^a(I, R^n) \times L_2^a(I, \mathcal{L}(R^m, R^n)).$$

We may call the pair (v, Σ) the infinitesimal generator (or simply the intensity) of the semi martingale m. Let $m_1 \in \mathcal{SM}_0^2$ corresponding to the intensity process $(v_1, \Sigma_1) \in L_2^a(I, R^n) \times L_2^a(I, \mathcal{L}(R^m, R^n))$ and $m_2 \in \mathcal{SM}_0^2$ corresponding to the intensity process $(v_2, \Sigma_2) \in L_2^a(I, R^n) \times L_2^a(I, \mathcal{L}(R^m, R^n))$ respectively. Clearly, $v_1 + v_2 \in L_2^a(I, R^n)$ and $\Sigma_1 + \Sigma_2 \in L_2^a(I, \mathcal{L}(R^m, R^n))$. Hence $m \equiv m_1 + m_2$, with intensity process $(v_1 + v_2, \Sigma_1 + \Sigma_2)$, is an element of $\mathcal{SM}_0^2$. For any real number α and any $m \in \mathcal{SM}_0^2$ with intensity process $(v, \Sigma) \in L_2^a(I, R^n) \times L_2^a(I, \mathcal{L}(R^m, R^n))$ we have $\alpha m \in \mathcal{SM}_0^2$ with intensity process $(\alpha v, \alpha \Sigma) \in L_2^a(I, R^n) \times L_2^a(I, \mathcal{L}(R^m, R^n))$. Thus $\mathcal{SM}_0^2$ is a linear vector space. We now furnish this with a scalar product and norm topology. Let $m_1, m_2 \in \mathcal{SM}_0^2$ with the intensity pairs $(v_1, \Sigma_1), (v_2, \Sigma_2)$ respectively and define

$$(m_1, m_2)_{\mathcal{SM}_0^2} \equiv \mathcal{E} \int_I (v_1(t), v_2(t))dt + \mathcal{E} \int_I tr(\Sigma_1^*(t)\Sigma_2(t))dt. \qquad (14)$$

The reader can easily verify that this gives a scalar product. Clearly taking $m_2 = m_1$ we have the norm square of m_1 given by

$$\| m_1 \|_{\mathcal{SM}_0^2}^2 = (m_1, m_1)_{\mathcal{SM}_0^2} \equiv \mathcal{E} \int_I |v_1(t)|_{R^n}^2 dt + \mathcal{E} \int_I |\Sigma_1(t)|_{\mathcal{L}(R^m, R^n)}^2 dt. \qquad (15)$$

It is easy to verify that the above expression defines a norm (modulo the null space). Thus $\mathcal{SM}_0^2$ is a scalar product space. To show that it is a Hilbert space, it suffices to verify that it is complete. Let $\{m_n\} \subset \mathcal{SM}_0^2$ be a Cauchy sequence corresponding to

the sequence of intensity pairs $\{(v_n, \Sigma_n)\} \subset L_2^a(I, R^n) \times L_2^a(I, \mathcal{L}(R^m, R^n))$. Let $p \geq 1$ and consider the expression

$$\| m_{n+p} - m_n \|_{\mathcal{SM}_0^2} = \left(\mathcal{E} \int_I |v_{n+p}(t) - v_n(t)|_{R^n}^2 dt + \mathcal{E} \int_I |\Sigma_{n+p}(t) - \Sigma_n|_{\mathcal{L}(R^m, R^n)}^2 dt \right)^{1/2}.$$

Since $\{m_n\}$ is a Cauchy sequence, $\lim_{n \to \infty} \| m_{n+p} - m_n \|_{\mathcal{SM}_0^2} = 0$ for every $p \geq 1$ and hence $\{(v_n, \Sigma_n)\}$ is a Cauchy sequence in $L_2^a(I, R^n) \times L_2^a(I, \mathcal{L}(R^m, R^n))$. But the later spaces are Hilbert and hence there exists a unique pair $(v_o, \Sigma_o) \in L_2^a(I, R^n) \times L_2^a(I, \mathcal{L}(R^m, R^n))$ to which (v_n, Σ_n) converges in norm (along a subsequence if necessary). Define the process m_0 by

$$m_0(t) = \int_0^t v_0(s) ds + \int_0^t \Sigma_0(s) dW(s), t \in I.$$

Clearly this is a semi martingale belonging to $\mathcal{SM}_0^2$ and it is the unique limit of the sequence of semi martingales $\{m_n\}$. This proves that $\mathcal{SM}_0^2$ is complete and hence a Hilbert space. Now we claim that for every $m \in \mathcal{SM}_0^2$ there exists a unique pair $(v, \Sigma) \in L_2^a(I, R^n) \times L_2^a(I, \mathcal{L}(R^m, R^n))$ such that

$$m(t) = \int_0^t v(s) ds + \int_0^t \Sigma(s) dW(s), t \in I.$$

Suppose this is false and there exists another pair $(v_1, \Sigma_1) \in L_2^a(I, R^n) \times L_2^a(I, \mathcal{L}(R^m, R^n))$ giving the same semi martingale m. This means that

$$0 = \int_0^t (v(s) - v_1(s)) ds + \int_0^t (\Sigma(s) - \Sigma_1(s)) dW(s), t \in I,$$

which is the same as

$$\int_0^t (v(s) - v_1(s)) ds = \int_0^t (\Sigma_1(s) - \Sigma(s)) dW(s), \quad t \in I.$$

But this is impossible since a martingale can never equal a function of bounded variation. Hence $v_1 = v$ and $\Sigma_1 = \Sigma$. Thus to every $m \in \mathcal{SM}_0^2$ there corresponds a unique pair $(v, \Sigma) \in L_2^a(I, R^n) \times L_2^a(I, \mathcal{L}(R^m, R^n))$ and conversely. The isometry follows from the expression (15). Hence $\mathcal{SM}_0^2 \cong L_2^a(I, R^n) \times L_2^a(I, \mathcal{L}(R^m, R^n))$. This completes the proof. •

5   Necessary Conditions of Optimality

Now we are prepared to develop the necessary conditions of optimality. The theory of relaxed controls is found to be a powerful technique for developing necessary conditions of optimality for deterministic systems [2], Theorem 8.3.5. Here we use the same technique for systems governed by stochastic differential equations driven by relaxed controls.

Below, we provide the main theorem. Later we use this result to derive a simplified minimum principle for both full as well as partial information.

Theorem 5.1 Consider the system (1) and the cost functional (5). An element $u^o \in \mathcal{U}_{ad}$, with the corresponding solution $x^o \in B_\infty^a(I, L_2(\Omega, R^n))$ to be optimal, it is necessary that there exists a semi martingale $m^o \in \mathcal{SM}_0^2$ with the intensity process $(\psi, Q) \in L_2^a(I, R^n) \times L_2^a(I, \mathcal{L}(R^m, R^n))$ such that the following inequality and the equations (SDE) hold:

$$(1) : \mathcal{E} \int_0^T \left\{ (b(t, x^o(t), u_t), \psi(t)) + tr(Q^*(t)\sigma(t, x^o(t), u_t)) + \ell(t, x^o(t), u_t) \right\} dt$$

$$\geq \mathcal{E} \int_0^T \left\{ (b(t, x^o(t), u_t^o), \psi(t)) + tr(Q^*(t)\sigma(t, x^o(t), u_t^o)) + \ell(t, x^o(t), u_t^o) \right\} dt$$

for all $\quad u \in \mathcal{U}_{ad}.$ \hfill (16)

$$(2) : dx^o(t) = b(t, x^o(t), u_t^o)dt + \sigma(t, x^o(t), u_t^o))dW(t)$$
$$x^o(0) = x_0 \hfill (17)$$

$$(3) : -d\psi(t) = b_x^*(t, x^o(t), u_t^o)\psi(t)dt + V_Q(t)dt + \ell_x(t, x^o(t), u_t^o)dt - Q(t)dW(t)$$
$$\psi(T) = \Phi_x(x^o(T)) \hfill (18)$$

where $V_Q \in L_2^a(I, R^n)$ is given by $(V_Q(t), \zeta) = tr(Q^*(t)\sigma_x(t, x^o(t), u_t^o; \zeta)), t \in I.$

Proof Suppose $u^o \in \mathcal{U}_{ad}$ is the optimal control and $u \in \mathcal{U}_{ad}$ any other control. Since $\mathcal{U}_{ad}$ is convex, for any $\varepsilon \in [0, 1]$, the control $u^\varepsilon \equiv u^o + \varepsilon(u - u^o) \in \mathcal{U}_{ad}$. Let $x^\varepsilon, x^o \in B_\infty^a(I, L_2(\Omega, R^n))$ denote the (strong) solutions of the system equation (1) corresponding to the controls u^ε and u^o respectively. Since u^o is optimal it is clear that

$$J(u^\varepsilon) - J(u^o) \geq 0 \hfill (19)$$

for all $\varepsilon \in [0, 1]$ and for all $u \in \mathcal{U}_{ad}$. Let $dJ(u^o, u - u^0)$ denote the Gateaux differential of J at u^o in the direction $u - u^o$. Dividing the expression (19) by ε and letting $\varepsilon \downarrow 0$ it is easy to verify that

$$dJ(u^o, u - u^0) = L(y) + \mathcal{E} \int_0^T \ell(t, x^o(t), u_t - u_t^o)dt \geq 0, \quad \forall \ u \in \mathcal{U}_{ad} \hfill (20)$$

where $L(y)$ is given by the functional

$$L(y) = \mathcal{E}\left\{ \int_0^T (\ell_x(t, x^o(t), u_t^o), y(t)) \ dt + (\Phi_x(x^o(T)), y(T)) \right\}. \hfill (21)$$

Since by Lemma 4.1, the process $y \in B_\infty^a(I, L_2(\Omega, R^n))$ and it is also continuous P-a.s it follows from assumption (a2) of Theorem 3.2 and the assumption (NC1), that $y \longrightarrow L(y)$ is a continuous linear functional. Further, by Lemma 4.1, $\eta \longrightarrow y$ is a continuous linear map from the Hilbert space $\mathcal{SM}_0^2$ to the B-space $B_\infty^a(I, L_2(\Omega, R^n))$

given by the expression (10). Thus the composition map $\eta \longrightarrow y \longrightarrow L(y) \equiv \tilde{L}(\eta)$ is a continuous linear functional on $\mathcal{SM}_0^2$. Then by virtue of the classical Riesz representation theorem for Hilbert spaces, there exists a semi martingale $\varrho \in \mathcal{SM}_0^2$ with intensity $(\psi, Q) \in L_2^a(I, R^n) \times L_2^a(I, \mathcal{L}(R^m, R^n))$ such that

$$L(y) \equiv \tilde{L}(\eta) = (\varrho, \eta)_{\mathcal{SM}_0^2} = \mathcal{E} \int_0^T (\psi(t), b(t, x^o(t), u_t - u_t^o))dt$$

$$+ \mathcal{E} \int_0^T tr(Q^*(t)\sigma(t, x^o(t), u_t - u_t^o))dt. \quad (22)$$

Substituting the expression (22) into the expression (20) we obtain

$$dJ(u^o, u - u^0) = \mathcal{E} \int_0^T (\psi(t), b(t, x^o(t), u_t - u_t^o))dt$$

$$+ \mathcal{E} \int_0^T tr(Q^*(t)\sigma(t, x^o(t), u_t - u_t^o))$$

$$+ \mathcal{E} \int_0^T \ell(t, x^o(t), u_t - u_t^o)dt \geq 0, \quad \forall \ u \in \mathcal{U}_{ad}. \quad (23)$$

The necessary condition given by the expression (16) readily follows from this. Equation (17) is the system equation along the optimal control state pair (u^o, x^o), so nothing to prove. We prove that the pair (ψ, Q) is given by the solution of the adjoint equation (18). Computing the Ito differential of the scalar product (y, ψ) we have the general expression

$$d(y(t), \psi(t)) = (dy(t), \psi(t)) + (y(t), d\psi(t)) + < dy(t), d\psi(t) > \quad (24)$$

where the last bracket denotes the classical quadratic variation term. Integrating this over $I = [0, T]$ and using the fact that $y(0) = 0$, it follows from the variational equation (9) that

$$\mathcal{E}(y(T), \psi(T)) = \mathcal{E}\left\{ \int_0^T (y(t), b_x^*\psi(t)dt + \sigma_x^*(\psi(t))dW(t) + d\psi(t)) \right.$$

$$\left. + \int_0^T (b^o, \psi(t))dt + \int_0^T ((\sigma^o)^*\psi(t), dW(t)) \right\} + \mathcal{E} \int_0^T < dy(t), d\psi(t) >, (25)$$

where for convenience of notation we have used

$$b_x \equiv b_x(t, x^o(t), u_t^o), \quad \sigma_x(\xi) \equiv \sigma_x(t, x^o(t), u_t^o; \xi), \xi \in R^n,$$
$$b^o = b(t, x^o(t), u_t - u_t^o), \quad \sigma^o \equiv \sigma(t, x^o(t), u_t - u_t^o).$$

Note that the stochastic integrals in (25) equal zero and hence make no contribution. This follows from the facts that $\sigma_x^*(\psi(t)) \in L_2^a(I, \mathcal{L}(R^m, R^n))$ and $(\sigma^o)^*\psi \in L_2^a(I, R^m)$ as seen later. So we can eliminate them giving the following expression

$$\mathcal{E}(y(T), \psi(T)) = \mathcal{E}\left\{ \int_0^T (y(t), b_x^*\psi(t)dt + d\psi(t)) + \int_0^T (b^o, \psi(t))dt \right\}$$

$$+ \mathcal{E} \int_0^T < dy(t), d\psi(t) > . \quad (26)$$

Before we consider the quadratic variation term, let us recall that the Ito derivatives of the variation process y and the adjoint process ψ are of the following form:

$$
\begin{aligned}
dy(t) &= \text{ bounded variation terms} + \sigma_x(t, x^o(t), u_t^o; y(t))dW(t) \\
&+ \sigma(t, x^o(t), u_t - u_t^o)dW(t), \\
d\psi(t) &= \text{ bounded variation terms} + Q(t)dW(t).
\end{aligned}
$$

Considering now the quadratic variation term it is easy to verify that

$$
\mathcal{E} \int_0^T < dy(t), d\psi(t) > = \mathcal{E} \int_0^T \{tr(Q^*(t)\sigma_x(y)) + tr(Q^*(t)\sigma^o)\}dt. \tag{27}
$$

Clearly, the first term on the right hand side of the above expression is linear in y. Thus there exists a process $V_Q(t), t \in I$, given by the following expression

$$
(V_Q(t), y(t)) \equiv tr(Q^*(t)\sigma_x(y)) \equiv tr(Q^*(t)\sigma_x(t, x^o(t), u_t^o; y(t))). \tag{28}
$$

By assumption (NC1), σ has uniformly bounded spatial first derivative and it follows from the semi martingale representation Theorem 4.3 that $Q \in L_2^a(I, \mathcal{L}(R^m, R^n))$ and hence $V_Q \in L_2^a(I, R^n)$. Substituting (28) into (27) and then (27) into (26), we obtain

$$
\mathcal{E}(y(T), \psi(T)) = \mathcal{E}\Bigg\{ \int_0^T (y(t), b_x^*\psi(t)dt + V_Q(t)dt - Q(t)dW(t) + d\psi(t)) \\
+ \int_0^T (b^o, \psi(t))dt + tr(Q^*(t)\sigma^o)dt \Bigg\}. \tag{29}
$$

By setting

$$
\begin{aligned}
b_x^*(t, x^o(t), u_t^o)\psi(t)dt + V_Q(t)dt - Q(t)dW(t) + d\psi(t) &= -\ell_x(t, x^o(t), u_t^o)dt \\
\psi(T) &= \Phi_x(x^o(T)),
\end{aligned} \tag{30}
$$

it follows from (29) and the expression for the functional L given by (21) that

$$
\begin{aligned}
L(y) &= \mathcal{E}(y(T), \psi(T)) + \mathcal{E} \int_0^T (y(t), \ell_x(t, x^o(t), u_t^o))dt \\
&= \mathcal{E} \int_0^T \{(b(t, x^o(t), u_t - u_t^o)\psi(t)) + tr(Q^*(t)\sigma(t, x^o(t), u_t - u_t^o))\}dt. \tag{31}
\end{aligned}
$$

This is precisely what was obtained by the semi martingale argument giving (22). Thus the pair (ψ, Q) must satisfy the backward stochastic differential equation (30) which is precisely the adjoint equation given by (18) as stated. Since ψ satisfies the stochastic differential equation and T is finite, it follows from the classical theory of Ito differential equations that ψ is actually an element of $B_\infty^a(I, L_2(\Omega, R^n)) \subset L_2^a(I, R^n)$. In other words, ψ is more regular than predicted by semi martingale theory. Hence by

our assumption on σ it is easy to verify that $\sigma_x^*(\psi) \in L_2^a(I, \mathcal{L}(R^m, R^n))$ and $(\sigma^o)^*\psi \in L_2^a(I, R^m)$ as stated before. Thus we have completed the proof. $\bullet$

Remark 5.2 Define the Hamiltonian

$$H : I \times R^n \times R^n \times \mathcal{L}(R^m, R^n) \times \mathcal{M}_1(U) \longrightarrow R$$

by

$$H(t, \xi, \zeta, M, \nu) = (b(t, \xi, \nu), \zeta) + tr(M^*\sigma(t, \xi, \nu)) + \ell(t, \xi, \nu).$$

In terms of this Hamiltonian, the necessary conditions of optimality (16)-(18) can be written compactly as follows

$$\mathcal{E} \int_0^T H(t, x^o(t), \psi(t), Q(t), u_t)dt \geq \mathcal{E} \int_0^T H(t, x^o(t), \psi(t), Q(t), u_t^o)dt$$
$$\text{for all } u \in \mathcal{U}_{ad}, \tag{32}$$

where the triple $\{x^o, \psi, Q\}$ is the unique solution of the following Hamiltonian system

$$dx^o(t) = H_\psi(t, x^o(t), \psi(t), Q(t), u_t^o)dt + \sigma(t, x^o(t), u_t^o)dW(t), \quad x^o(0) = x_0, \tag{33}$$

$$d\psi(t) = -H_x(t, x^o(t), \psi(t), Q(t), u_t^o)dt + Q(t)dW(t), \quad \psi(T) = \Phi_x(x^o(T)). \tag{34}$$

Note the similarity in appearance with the Pontryiagin minimum principle. In fact we recover the Pontryagin minimum principle for relaxed controls in [5, 2] by setting $\sigma = 0$.

For controls based on full-information which are $\mathcal{F}_t$ adapted, and under the condition that $\{\mathcal{F}_t, t \in [0, T]\}$ is the natural filtration generated by the Brownian motion $\{W(t), t \in [0, T]\}$, augmented by all $P-$null sets in $\mathcal{F}$, given by the inequality (16) (or equivalently (32)) is equivalent to the following point wise almost sure inequality (the derivation is similar to that of Corollary 5.3):

$$H(t, x^o(t), \psi(t), Q(t), \mu) \geq H(t, x^o(t), \psi(t), Q(t), u_t^o),$$
$$\forall \mu \in \mathcal{M}_1(U), \ a.e. \ t \in [0, T], \ P - a.s.$$

or equivalently,

$$H(t, x^o(t), \psi(t), Q(t), u_t^o) = \min_{\mu \in \mathcal{M}_1(U)} H(t, x^o(t), \psi(t), Q(t), \mu),$$
$$a.e. \ t \in [0, T], \ P - a.s.$$

subject to the Hamiltonian system (33)-(34).

For the partial information case, the point wise necessary conditions of optimality for controls are given in the next Corollary.

Corollary 5.3 Suppose the assumptions of Theorem 5.1 hold and consider controls which are $\mathcal{G}_t$ adapted. Then the inequality (16) (or equivalently (32)) is equivalent to the following point wise almost sure inequality with respect to the σ-algebra $\mathcal{G}_t \subset \mathcal{F}_t$:

$$\mathcal{E}\{H(t, x^o(t), \psi(t), Q(t), \mu)|\mathcal{G}_t\} \geq \mathcal{E}\{H(t, x^o(t), \psi(t), Q(t), u_t^o)|\mathcal{G}_t\} \tag{35}$$

for all $\mu \in \mathcal{M}_1(U)$, $a.e.t \in [0, T]$, $P-a.s.$ subject to the Hamiltonian system (33)-(34).

Proof. Since the admissible controls are vaguely $\mathcal{G}_t$ measurable, we can rewrite the inequality (32) in the following equivalent form,

$$\mathcal{E} \int_0^T \mathcal{E}\{H(t, x^o(t), \psi(t), Q(t), u_t)|\mathcal{G}_t\} \, dt$$

$$\geq \mathcal{E} \int_0^T \mathcal{E}\{H(t, x^o(t), \psi(t), Q(t), u_t^o)|\mathcal{G}_t\} \, dt. \tag{36}$$

Let $t \in (0, T)$, $\omega \in \Omega$ and $\varepsilon > 0$ and consider the sets $I_\varepsilon \equiv [t, t + \varepsilon] \subset I$ and $\Omega_\varepsilon (\subset \Omega) \in \mathcal{G}_t$ containing ω such that $|I_\varepsilon| \to 0$ and $P(\Omega_\varepsilon) \to 0$ as $\varepsilon \to 0$. For any subsigma algebra $\mathcal{G} \subset \mathcal{F}$, let $P_\mathcal{G}$ denote the restriction of the probability measure P on to the σ-algebra $\mathcal{G}$. For any (vaguely) $\mathcal{G}_t$-measurable $\nu \in \mathcal{M}_1(U)$, construct the control

$$u_t = \begin{cases} \nu & \text{for } (t, \omega) \in I_\varepsilon \times \Omega_\varepsilon \\ u_t^o & \text{otherwise.} \end{cases}.$$

Clearly, it follows from the above construction that $u \in \mathcal{U}_{ad}$. Using this control in (35) we obtain the following inequality

$$\int_{\Omega_\varepsilon \times I_\varepsilon} \mathcal{E}\{H(t, x^o(t), \psi(t), Q(t), \nu)|\mathcal{G}_t\} \, dt \geq \int_{\Omega_\varepsilon \times I_\varepsilon} \mathcal{E}\{H(t, x^o(t), \psi(t), Q(t), u_t^o)|\mathcal{G}_t\} \, dt. \tag{37}$$

Letting $|I_\varepsilon|$ denote the Lebesgue measure of the set I_ε and dividing the above expression by the product measure $P(\Omega_\varepsilon)|I_\varepsilon|$ and letting $\varepsilon \to 0$ we arrive at the following in equality,

$$\mathcal{E}\{H(t, x^o(t), \psi(t), Q(t), \nu)|\mathcal{G}_t\} \geq \mathcal{E}\{H(t, x^o(t), \psi(t), Q(t), u_t^o)|\mathcal{G}_t\}$$

which holds for almost all $t \in I$ and $P_{\mathcal{G}_t}$ almost all $\omega \in \Omega$. Thus we have completed the proof. ●

Remark 5.4 Define

$$g_t(\xi) \equiv \mathcal{E}\{H(t, x^o(t), \psi(t), Q(t), \xi)|\mathcal{G}_t\}, t \in I, \xi \in U.$$

The reader can easily verify from the basic assumptions on the parameters $\{b, \sigma, \ell, \Phi\}$ that the random process g is an element of $L_1^a(I, C(U))$ and that it is adapted to the σ-algebra $\mathcal{G}_t$. Clearly, the necessary condition given by the inequality (35) can be written as follows

$$\int_U g_t(\xi)\mu(d\xi) \geq \int_U g_t(\xi)u_t^o(d\xi),$$

and this must hold for all $\mathcal{M}_1(U)$-valued $\mathcal{G}_t$-adapted (vaguely measurable) random variables μ. Define

$$\Lambda_t(\mu) \equiv \int_U g_t(\xi)\mu(d\xi).$$

This is a $\mathcal{G}_t$- measurable continuous linear functional on $\mathcal{M}_1(U)$. Since the later space is vaguely compact, it attains its minimum on $\mathcal{M}_1(U)$ and from the above inequality it follows that u_t^o is one such element. Because the functional Λ_t is not strictly convex there may be multiplicities of minima $M^o(t)$. It is easy to verify that the set

$$M^o(t) \equiv \{\mu \in M_1(U) : \mu \text{ is } \mathcal{G}_t - \text{measurable and } \Lambda_t(\mu) = \Lambda_t(u_t^o)\}$$

is convex and a vaguely (weak star) closed subset of $\mathcal{M}_1(U)$ and hence vaguely compact. Thus $t \longrightarrow M^o(t)$ is a measurable multi function with convex compact values in $\mathcal{M}_1(U)$. By our assumption U is compact and hence $\mathcal{M}_1(U)$ is a compact Polish space and hence a compact Souslin space. Thus it follows from the well known Yankov-Von Neumann-Auman selection theorem [[26], Theorem 2.14, p158] that the multi function $t \longrightarrow M^o(t)$ has a $\mathcal{G}_t$ measurable selection. Hence we have a $\mathcal{G}_t$ measurable optimal relaxed control.

6   Extension to Jump Processes

The necessary conditions of optimality given in the previous section can be easily extended to control problems involving stochastic differential equations driven both by Brownian motion and Leṽy process or Poisson jump process. Let $Z \equiv R^n \setminus \{0\}$ and $\mathcal{B}(Z)$ the Borel algebra of subsets of the set Z. Let $p(dv \times dt)$ denote the Poisson counting measure on $\mathcal{B}(Z) \times \sigma(I)$. Physical interpretation of this measure is simple. For each $\Gamma \in \mathcal{B}(Z)$ and any interval $\Delta \in \sigma(I)$, $p(\Gamma \times \Delta)$ gives the number of jumps over the interval Δ of sizes confined in Γ. This is a Poisson random variable with mean $\mathcal{E}p(\Gamma \times \Delta) = \pi(\Gamma)\lambda(\Delta)$ where λ is the Lebesgue measure on the real line and π is the Leṽy measure on Z. Here π is a countably additive bounded positive measure. The compensated Poisson random measure is given by

$$q(dv \times dt) = p(dv \times dt) - \pi(dv)dt.$$

There is no loss of generality considering the compensated Poission random measure in modeling SDE. As usual, we assume that all the random processes considered in this paper are based on the filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_{t\geq0}, P)$ where $\{\mathcal{F}_t, t \geq 0\}$ is an increasing family of subsigma algebras of σ-algebra $\mathcal{F}$ and that they are right continuous with left limits. A controlled stochastic differential equation driven both by Brownian motion and the compensated jump process described above is given by the following stochastic differential equation

$$dx(t) = b(t, x(t), u_t)dt + \sigma(t, x(t), u_t)dW(t) + \int_Z C(t, x(t), v, u_t)q(dv \times dt), t \in I \ (38)$$

for $x(0) = x_0$. Throughout the rest of the paper it is assumed without any further notice that $\{x_0, W, q\}$ are independent random elements. Again our controls are

relaxed controls which, for the partial information case, are weakly $\mathcal{G}_t$ adapted $\mathcal{M}_1(U)$ valued random processes denote by $\mathcal{U}_{ad}$. The cost functional is given by

$$J(u) \equiv \mathcal{E}\left\{ \int_I \ell(t, x(t), u_t)dt + \Phi(x(T)) \right\}. \tag{39}$$

Objective is to find a control from the admissible set $\mathcal{U}_{ad}$ at which the functional (39) attains its minimum. The method of proof of the necessary conditions of optimality for this model is no different from the one given for the continuous case. Hence we present the results without repeating the detailed proof.

For the problem involving jump process, we introduce the following Hilbert space of discontinuous square integrable semi martingales denoted by $\mathcal{DSM}_0^2$ and this is given by

$$\mathcal{DSM}_0^2 \equiv \left\{ m : m(t) = \int_0^t v(s)ds + \int_0^t Q(s)dW(s) + \int_0^t \int_Z \varphi(v,t)q(dv \times dt) \right.$$
$$\left. v \in L_2^a(I, R^n), Q \in L_2^a(I, \mathcal{L}(R^m, R^n)), \varphi \in L_2^a(I, L_2^n(Z, \pi)) \right\} \tag{40}$$

where $L_2^n(Z, \pi)$ denotes the Hilbert space of R^n-valued functions defined on Z which are square integrable with respect to the Levy measure π. In this case the norm topology is given by

$$\| m \|_{\mathcal{DSM}_0^2} = \left(\mathcal{E} \int_I |v(t)|_{R^n}^2 dt + \mathcal{E} \int_I tr(Q^*(t)Q(t))dt + \mathcal{E} \int_I \int_Z |\varphi(v,t)|_{R^n}^2 \pi(dv)dt \right)^{1/2}. \tag{41}$$

Now we are prepared to present the necessary conditions of optimality. Before we do so we need the following assumptions for C.

The function $C : I \times R^n \times R^n \times U \longrightarrow R^n$ is measurable in t on I and continuous in the rest of the arguments satisfying, uniformly with respect to $\xi \in U$, the following assumptions

$$(A6): \qquad \left(\int_Z |C(t,x,v,\xi)|_{R^n}^2 \pi(dv) \right)^{1/2} \leq K(t)(1 + |x|_{R^n})$$

$$(A7): \qquad \left(\int_Z |C(t,x,v,\xi) - C(t,y,v,\xi)|_{R^n}^2 \pi(dv) \right)^{1/2} \leq K(t)(|x-y|_{R^n}).$$

Theorem 6.1 Consider the system (38) with the cost functional (39) and the admissible controls $\mathcal{U}_{ad}$. Suppose $\{b, \sigma, C\}$ satisfy the assumptions (A1)-(A7) and that their first derivatives with respect to the state variable $x \in R^n$ are uniformly bounded. An element $u^o \in \mathcal{U}_{ad}$, with the corresponding solution $x^o \in B_\infty^a(I, L_2(\Omega, R^n))$ to be optimal, it is necessary that there exists a semi martingale $m^o \in \mathcal{DSM}_0^2$ with the intensity process $(\psi, Q, \varphi) \in L_2^a(I, R^n) \times L_2^a(I, \mathcal{L}(R^m, R^n)) \times L_2^a(I, L_2^n(Z, \pi))$ such that the following inequality and the stochastic differential equations hold:

$$(1): \mathcal{E} \int_0^T \left\{ (b(t, x^o(t), u_t - u_t^o), \psi(t)) + tr(Q^*(t)\sigma(t, x^o(t), u_t - u_t^o)) \right.$$
$$\left. + \int_Z (C(t, x^o(t), v, u_t - u_t^o), \varphi(t, v))\pi(dv) + \ell(t, x^o(t), u_t - u_t^o) \right\} dt \geq 0 \ (42)$$

for all $u \in \mathcal{U}_{ad}$.

$$(2): dx^o(t) = b(t, x^o(t), u_t^o)dt + \sigma(t, x^o(t), u_t^o)dW(t) + \int_Z C(t, x^o(t), v, u_t^o)q(dv \times dt)$$
$$x^o(0) = x_0 \tag{43}$$

$$(3): \ -d\psi(t) = b_x^*(t, x^o(t), u_t^o)\psi(t)dt + V_Q(t)dt - Q(t)dW(t)$$
$$+ \int_Z C_x^*(t, x^o(t), v, u_t^o)\varphi(t, v)\pi(dv)dt - \int_Z \varphi(t, v)q(dv \times dt) + \ell_x(t, x^o(t), u_t^o)dt$$
$$\psi(T) = \Phi_x(x^o(T)) \tag{44}$$

where $V_Q \in L_2^a(I, R^n)$ is given by $(V_Q(t), \zeta) = tr(Q^*(t)\sigma_x(t, x^o(t), u_t^o; \zeta)), t \in I$.

Remark 6.2 Define the Hamiltonian

$$H: I \times R^n \times R^n \times \mathcal{L}(R^m, R^n) \times L_2^n(Z, \pi) \times \mathcal{M}_1(U) \longrightarrow R$$

by the following expression

$$H(t, x, \psi, Q, \varphi, \mu) \equiv (b(t, x, \mu), \psi) + tr(Q^*\sigma(t, x, \mu))$$
$$+ \int_Z (C(t, x, v, \mu), \varphi(t, v))_{R^n}\pi(dv) + \ell(t, x, \mu), \tag{45}$$

where $\varphi \in L_2^n(Z, \pi)$. We write $t \longrightarrow \varphi(t)$ for the $L_2^n(Z, \pi)$ valued function. In terms of this Hamiltonian, the necessary conditions of Theorem 6.1 can be written in the following canonical form:

$$\mathcal{E} \int_I H(t, x^o(t), \psi(t), Q(t), \varphi(t), u_t)dt \geq \mathcal{E} \int_I H(t, x^o(t), \psi(t), Q(t), \varphi(t), u_t^o)dt$$
$$\text{for all } u \in \mathcal{U}_{ad}, \tag{46}$$
$$dx^o(t) = H_\psi dt + \sigma(t, x^o(t), u_t^o)dW(t)$$
$$+ \int_Z C(t, x^o(t), v, u_t)q(dv \times dt), \quad x^o(0) = x_0 \tag{47}$$

$$d\psi(t) = -H_x dt + Q(t)dW(t) + \int_Z \varphi(t, v)q(dv \times dt), \quad \psi(T) = \Phi_x(x^o(T)). \tag{48}$$

Similarly as before, one can also obtain point wise almost sure variational inequalities.

7   Necessary conditions with Regular Controls

In the development of the necessary conditions of optimality given in the preceding two sections we have tacitly used the existence Theorem 3.2 which asserts the existence of optimal controls from the class of relaxed controls $\mathcal{U}_{ad}$. Let $L_\infty^a(I \times \Omega, U)$ denote the class of $\mathcal{G}_t$ adapted random processes defined on the interval I and taking values from the closed bounded set $U \subset R^d$. This is the class of regular controls and we denote this by $\mathcal{U}^r$. It is clear that this embeds continuously into the class of relaxed controls through the map $u \ni \mathcal{U}^r \longrightarrow \delta_{u(t,\omega)} \in \mathcal{U}_{ad}$. Clearly, for every $\vartheta \in L_1^a(I \times \Omega, C(U))$

$$\mathcal{E} \int_{I \times U} \vartheta(t, \omega, \xi) \delta_{u(t,\omega)}(d\xi) dt = \mathcal{E} \int_I \vartheta(t, \omega, u(t, \omega)) dt. \tag{49}$$

Theorem 7.1 Consider the class of regular controls $\mathcal{U}^r$ with U assumed to be closed bounded and convex. Suppose Theorem 3.2 holds for regular controls in the sense that an optimal control exists from the class $\mathcal{U}^r$. Then all the necessary conditions involving relaxed controls (Theorem 5.1, Theorem 6.1) reduce to the classical minimum principle for stochastic systems.

Proof. The proof is direct. In fact it follows from straightforward application of the embedding mentioned above and the definition (49). Considering the necessary conditions of optimality given by Theorem 5.1, and using the embedding mentioned above it is easy to derive the following necessary conditions of optimality

$$(1) : \mathcal{E} \int_0^T \left\{ (b(t, x^o(t), u_t), \psi(t)) + tr(Q^*(t)\sigma(t, x^o(t), u_t)) + \ell(t, x^o(t), u_t) \right\} dt$$

$$\geq \mathcal{E} \int_0^T \left\{ (b(t, x^o(t), u_t^o), \psi(t)) + tr(Q^*(t)\sigma(t, x^o(t), u_t^o)) + \ell(t, x^o(t), u_t^o) \right\} dt,$$

$$\text{for all} \quad u \in \mathcal{U}^r. \tag{50}$$

$$(2) : dx^o(t) = b(t, x^o(t), u_t^o) dt + \sigma(t, x^o(t), u_t^o)) dW(t)$$
$$x^o(0) = x_0 \tag{51}$$

$$(3) : -d\psi(t) = b_x^*(t, x^o(t), u_t^o)\psi(t) dt + V_Q(t) dt + \ell_x(t, x^o(t), u_t^o) dt - Q(t) dW(t)$$
$$\psi(T) = \Phi_x(x^o(T)) \tag{52}$$

where $V_Q \in L_2^a(I, R^n)$ is given by $(V_Q(t), \zeta) = tr(Q^*(t)\sigma_x(t, x^o(t), u^o(t); \zeta)), t \in I$.

Remark 7.2 Using precisely similar arguments for the SDE with jumps, one can obtain the minimum principle for regular controls from those of relaxed controls given by Theorem 6.1.

8 Realizability of Relaxed Controls by Regular Controls

We proved existence of optimal relaxed controls in Theorem 3.2 without requiring convexity of the control domain U. In any application it is much easier to construct regular controls. So one may be interested to find a regular control corresponding to which the performance of the system is close to that realized by optimal relaxed control. In this regard we have the following result.

Theorem 8.1 Consider the regular controls $\mathcal{U}^r$ with U closed bounded but not necessarily convex as in Theorem 7.1. Suppose the basic assumptions of Lemma 3.1 and Theorem 3.2 hold and consider the control problem as stated in Theorem 3.2. Further, suppose that $x \longrightarrow \Phi(x)$ is continuous. Let $u^o \in \mathcal{U}_{ad}$ be the optimal relaxed control. Then, for every $\varepsilon > 0$ there exists a regular control $u_r \in \mathcal{U}^r$ such that

$$J(u_r) \leq \varepsilon + J(u^o).$$

Proof Since $\mathcal{U}_{ad} \equiv L_\infty^a(I \times \Omega, \mathcal{M}_1(U)) \subset L_\infty^a(I \times \Omega, \mathcal{M}(U))$ is compact in the vague topology (that is weak star topology) and convex (because $\mathcal{M}_1(U)$ is convex, it follows from the well known Krein-Millman theorem that

$$\mathcal{U}_{ad} = cl^v conv(\text{ext}(\mathcal{U}_{ad})),$$

that is, $\mathcal{U}_{ad}$ is the weak star closed convex hull of its extreme points. Considering the embedding $\mathcal{U}^r \hookrightarrow \mathcal{U}_{ad}$ as mentioned above, it is easy to verify that the extreme points of $\mathcal{U}_{ad}$ are precisely the set of regular controls $\mathcal{U}^r$ through the map $u \ni \mathcal{U}^r \longrightarrow \delta_u \in \mathcal{U}_{ad}$. Thus, if $u^o \in \mathcal{U}_{ad}$ is the optimal (relaxed) control there exists a sequence $\{u^n\}$ of the form

$$u^n \equiv \sum_{i=1}^n \alpha_i^n u_i, u_i \in \mathcal{U}^r, \alpha_i^n \geq 0, \sum_{i=1}^n \alpha_i^n = 1, n \in N$$

such that $u^n \xrightarrow{v} u^o$. Let $\{x^n, x^o\} \subset B_\infty^a(I, L_2(\Omega, R^n))$ denote the solutions of the system equation (1) corresponding to the controls $\{u^n, u^o\}$ respectively. Then it follows from Lemma 3.1 that, along a subsequence if necessary, $x^n \xrightarrow{s} x^o$ in $B_\infty^a(I, L_2(\Omega, R^n))$. Consequently, it follows from continuity of ℓ and Φ in the state variable x and the assumptions (a1)-(a3) and Lebesgue dominated convergence theorem that $\lim_{n\to\infty} J(u^n) = J(u^o)$. Note that for every $n \in N$, $u^n \in \mathcal{U}^r$, and so, for every $\varepsilon > 0$, there exists an $n_\varepsilon \in N$ such that $|J(u^n) - J(u^o)| < \varepsilon$ for all $n \geq n_\varepsilon$. Taking $u_r = u^{n_\varepsilon}$ we have $J(u_r) \leq \varepsilon + J(u^o)$. This completes the proof. ●

Remark 8.2 In view of the above result it is evident that an ϵ-optimal control can be found from the class of regular controls (measurable functions with values in U) though the limit of such controls may be a relaxed control. More specifically if $U \subset R^d$ consists of a finite set of points, it is clearly non-convex, and optimal control may not exist from the class of regular controls $\mathcal{U}^r$ based on the set U. However, optimal relaxed controls do exist. In this case the sequence of regular controls approximating the optimal relaxed control may oscillate violently between the finite set of points of U with increasing frequency (converging to infinity). This is known as chattering.

References

[1] L. Cesari, *Optimization Theory and Applications.* Springer-Verlag, 1983.

[2] N. U. Ahmed, *Dynamic Systems and Control with Applications.* World Scientific, New Jersey London, Singapore Beijing Shanghai, Hong Kong, Taipei, Chenna, 2006.

[3] N. U. Ahmed and K. L. Teo, *Optimal Control of Distributed Parameter Systems.* Elsevier North Holland, New York, Oxford, 1981.

[4] H. O. Fattorini, *Infinite Dimensional Optimization and Control Theory.* Encyclopedia of Mathematics and Its Applications, 62, Cambridge University Press., 1999.

[5] N. U. Ahmed, "Optimal relaxed controls for systems governed by impulsive differential inclusions," *Nonlinear Functional Analysis & Applications*, vol. 10, no. 3, pp. 427–460, 2005.

[6] N. U. Ahmed and C. D. Charalambous, "Minimax games for stochastic systems subject to relative entropy uncertainty: Applications to SDE's on Hilbert spaces," *Journal of Mathematics of Control, Signals and Systems*, vol. 19, pp. 65–91, 2001.

[7] H. J. Kushner, "Necessary conditions for continuous parameter stochastic optimization problems," *SIAM Journal on Control and Optimization,*, 1972.

[8] U. Haussmann, *A Stochastic Maximum Principle for Optimal Control of Diffusions*, ser. Longman Sci. & Tech., Harlow, UK. Pitman Research Notes in Mathematics, 1986, vol. 151.

[9] J. Bismut, "An introductory approach to duality in optimal stochastic contro," *SIAM Review*, vol. 30, pp. 62–78, 1978.

[10] A. Bensoussan, *Lecture on Stochastic Control*, ser. Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1982.

[11] ——, "Stochastic maximum principle for distributed parameter systems," *Journal of the Franklin Institute*, vol. 315, pp. 387–406, 1999.

[12] S. Peng, "A general stochastic maximum principle for optimal control problems," *SIAM Journal on Control and Optimization*, vol. 28, no. 4, pp. 966–979, 1990.

[13] A. Cadenillas and I. Karatzas, "The stochastic maximum principle for linear convex optimal control with random coefficients," *SIAM Journal on Control and Optimization*, 1995.

[14] R. Elliott and M. Kohlmann, "The second order minimum principle and adjoint process," *Stochastics & Stochastic Reports,*, vol. 46, pp. 25–39, 1994.

[15] X. Y. Zhou, "A unified treatment of maximum principle and dynamic programming in stochastic controls," *Stochastics & Stochastic Reports*, vol. 36, pp. 137–161, 1991.

[16] J. Yong and X. Zhou, *Stochastic Controls, Hamiltonian Systems and HJB Equations.*   Springer-Verlag, 1999.

[17] B. Mezerdi and S. Bahlali, "Necessary conditions for optimality in relaxed stochastic control problems," *Stochastics & Stochastic Reports*, vol. 73, pp. 201–218, 2002.

[18] B. M. S. Bahlali and B. Djehiche, "Approximation and optimality necessary conditions in relaxed stochastic control problems," *Journal of Applied Mathematics and Stochastic Analysis*, p. 23, 2006, iD 72762.

[19] B. D. S. Bahlali and B. Mezerdi, "The relaxed stochastic maximum principle in singular optimal control of diffusions," *SIAM Journal on Control and Optimization*, vol. 46, pp. 427–444, 2007.

[20] K. D. S. Bahlali and B. Mezerdi, "On the stochastic maximum principle in optimal control of degenerate diffusions with lipschitz coefficients," *Applied Mathematics and Optimization*, vol. 56, pp. 364–378, 2007.

[21] S. Bahlali, "Necessary and sufficient optimality coditions for relaxed and strict control problems," *SIAM Journal on Control and Optimization*, 2008.

[22] B. D. R. Buckdahn and J. Li, "A general stochastic maximum principle for sde's of mean-field type," *Applied Mathematics and Optimization*, vol. 64, no. 197-216, 2011.

[23] L. Zhang and Y. Shi, "Maximum principle for forward-backward doubly stochastic control systems and applications," *ESAIM: Control,Optimization and Calculus of Variations*, vol. COCV 17, pp. 1174–1197, 2011, dOI: 10,1051/cocv/2010042.

[24] R. J. E. X. Zhang and T. K. Siu, "A stochastic maximum principle for a markov regime-switching jump-diffusion model and its applications to finance," *SIAM Journal on Control and Optimization*, vol. 50, no. 2, pp. 964–990, 2012.

[25] C. D. Charalambous and J. L. Hibey, "Minimum principle for partially observable nonlinear risk-sensitive control problems using measure-valued decompositions," *Stochastics & Stochastic Reports*, 1996.

[26] S. Hu and N. S. Papageorgiou, *Handbook of Multivalued Analysis.*   Kluwer Academic Publishers, Dordrecht, Boston, London., 1997.