New Approach to Stochastic Optimal Control

R. Josa-Fombellida · J.P. Rincón-Zapatero

Abstract This paper provides new insights into the solution of optimal stochastic control problems by means of a system of partial differential equations, which characterize directly the optimal control. This new system is obtained by the application of the stochastic maximum principle at every initial condition, assuming that the optimal controls are smooth enough. The type of problems considered are those where the diffusion coefficient is independent of the control variables, which are supposed to be interior to the control region.

Keywords Optimal stochastic control · Itô's formula · Hamilton–Jacobi–Bellman equation · Semilinear parabolic equation

1 Introduction

Three major approaches in stochastic optimal control can be differentiated: dynamic programming, duality and the maximum principle.

Dynamic programming obtains, by means of the optimality principle of Bellman, the Hamilton–Jacobi–Bellman equation, which characterizes the value function; see [1–5]. Under some smoothness and regularity assumptions on the solution, it is pos-

R Josa-Fombellida
Departamento de Estadística e Investigación Operativa, Universidad de Valladolid, Valladolid, Spain
e-mail: ricar@eio uva es

J P Rincón-Zapatero (✉)
Departamento de Economía, Universidad Carlos III de Madrid, Getafe, Madrid, Spain
e-mail: jrincon@eco uc3m es

sible to obtain, at least implicitly, the optimal control. This is the content of the so-called *verification theorems* which appear in Fleming and Rishel [1] or Fleming and Soner [3]. However, the problem of recovering the optimal control from the gradient of the value function by means of solving a static optimization problem remains and this can be difficult to do.

Duality methods, also known in stochastic control theory as the Martingale approach, have become very popular in recent years because they provide powerful tools for the study of some classes of stochastic control problems. Martingale methods are particularly useful for problems appearing in finance, such as the model of Merton [6]. Duality reduces the original problem to one of finite dimension. The approach is based on the martingale representation theorem and the Girsanov transformation. We refer the reader to Bismut [7, 8] and the monograph by Duffie [9] for an account of the theory and the references therein.

The stochastic maximum principle has been developed completely in recent years in Peng [10] and Yong and Zhou [5]. It is the counterpart of the maximum principle for deterministic problems. The distinctive feature is the use of the concept of forward–backward stochastic differential equations, which naturally arise, governing the evolution of the state variables and the corresponding adjoint variables. Antecedents of the maximum principle are found in Kushner [11], Bismut[1] [7] or Haussmann [12]. Other developments, applicable to problems with differential equations with random coefficients can be found in Marti [13].

It is the aim of this paper to develop a new approach to stochastic control. The novelty comes from the fact that we obtain a system of PDEs that a smooth Markov control must satisfy and that also provides sufficient condition for optimality in the spirit of the verification theorems. Although the system is obtained using classical methods—the maximum principle applied to every initial condition—the authors have not found any reference in the literature to the possibility of establishing a system of PDEs to characterize the optimal control directly. The equations of this new system are of a different type than the HJB. In the case considered in this paper, where the control does not affect the diffusion coefficient in the state equation, both the HJB and the equations of the new system are semilinear. There is an important difference, however, because the nonlinearities in the first-order derivatives in the former equation can be very general, whereas in the latter they are always of quadratic type. This fact can be used to establish the existence and uniqueness of smooth optimal Markov controls as shown in Josa-Fombellida and Rincón-Zapatero [14].

Our approach has the following limitations:

  (i) we consider only problems where the diffusion coefficient is independent of the control variables;
 (ii) the optimal control is interior to the control region;
(iii) controls are Markovian and of class $\mathcal{C}^{1,2}$;
(iv) the number of control variables is greater than or equal to the number of state variables.

---

[1] The maximum principle, duality methods and the concept of forward–backward stochastic differential equations have its roots in the work of Bismut, who gave a very complete and rigorous theory regarding these topics

It is worth noting that many control problems share these properties, specially some important models arising in economics.

The idea to systematically obtain a system of PDEs for the optimal control dates back to the paper by Bourdache-Siguerdidjane and Fliess [15] for deterministic control problems. The method was later extended to differential games in Rincón-Zapatero et al. [16] and Rincón-Zapatero [17].

The paper is organized as follows. In Sect. 2, we present the control problem and the first hypotheses and notations. In Sect. 3, we find a system of partial differential equations that a vector of optimal controls of class $\mathcal{C}^{1,2}$ must satisfy. Section 4 is devoted to establishing sufficient conditions to guarantee that a vector of admissible controls satisfying the system is an optimal control of the problem. Hence, Sects. 3 and 4 respectively, provide necessary and sufficient conditions for optimality. Concluding remarks are stated in Sect. 5.

2 Control Problem

In this section the framework for the stochastic control problem to be considered is presented. First, we introduce some useful notation. The partial derivatives are indicated by subscripts and ∂_x stands for *total derivation*; the partial derivative of a scalar function with respect to a vector is a column vector; given a real vector function $g : \mathbb{R}^n \to \mathbb{R}^m$ and a vector $z \in \mathbb{R}^n$, g_z is defined as the matrix $(\partial g^i / \partial z^j)_{i,j}$; for a matrix A, $A^{(i)}$ denotes the ith column and A^{ij} denotes the (i, j)th element; the vector $v \in \mathbb{R}^n$ is a column vector and v^i is the ith component; finally, $^\mathrm{T}$ denotes transposition.

Let a time interval $[0, T]$, with $0 < T \leq \infty$, and let $(\Omega, \mathcal{F}, \mathrm{P})$ be a complete probability space. Assume that, on this space, a d-dimensional Brownian motion $\{w(t), \mathcal{F}_t\}_{t \in [0,T]}$ is defined with $\{\mathcal{F}_t\}_{t \in [0,T]}$ being the Brownian filtration. Let E denote expectation under the probability measure P.

The state space is $\mathbb{R}^n$ and the control region is some subset $U \subseteq \mathbb{R}^m$, with $m \geq n$. This assumption will be explained later in Remark 3.2. A U-valued *control process* $\{(u(s), \mathcal{F}_s)\}$ defined on $[t, T] \times \Omega$ is an $\mathcal{F}_s$-progressively measurable map $(r, \omega) \to u(r, \omega)$ from $[t, s] \times \Omega$ into U, that is, $u(t, \omega)$ is $\mathcal{B}_s \times \mathcal{F}_s$-measurable for each $s \in [t, T]$, where $\mathcal{B}_s$ denotes the Borel σ-field in $[t, s]$. For simplicity, we use $u(t)$ for $u(t, \omega)$.

The state process $\xi \in \mathbb{R}^n$ obeys a system of controlled stochastic differential equations of the form

$$d\xi(s) = f(s, \xi(s), u(s)) \, ds + \sigma(s, \xi(s)) \, dw(s), \quad s \geq t, \tag{1}$$

with initial condition $\xi(t) = x$. ξ^u is sometimes used to indicate the dependence of the state variable with respect to the control u. An important feature of the above system is that the noise coefficient σ is independent of the control variable u. Here, $\sigma = (\sigma_{ij})$ is an $n \times d$ matrix.

Definition 2.1 (Admissible Control) A control $\{(u(t), \mathcal{F}_t)\}_{t \in [0,T]}$ is called admissible if:

(i) for every (t, x), the system of SDE (1) with initial condition $\xi(t) = x$ admits a pathwise unique strong solution;

(ii) there exists some function $\phi : [0, T] \times \mathbb{R}^n \to U$ of class $\mathcal{C}^{1,2}$ such that u is in relative feedback to ϕ, i.e. $u(s) = \phi(s, \xi(s))$ for every $s \in [0, T]$.

Let $\mathcal{U}(t, x)$ denote the set of admissible controls corresponding to the initial condition $(t, x) \in [0, T] \times \mathbb{R}^n$.

According to the definition, we are considering Markovian controls. If ϕ is time independent, the corresponding control will be called a stationary Markov control. u and ϕ will sometimes be identified in the notation.

Given the initial data $(t, x) \in [0, T] \times \mathbb{R}^n$, the criterion to be maximized is

$$J(t, x; u) = \mathrm{E}_{tx} \left\{ \int_t^T L(s, \xi(s), u(s)) \, ds + S(T, \xi(T)) \right\}, \qquad (2)$$

where E_{tx} denotes conditional expectation with respect to the initial condition (t, x). In the following, the subscript will be eliminated if there is no confusion. The functions $f : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}^n$, $\sigma : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n \times d}$, $L : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}$, $S : [0, T] \times \mathbb{R}^n \to \mathbb{R}$, are all assumed to be of class $\mathcal{C}^2$ with respect to (x, u) and of class $\mathcal{C}^1$ with respect to t. The hypotheses established so far will be assumed to hold throughout the paper. Given that our aim is to solve the problem *for every* $(t, x) \in [0, T] \times \mathbb{R}^n$, $\mathcal{U}$ will often be written instead of $\mathcal{U}(t, x)$.

In the specification of the problem, we have supposed $m \geq n$, that is, the dimension of the control variable is *greater than or equal to* the dimension of the state variable. This is a crucial assumption for the following developments. However, for the sake of simplicity, the case $m = n$ will be considered first and then we will show in Remark 3.2 that the case $m > n$ can be reduced to the equality situation.

With a view to applying the stochastic maximum principle as it is stated in Yong and Zhou [5], an additional assumption will be imposed.

Assumption A1 There exists a constant $C > 0$ and a modulus of continuity $\bar{\omega} : [0, \infty) \to [0, \infty)$ such that, for $\psi = f, \sigma, L, S$, we have

$$|\psi(t, x, u) - \psi(t, \widehat{x}, \widehat{u})| \leq C \|x - \widehat{x}\| + \bar{\omega}(\|u - \widehat{u}\|),$$
$$|\psi_x(t, x, u) - \psi_x(t, \widehat{x}, \widehat{u})| \leq C \|x - \widehat{x}\| + \bar{\omega}(\|u - \widehat{u}\|),$$
$$|\psi_{xx}(t, x, u) - \psi_{xx}(t, \widehat{x}, \widehat{u})| \leq \bar{\omega}(\|x - \widehat{x}\| + \|u - \widehat{u}\|),$$
$$\forall t \in [0, T], \; x, \widehat{x} \in \mathbb{R}^n, \; u, \widehat{u} \in U,$$
$$|\psi(t, 0, u)| \leq C, \quad \forall (t, u) \in [0, T] \times U.$$

Consider a control satisfying property (ii) of Definition 2.1. Then, the Lipschitz and linear growth conditions on f and σ postulated in Assumption A1 imply that the control also satisfies (i), that is, it is admissible; see Yong and Zhou [5], p. 114. However, the assumptions are quite stringent and will only be used in the derivation of the quasilinear system as a necessary condition for optimality. Sufficiency conditions, which will be established in Sect. 4, do not make use of Assumption A1.

The backward evolution operator associated with (1) is given by

$$\mathcal{A}^\phi W(t, x) = W_t(t, x) + W_x^{\mathrm{T}}(t, x) f(t, x, \phi(t, x)) + (1/2) \operatorname{Tr}\{(\sigma \sigma^{\mathrm{T}} W_{xx})(t, x)\},$$

with $W : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n$ of class $\mathcal{C}^{1,2}$ and where

$$\operatorname{Tr}\{\sigma \sigma^{\mathrm{T}} W_{xx}\} := \big(\operatorname{Tr}\{\sigma \sigma^{\mathrm{T}} W_{xx}^1\}, \ldots, \operatorname{Tr}\{\sigma \sigma^{\mathrm{T}} W_{xx}^n\}\big)^{\mathrm{T}}.$$

The *value function* is defined as $V(t, x) = \sup_{u \in \mathcal{U}(t,x)} J(t, x; u)$. An admissible control $\widehat{u} \in \mathcal{U}$ is *optimal* if $V(t, x) = J(t, x; \widehat{u})$ for every initial condition (t, x).

The standard approach adopted in the literature to determine an optimal control is to solve the HJB equation

$$V_t(s, x) + \max_{u \in U}\big\{L(s, x, u) + V_x(s, x)^{\mathrm{T}} f(s, x, u)$$

$$+ (1/2) \operatorname{Tr}\big\{(\sigma \sigma^{\mathrm{T}} V_{xx})(s, x)\big\}\big\} = 0, \tag{3}$$

$$V(T, x) = S(T, x), \quad t \le s \le T, \ x \in \mathbb{R}^n. \tag{4}$$

3 Necessary Conditions

Our purpose in this section is to find a system of PDEs that an optimal control must satisfy. Let $L_{\mathcal{F}}^2([0, T]; \mathbb{R}^n)$ be the set of all processes $X(\cdot)$ with values in $\mathbb{R}^n$ adapted to filtration $\{\mathcal{F}_t\}_{t \ge 0}$ such that

$$\mathrm{E} \int_0^T \|X(t)\|^2 \, dt < \infty.$$

As previously stated, Assumption A1 allows us to apply the stochastic maximum principle so that, if given the initial condition (t, x), the pair (ξ, u) is optimal, with $u \in \mathcal{U}(t, x)$, then there exist processes $p \in L_{\mathcal{F}}^2([0, T]; \mathbb{R}^n)$, $q \in (L_{\mathcal{F}}^2([0, T]; \mathbb{R}^n))^d$ satisfying for $s \in [t, T]$ the first-order adjoint equations

$$dp(s) = -\left(H_x(s, \xi(s), \phi(s, \xi(s)), p(s)) + \sum_{i=1}^d \sigma_x^{(i)}(s, \xi(s))^{\mathrm{T}} q^{(i)}(s)\right) ds$$

$$+ q(s) \, dw(s), \tag{5}$$

$$p(T) = S_x(T, \xi(T)), \tag{6}$$

where $H(t, x, u, p) = L(t, x, u) + p^{\mathrm{T}} f(t, x, u)$ is the deterministic Hamiltonian function, corresponding to the associated deterministic problem, with $\sigma \equiv 0$. A more precise notation for the adjoint processes is $p(s; t, x)$ and $q(s; t, x)$ with $s \in [t, T]$, though in the following, we suppress the dependence with respect to the initial condition (t, x).

Furthermore, the following maximization condition:

$$H(s, \xi(s), \phi(s, \xi(s)), p(s)) = \max_{u \in U} H(s, \xi(s), u, p(s)) \tag{7}$$

holds for every $s \in [t, T]$, P-a.s.

For the next result, which establishes a necessary condition of optimality in terms of a new system of PDEs, we define

$$\Gamma(t, x, u) := -f_u^{-\mathrm{T}} L_u(t, x, u), \tag{8}$$

and $\Sigma^{\mathrm{T}} := \mathrm{diag}(\sigma^{\mathrm{T}} \ \sigma^{\mathrm{T}} \partial_x \Gamma \ (\phi_x \sigma)^{\mathrm{T}})$. $\widehat{H}_{\{\cdot\}}$ will denote $H_{\{\cdot\}}$ once (8) is substituted into it.

Theorem 3.1 (Necessary Conditions) *Let Assumption A1 on the coefficient functions be satisfied. Let $\phi \in \mathcal{U}$ be an interior optimal Markov control such that $\det f_u(t, x, \phi) \neq 0$ for all $(t, x) \in [0, T] \times \mathbb{R}^n$. Then, ϕ satisfies*

$$0 = \widehat{H}_{ut} + \widehat{H}_{ux}^{\mathrm{T}} f + \widehat{H}_{up}^{\mathrm{T}} \left(-\widehat{H}_x - \sum_{i=1}^{m} \sigma_x^{(i)} (\partial_x \Gamma \sigma)^{(i)} \right) + \widehat{H}_{uu} \mathcal{A}^\phi \phi + \frac{1}{2} \mathrm{Tr}\{\Sigma \Sigma^{\mathrm{T}} \nabla^2 \widehat{H}_u\} \tag{9}$$

and the final condition

$$L_{u^i}(T, x, \phi(T, x)) + S_x(T, x)^{\mathrm{T}} f_{u^i}(T, x, \phi(T, x)) = 0, \quad i = 1, \ldots, n. \tag{10}$$

Proof Since by assumption the maximizing argument is interior to U, (7) implies

$$H_{u^i}(s, \xi(s), \phi(s, \xi(s)), p(s)) = 0, \quad \forall s \in [t, T], \ P\text{-a.s.}, \tag{11}$$

for all $i = 1, \ldots, n$. Assuming that f_u is invertible for all $(t, x, u) \in [0, T] \times \mathbb{R}^n \times U$, it is possible to obtain the unique solution of the above *linear system* in the adjoint variable p,

$$L_u + f_u^{\mathrm{T}} p = 0,$$

as

$$p = -f_u^{-\mathrm{T}} L_u. \tag{12}$$

An obvious consequence of (11) is $dH_{u^i}(s, \xi(s), u(s), p(s)) = 0$ a.s. For an admissible feedback ϕ, Itô's rule is applicable to $u(s) = \phi(s, \xi(s))$; hence, omitting the arguments and in differential notation,

$$du^i = d\phi^i = \phi_t^i dt + \phi_x^i d\xi + (1/2) d\xi^{\mathrm{T}} \phi_{xx}^i d\xi, \quad i = 1, \ldots, n. \tag{13}$$

Applying again Itô's rule to H_{u^i} for $i = 1, \ldots, n$, we have

$$0 = dH_{u^i} = H_{u^i t} \, dt + \nabla H_{u^i} \begin{pmatrix} d\xi \\ dp \\ d\phi \end{pmatrix} + \frac{1}{2} (d\xi^{\mathrm{T}} \ dp^{\mathrm{T}} \ d\phi^{\mathrm{T}}) \nabla^2 H_{u^i} \begin{pmatrix} d\xi \\ dp \\ d\phi \end{pmatrix}, \tag{14}$$

where ∇ and ∇^2 denote the gradient and the Hessian matrix operators respectively, with respect to the variables (x, p, u). Substituting (1, 5, 13) in the equality (14) and taking into account that $H_{u^i pp} = 0$ for all i, because the Hamiltonian is linear in p, the following system of stochastic differential equations holds along $(s, \xi(s), u(s), p(s))$ a.s.:

6

$$0 = \left(H_{u^i t} + H_{u^i x}^{\mathrm{T}} f + H_{u^i p}^{\mathrm{T}} \left(-H_x - \sum_{i=1}^{m} \sigma_x^{(i)} q^{(i)} \right) + H_{u^i u}^{\mathrm{T}} \mathcal{A}^{\phi} \phi \right.$$
$$\left. + \frac{1}{2} \operatorname{Tr} \left\{ \Sigma \Sigma^{\mathrm{T}} \nabla^2 \widehat{H}_u \right\} \right) ds + \left(H_{u^i x}^{\mathrm{T}} \sigma + H_{u^i p}^{\mathrm{T}} q + H_{u^i u}^{\mathrm{T}} \phi_x \sigma \right) dw(s). \quad (15)$$

Therefore, both the drift term and the diffusion coefficient of this system of SDEs must be identically null a.s. In order to obtain a system of PDEs for the optimal control, we must eliminate the adjoint vector p by means of (12). Equating the diffusion coefficient to zero, we get

$$q = -f_u^{-\mathrm{T}} (\widehat{H}_{ux} \sigma + \widehat{H}_{uu} \phi_x \sigma) \text{ a.s.}, \quad s \ge t.$$

In fact, q can be expressed as

$$q = (\partial_x \Gamma) \sigma. \quad (16)$$

The possibility to write q as shown in (16) follows from the identities

$$H_{u^i p} = f_{u^i}, \qquad \Gamma_x = -f_u^{-\mathrm{T}} \widehat{H}_{ux}, \qquad \Gamma_u = -f_u^{-\mathrm{T}} \widehat{H}_{uu}.$$

The drift term in (15) also vanishes a.e.; hence, after substitution of (12) and (16), the system of PDEs (9) characterizing an admissible optimal control is obtained. Note that (9) and (16) are valid a.s. along the optimal trajectory, but that at $(t, \xi(t)) = (t, x)$, (9) holds with certainty.

The stochastic maximum principle also provides a boundary condition at time T for the system of PDEs, which is given implicitly by (10). This follows from (6) and (12) evaluated at $t = T$; we will suppose that it is possible to obtain $\phi(T, x) := \varphi(x)$ for a function φ sufficiently regular. This is enough to check if the hypotheses of the implicit function theorem are fulfilled. □

Some comments about the structure of the system and comparison with the HJB equation (3) are pertinent here. The system is semilinear because the terms involving the second-order derivatives of ϕ are independent of the solution. Furthermore, assuming the invertibility of $\widehat{H}_{uu}$, the system is weakly coupled; that is, the second-order derivatives of ϕ^i appear only in equation i. The first-order derivatives are coupled and appear in a nonlinear way derived from the quadratic-type terms

$$\operatorname{Tr} \left\{ \phi_x \sigma (\phi_x \sigma)^{\mathrm{T}} \widehat{H}_{uuu} \right\} \quad \text{and} \quad \operatorname{Tr} \left\{ \phi_x \sigma (\phi_x \sigma)^{\mathrm{T}} \widehat{H}_{upu} (-f_u^{-\mathrm{T}} \widehat{H}_{uu}) \right\}.$$

This is a very interesting feature that has been used by the authors in Josa-Fombellida and Rincón-Zapatero [14] to study an economic model of determining the optimal consumption subject to stochastic returns. With respect to the HJB equation, it is also of semilinear type, but the nonlinearity with respect to ϕ_x can be much more general and not only of quadratic type. On the other hand, it must be pointed out that the HJB equation is a *single* equation, whereas we have obtained a *system* of n PDEs, but with a simpler structure.

It would be possible to replace the smoothness assumption on ϕ for a weaker one. Given that all that is needed is to apply Itô's rule, Theorem 3.1 is true if the class of

Markov controls is $(W_{l,\,\mathrm{loc}}^{1,2}([0,T]\times\mathbb{R}^n))^n$, $l \geq 2$, the space of functions such that the weak partial derivatives of order 1 with respect to time and order 2 with respect to x are in $(L_{\mathrm{loc}}^l([0,T]\times\mathbb{R}^n))^n$; see Krylov [2]. Note that the hypotheses imposed imply that H_u belongs to $W_{l,\,\mathrm{loc}}^{1,2}([0,T]\times\mathbb{R}^n)$.

Remark 3.1 There is in the literature a different but closely related system of PDEs which characterize the vector of adjoint variables under some regularity assumptions, see (17) below. This system was obtained for the first time in Bismut [7] and later in[2] Elliot [18]. It is important to note that the system below depends also on the optimal control and for this reason it appears with a simple structure. To obtain the system for the adjoint variables, we can proceed as follows. Let us suppose the existence of a vector function γ of class $\mathcal{C}^{1,2}$ depending of the variables (s,y) and such that $p(s) = \gamma(s,\xi(s))$, where p is the adjoint variable of the problem with initial condition (t,x). Applying Itô's rule to $\gamma(s,\xi(s))$, we have

$$d\gamma = \left(\gamma_t + \gamma_x f + \frac{1}{2}\operatorname{Tr}\{\sigma\sigma^{\mathrm{T}}\gamma_{xx}\}\right)ds + \gamma_x\sigma\,dw. \tag{17}$$

Once the validity of the maximum principle is established, by the uniqueness of the solutions of (5), we can match the diffusion terms and drift terms in the expressions (5) and (17) to obtain $q = \gamma_x\sigma$ and

$$-\left(H_x + \sum_{i=1}^n (\sigma_x^{(i)})^{\mathrm{T}} q^{(i)}\right) = \gamma_t + \gamma_x f + \frac{1}{2}\operatorname{Tr}\{\sigma\sigma^{\mathrm{T}}\gamma_{xx}\}.$$

Of course, the first equality is nothing but (16). Therefore, we find that the second identity is transformed into

$$\gamma_t + \gamma_x f + H_x + \sum_{i=1}^m (\sigma_x^{(i)})^{\mathrm{T}} \gamma_x \sigma^{(i)} + \frac{1}{2}\operatorname{Tr}\{\sigma\sigma^{\mathrm{T}}\gamma_{xx}\} = 0. \tag{18}$$

For the derivation of this identity, the *equal dimension* condition between the state and control variables is not needed. Furthermore, the equality $q = \gamma_x\sigma$ allows situations to be handled where the diffusion parameter σ depends also on the control variables u. In this case, the elimination of the optimal control variables is not so straightforward. As already observed, the system (18) depends also on the unknown vector of optimal controls. Supposing that it is possible to obtain a sufficiently regular function $\widetilde{u}(t,x,z)$ such that $\widehat{\phi}(t,x) = \widetilde{u}(t,x,\gamma(t,x))$, that is, $\widetilde{u}$ is the inverse function of Γ with respect to its third component for all (t,x), then by substituting in the previous system of equations, we obtain a system of PDEs that truly characterize the vector of adjoint variables. However, writing the system for the optimal control does not require the inverse function $\widetilde{u}$ to be found, which can be hard or impossible to do, even in scalar problems. Under the conditions contemplated in this paper,

---

[2]In Elliot [18], a misprint is registered making the equation shown different from that appearing in Bismut [7] and in (18) of the present paper

it is only necessary to solve the *linear* system (12) to obtain $\gamma(t, x) = \Gamma(t, x, \phi(t, x))$ and by substituting in (18), to arrive at the desired PDE system for ϕ. Of course, this is simply the system (9).

Remark 3.2 Case $m > n$. When the number of control variables is greater than the number of state variables, $m > n$, the linear system in p is overdeterminate. Because the maximum principle holds, the existence of a solution to the overdetermined system is assured. This solution can be obtained as follows. Suppose that f_u has rank n for all t, x, u; then $f_u^{\mathrm{T}} f_u$ has full rank n; hence, from (11),

$$p = -(f_u f_u^{\mathrm{T}})^{-1} f_u L_u.$$

Now, the argument runs as shown above, obtaining a system of n PDEs for $m > n$ unknowns. These equations can be complemented with an algebraic relationship between the controls, which is obtained from the fact that the system $H_u = 0$ admits a solution in p. In this way, $m - n$ control variables can be formally expressed by means of n of them.

In the case $m < n$, elimination of p is not so straightforward. Now, the procedure to obtain a system of PDEs for the control would be to take $n - m + 1$ Itô differentials in the identity $H_u \equiv 0$. This leads to PDE equations for the optimal control of higher order and of a very different nature than (9); hence, this case will not be pursued in this paper.

4 Sufficient Conditions

The main objective in this section is to show that a solution of class $\mathcal{C}^{1,2}$ of (9, 10), maximizing the Hamiltonian function for all (t, x) and satisfying some additional assumptions, is an optimal Markov control for problem (1, 2). This result is, therefore, similar to the *verification theorems* in Fleming and Rishel [1] or Fleming and Soner [3]. The applicability of Theorem 4.1 established in this section has been shown in an economic model studied in Josa-Fombellida and Rincón-Zapatero [14].

The process ξ depends on the initial condition (t, x). In the following, $\xi_{x^i}^j$ will denote the partial derivative of ξ^j with respect to x^i.

We consider the following assumption.

Assumption A2

$$\mathrm{E}\left\{\int_t^T (\gamma^j \sigma_{x^l}^{(j)} + q^{(j)}) \xi_{x^i}^j \, dw(s)\right\} = 0, \quad \text{for every } i, j, l = 1, \ldots, n.$$

The following result establishes that the adjoint process $p(s) = \gamma(s, \xi(s))$ is the gradient with respect to x of the objective functional. This result, of independent interest, is a previous step in the formulation of the sufficiency theorem that will be stated later.

Recall from the previous section that Γ denotes $f_u^{-\mathrm{T}} L_u$ and $\gamma(t, x) = \Gamma(t, x, \widehat{\phi}(t, x))$, where $\widehat{\phi}$ is an admissible Markov control solving the semilinear system. The proof of the following result can be found in Josa-Fombellida and Rincón-Zapatero [14].

Proposition 4.1 (Shadow Price) *Let $\widehat{\phi} \in \mathcal{U}$ be a solution of (9, 10) such that Assumption A2 is satisfied and $f_u(t, x, \widehat{\phi})$ is invertible for all $(t, x) \in [0, T] \times \mathbb{R}^n$. Then,*

$$J_x(t, x; \widehat{\phi}) = \Gamma(t, x, \widehat{\phi}(t, x)) = p(t),$$
$$J_{xx}(t, x; \widehat{\phi})\sigma(t, x) = q(t),$$

for every $(t, x) \in [0, T] \times \mathbb{R}^n$.

Once we have identified the vector of adjoint variables with the gradient of the objective functional, the system (9) can be expressed in *conservative form*. Since γ is the gradient with respect to the variable x of the function $J(t, x; \widehat{\phi})$, which is of class $\mathcal{C}^3$, $\gamma_{x^j}^i = \gamma_{x^i}^j$ is satisfied for every $i, j = 1, \dots, n$, because the crossed second-order partial derivatives of the function J coincide. By the same argument, $\gamma_{x^r x^j}^i = \gamma_{x^j x^i}^r$ for all $i, j, r = 1, \dots, n$. On the other hand, after some tedious calculations, we find that

$$(\mathrm{Tr}\{\sigma\sigma^{\mathrm{T}}\gamma_x\})_{x^r} = \mathrm{Tr}\{\sigma\sigma^{\mathrm{T}}\gamma_{xx}^r\} + 2\sum_{i=1}^{m}(\sigma_{x^r}^{(i)})^{\mathrm{T}}\left(\sum_{j=1}^{n}\gamma_x^j\sigma^{ji}\right),$$

and substituting this expression in (18), we obtain

$$\gamma_t + \partial_x\left(L + \gamma^{\mathrm{T}}f + \frac{1}{2}\mathrm{Tr}\{\sigma\sigma^{\mathrm{T}}\gamma_x\}\right) = 0, \tag{19}$$

where the fact that $H_u = 0$ holds at the optimal control has been used. It is interesting to compare the structure of (19), which is expressed in conservative form, with that of (18), which appears in nonconservative form.

In terms of $\Gamma(t, x, \phi)$, (19) can be rewritten as

$$\partial_t\Gamma(t, x, \phi(t, x))$$
$$+ \partial_x\left(\mathcal{H}(t, x, \phi(t, x)) + \frac{1}{2}\mathrm{Tr}\{\sigma(t, x)\sigma(t, x)^{\mathrm{T}}\partial_x\Gamma(t, x, \phi(t, x))\}\right) = 0, \tag{20}$$

with

$$\mathcal{H}(t, x, u) := H(t, x, u, \Gamma(t, x, u)).$$

Taking total derivatives, a system of partial differential equations of second order arise, which is the same as (9). Expressing the system in conservative form is useful, because it allows us in the next theorem to establish a sufficient result of optimality. It also makes possible to obtain the value function from the control, as will be shown in the following section.

Now, we are in position to establish the following sufficient condition for optimality.

Theorem 4.1 (Verification Theorem) *Let $\widehat{\phi} \in \mathcal{U}$ be a solution of (9, 10) such that Assumption A2 is satisfied and $f_u(t, x, \widehat{\phi})$ is invertible for all $(t, x) \in [0, T] \times \mathbb{R}^n$.*

Suppose further that the following maximization property holds for all $(t, x) \in [0, T] \times \mathbb{R}^n$, for every admissible Markov control u:

$$H(t, x, \widehat{\phi}, \Gamma(t, x, \widehat{\phi})) \geq H(t, x, u, \Gamma(t, x, \widehat{\phi})). \tag{21}$$

Then, $\widehat{\phi}$ is an optimal Markov control for the problem (1, 2).

Proof Let u be any admissible Markov control and let ξ^u be the associated process with initial condition (t, x). We will omit the dependence of ξ^u on the initial condition in order to facilitate the exposition. Let $u(s)$ be $u(s, \xi^u(s))$. Applying Itô's rule to $J(s, \xi^u(s); u)$, $s \geq t$, we have

$$dJ(s, \xi^u(s); u) = \mathcal{A}^u J(s, \xi^u(s); u)\, ds + J_x(s, \xi^u(s); u)^{\mathrm{T}} \sigma(s, \xi^u(s))\, dw(s). \tag{22}$$

On the other hand, as shown in Yong and Zhou [5], we can write the objective functional as

$$J(s, \xi^u(s); u) = \mathrm{E}\left\{ \int_s^T L(r, \xi^u(r), u(r))\, dr + S(T, \xi^u(T)) \,|\, \mathcal{F}_s^t \right\},$$
$$\forall s \in [t, T], \ \ P\text{-a.s.}, \tag{23}$$

where $\{\mathcal{F}_s^t\}_{s \geq t}$ is the filtration of the σ-fields generated by Brownian motion in the interval $[t, s]$. The process

$$m(s) = \mathrm{E}\left\{ \int_t^T L(r, \xi^u(r), u(r))\, dr + S(T, \xi^u(T)) \,|\, \mathcal{F}_s^t \right\}, \quad s \in [s, T],$$

is a square-integrable $\{\mathcal{F}_s^t\}_{s \in [t, T]}$-martingale; hence, by the martingale representation theorem, we have

$$m(s) = m(t) + \int_t^s M(r)\, dw(r),$$

with $M \in (L_{\mathcal{F}}^2(t, T; \mathbb{R}^n))^d$. Let us observe that $m(t) = J(t, x; u)$. Therefore,

$$m(s) = J(t, x; u) + \int_t^s M(r)\, dw(r). \tag{24}$$

By (23) and (24), we have

$$J(s, \xi^u(s); u) = m(s) - \mathrm{E}\left\{ \int_t^s L(r, \xi^u(r), u(r))\, dr \right\}$$
$$= J(t, x; u) - \mathrm{E}\left\{ \int_t^s L(r, \xi^u(r), u(r))\, dr \right\} + \int_t^s M(r)\, dw(r).$$

It then follows that

$$dJ(s, \xi^u(s); u) = -\mathrm{E}\{L(s, \xi^u(s), u(s))\}\, ds + M(s)\, dw(s). \tag{25}$$

We get, from (22) and (25),

$$\mathrm{E}\left\{ J_s(s, \xi^u(s); u) + L(s, \xi^u(s), u(s)) + J_y^{\mathrm{T}}(s, \xi^u(s); u) f(s, \xi^u(s), u(s)) \right.$$

$$\left. + \frac{1}{2} \mathrm{Tr}\left\{ (\sigma\sigma^{\mathrm{T}})(s, \xi^u(s)) J_{yy}(s, \xi^u(s); u) \right\} \right\} = 0.$$

This equality holds for all admissible $u \in \mathcal{U}$, for all $s \in [t, T]$. In particular, it holds for $\widehat{\phi}$; hence,

$$0 = \mathrm{E}\left\{ J_s(s, \xi^u(s); \widehat{\phi}) + H(s, \xi^u(s), \widehat{\phi}, J_y(s, \xi^u(s); \widehat{\phi})) \right.$$

$$\left. + \frac{1}{2} \mathrm{Tr}\left\{ (\sigma\sigma^{\mathrm{T}})(s, \xi^u(s)) J_{yy}(s, \xi^u(s); \widehat{\phi}) \right\} \right\}$$

$$\geq \mathrm{E}\left\{ J_s(s, \xi^u(s); \widehat{\phi}) + H(s, \xi^u(s), u, J_y(s, \xi^u(s); \widehat{\phi})) \right.$$

$$\left. + \frac{1}{2} \mathrm{Tr}\left\{ (\sigma\sigma^{\mathrm{T}})(s, \xi^u(s)) J_{yy}(t, \xi^u(s); \widehat{\phi}) \right\} \right\},$$

because $J_y \equiv \Gamma$, an identity which is proven in Proposition 4.1, and because of (21). Expanding the Hamiltonian function we find that the latter inequality is equivalent to

$$\mathrm{E}\{L(s, \xi^u(s), u(s)) + \mathcal{A}^u J(s, \xi^u(s); \widehat{u})\} \leq 0.$$

Integrating and exchanging expectation and integration, we have

$$\mathrm{E}\left\{ \int_t^T (L(s, \xi^u(s), u(s)) + \mathcal{A}^u J(s, \xi^u(s); \widehat{\phi})) \, ds \right\} \leq 0. \tag{26}$$

Given that, by the assumptions made,

$$\int_t^T J_x(s, \xi^u(s); \widehat{\phi})^{\mathrm{T}} \sigma(s, \xi^u(s)) \, dw(s)$$

is a martingale, the application of the Dynkin formula to (25) leads to

$$\mathrm{E}\{S(T, \xi^u(T))\} - J(t, x; \widehat{\phi}) = \mathrm{E}\left\{ \int_t^T \mathcal{A}^u J(s, \xi^u(s); \widehat{\phi}) \, ds \right\}. \tag{27}$$

Substituting (27) into (26), we obtain

$$\mathrm{E}\left\{ \int_t^T L(s, \xi^u(s), u(s)) \, ds \right\} \leq J(t, x; \widehat{\phi}) - \mathrm{E}\{S(T, \xi^u(T))\}, \tag{28}$$

that is, $J(t, x; u) \leq J(t, x; \widehat{\phi})$. $\qquad\qquad\square$

Remark 4.1 Condition (21) holds automatically when $\widehat{\phi}$ is interior to the control set U and the Hamiltonian function is concave with respect to u for every t, x, p. To see this, note that

$$H_u(t, x, \widehat{\phi}, \Gamma(t, x, \widehat{\phi})) = 0$$

is trivially fulfilled by the definition of Γ; hence, $\widehat{\phi}$ is a critical point of the concave function $u \mapsto H(\cdot, \cdot, u, \cdot)$, so $\widehat{\phi}$ is a global maximum of H. On the other hand, it is worth noting that the full strength of (21) is not really needed in the proof. It only suffices that, for every initial condition (t, x) and for every admissible Markov control u, the following holds:

$$\mathrm{E}\left\{H(s, \xi^u(s), \Gamma(s, \xi^u(s), \widehat{\phi}), \widehat{\phi})\right\} \geq \mathrm{E}\left\{H(s, \xi^u(s), \Gamma(s, \xi^u(s), \widehat{\phi}), u)\right\}, \quad (29)$$

where ξ^u is the state variable process associated to u.

Remark 4.2 (Infinite Horizon) Proposition 4.1 can be extended to the infinite-horizon case, $T = \infty$, when the following transversality condition holds:

$$\lim_{T \to \infty} \mathrm{E}\left\{\gamma^{\mathrm{T}}(T, \xi(T)) \, \xi_x(T)\right\} = 0. \quad (30)$$

By Proposition 4.1, (30) is the same as

$$\lim_{T \to \infty} \mathrm{E}\{\partial_x V(T, \xi(T))\} = 0,$$

for every initial condition (t, x), that is, the long run behavior of the expected value function along the optimal trajectory is independent of the initial condition x, for every x. With respect to Theorem 4.1, two assumptions about the limit of $J(t, x; \widehat{\phi})$ as $t \to \infty$ must be added to the hypotheses, in order to assure the optimality of $\widehat{\phi}$. One of them is (30), which assures the equality between p^i and $J_{x^i}(t, x; \widehat{\phi})$, for $i = 1, \ldots, n$. The other one is obtained by substituting $\mathrm{E}\{S(T, \xi^u(T))\}$ by $J(T, \xi^u(T); \widehat{\phi})$ in (28), given that, in the infinite-horizon problem, there is no residual function S. Taking limits when T tends to infinity in the expression (28), if the conditions

$$\limsup_{T \to \infty} J(T, \xi^u(T); \widehat{\phi}) = \limsup_{T \to \infty} V(T, \xi^u(T)) \geq 0 \quad (31)$$

and

$$\lim_{T \to \infty} \mathrm{E}\left\{\int_t^T L(s, \xi^u(s), u(s)) \, ds\right\} = \mathrm{E}\left\{\int_t^\infty L(s, \xi^u(s), u(s)) \, ds\right\} < \infty$$

hold for all admissible control u, then

$$J(t, x; u) \leq J(t, x; \widehat{\phi}).$$

The latter equality simply means that the cost functional of the infinite-horizon problem makes sense for the class of admissible controls.

Remark 4.3 In the deterministic case, $\sigma \equiv 0$, the system of partial differential equations (9) is of course of first order and quasilinear. The system for this case was first derived in Bourdache-Siguerdidjane and Fliess [15]. Clearly, the results remain valid now for $\mathcal{C}^1$ solutions and Assumption A2 is not needed. In Rincón-Zapatero et al. [16] and Rincón-Zapatero [17], an extension to differential games is provided.

5 Conclusions

This paper provides an alternative method for the analysis of stochastic optimal control problems to the classical ones based on dynamic programming, duality, and the maximum principle. The novelty of the approach that we propose in this paper does not consists in the tools that we use in the construction of the theoretical framework—which heavily depend on dynamic programming concepts and the maximum principle—but in the optimality conditions, necessary and sufficient, that are obtained. These are entirely new. We do not pretend to convey to the reader the idea that our approach is superior to the existing ones—we have remarked the limitations of the method in the Introduction—but to provide a different perspective, based in a system of PDEs which characterize directly the optimal controls, without resorting to the value function. A useful feature of the system of PDEs introduced in the paper is that the gradient of the optimal control enters in a quadratic way. In contradistinction, the gradient of the value function in the HJB equation enters in a nonlinear way.

References

1 Fleming, W H , Rishel, R W : Deterministic and Stochastic Optimal Control Springer, New York (1975)
2 Krylov, N V : Controlled Diffusion Processes Springer, New York (1980)
3 Fleming, W H , Soner, H M : Controlled Markov Processes and Viscosity Solutions Springer, New York (1993)
4 Hernández–Lerma, O : Lectures on continuous-time Markov control processes, Aportaciones Matemáticas Sociedad Matemática Mexicana, Mexico City (1994)
5 Yong, J , Zhou, X Y : Stochastic Controls Hamiltonian Systems and HJB Equations Springer, New York (1999)
6 Merton, R C : Optimum consumption and portfolio rules in a continuous time model J Econ Theory **3**, 373–413 (1971)
7 Bismut, J M : Conjugate convex functions in optimal stochastic control J Math Anal Appl **44**, 484–404 (1973)
8 Bismut, J M : An introductory approach to duality in optimal stochastic control SIAM Rev **20**, 62–78 (1978)
9 Duffie, D : Dynamic Asset Pricing Theory Princeton University Press, Princeton (1996)
10 Peng, S : A general stochastic maximum principle for optimal control problems SIAM J Control Optim **28**, 966–979 (1990)
11 Kushner, H J : Necessary conditions for continuous parameter stochastic optimization problems SIAM J Control **10**, 550–565 (1972)
12 Haussmann, U G : A Stochastic Maximum Principle for Optimal Control of Diffusions Pitman Research Notes in Mathematics, vol 151 Longman, Harlow (1986)
13 Marti, K : Approximative solutions of stochastic control problems by means of quasilinearization In: Hamza, M H (ed) Measurement and Control, pp 183–188 ACTA Press, Anaheim (1977)
14 Josa-Fombellida, R , Rincón-Zapatero, J P : New method to stochastic optimal control and applications to economics Universidad Carlos III de Madrid, Working Paper No 05-19 (2005)

15 Bourdache-Siguerdidjane, H , Fliess, M : Optimal feedback control of nonlinear systems Automatica **23**, 365–372 (1987)

16 Rincón-Zapatero, J P , Martínez, J , Martín-Herrán, G : New method to characterize subgame perfect Nash equilibria in differential games J Optim Theory Appl **96**, 377–395 (1998)

17 Rincón-Zapatero, J P : Characterization of Markovian equilibria in a class of differential games J Econ Dynam Control **28**, 1243–1266 (2004)

18 Elliot, R J : The optimal control of diffusions Appl Math Optim **22**, 229–240 (1990)