

Second Order Backward Stochastic Differential Equations and Fully Nonlinear Parabolic PDEs

Patrick Cheridito * H. Mete Soner † Nizar Touzi ‡ Nicolas Victoir §

August 2006

Abstract

For a d -dimensional diffusion of the form $dX_t = \mu(X_t)dt + \sigma(X_t)dW_t$, and continuous functions f and g , we study the existence and uniqueness of adapted processes Y , Z , Γ and A solving the second order backward stochastic differential equation (2BSDE)

$$\begin{aligned} dY_t &= f(t, X_t, Y_t, Z_t, \Gamma_t) dt + Z_t' \circ dX_t, \quad t \in [0, T], \\ dZ_t &= A_t dt + \Gamma_t dX_t, \quad t \in [0, T], \\ Y_T &= g(X_T). \end{aligned}$$

If the associated PDE

$$\begin{aligned} -v_t(t, x) + f(t, x, v(t, x), Dv(t, x), D^2v(t, x)) &= 0, \quad (t, x) \in [0, T] \times \mathbb{R}^d, \\ v(T, x) &= g(x), \end{aligned}$$

has a sufficiently regular solution, then it follows directly from Itô's formula that the processes

$$v(t, X_t), Dv(t, X_t), D^2v(t, X_t), \mathcal{L}Dv(t, X_t), \quad t \in [0, T],$$

solve the 2BSDE, where \mathcal{L} is the Dynkin operator of X without the drift term.

The main result of the paper shows that if f is Lipschitz in Y as well as decreasing in Γ and the PDE satisfies a comparison principle as in the theory of viscosity solutions, then the existence of a solution (Y, Z, Γ, A) to the 2BSDE implies that the associated PDE has a unique continuous viscosity solution v and the process Y is of the form $Y_t = v(t, X_t)$, $t \in [0, T]$. In particular, the 2BSDE has at most one solution. This provides a stochastic representation for solutions of fully nonlinear parabolic PDEs. As a consequence, the numerical treatment of such PDE's can now be approached by Monte Carlo methods.

Key Words: Second order backward stochastic differential equations, Fully nonlinear parabolic partial differential equations, Viscosity solutions, Monte Carlo methods.

AMS 2000 Subject Classification: 60H30, 35K55, 60H35, 65C05.

*Princeton University, Princeton NJ, USA

†Koç University, Istanbul, Turkey, Member of the Turkish Academy of Sciences and this work was partly supported by the Turkish Academy of Sciences.

‡Ecole Polytechnique, Paris, France, and Tanaka Business School, Imperial College London, UK

§University of Oxford, Oxford, UK

1 Introduction

Since their introduction, backward stochastic differential equations (BSDEs) have received considerable attention in the probability literature. Interesting connections to partial differential equations (PDEs) have been obtained, and the theory has found wide applications in areas like stochastic control, theoretical economics and mathematical finance.

BSDEs were introduced by Bismut [7] in the linear case and by Pardoux and Peng [39] in the general case. According to these authors, a solution to a BSDE consists of a pair of adapted processes (Y, Z) taking values in \mathbb{R}^n and $\mathbb{R}^{d \times n}$, respectively, such that

$$\begin{aligned} dY_t &= f(t, Y_t, Z_t)dt + Z_t' dW_t, \quad t \in [0, T], \\ Y_T &= \xi, \end{aligned} \tag{1.1}$$

where T is a finite time horizon, $(W_t)_{t \in [0, T]}$ a d -dimensional Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, P)$, f a progressively measurable function from $\Omega \times [0, T] \times \mathbb{R}^n \times \mathbb{R}^{d \times n}$ to \mathbb{R}^n and ξ an \mathbb{R}^n -valued, \mathcal{F}_T -measurable random variable.

The key feature of BSDEs is the random terminal condition ξ that the solution is required to satisfy. Due to the adaptedness requirement on the processes Y and Z , this condition poses certain difficulties in the stochastic setting. But they have been overcome, and now an impressive theory is available; see for instance, Bismut [7, 8], Arkin and Saksonov [2], Pardoux and Peng [39, 40, 41], Peng [43, 44, 45, 46, 47, 48], Antonelli [1], Ma et al. [36, 37, 38], Douglas et al. [22], Cvitanić et al. [17, 18, 19], Chevance [15], Pardoux and Tang [42], Delarue [20], Bouchard and Touzi [9], Delarue and Menozzi [21], or the overview paper El Karoui et al. [23].

If the randomness in the parameters f and ξ in (1.1) is coming from the state of a forward SDE, then the BSDE is referred to as a forward-backward stochastic differential equation (FBSDE) and its solution can be written as a deterministic function of time and the state process. Under suitable regularity assumptions, this function can be shown to be the solution of a parabolic PDE. FBSDEs are called uncoupled if the solution of the BSDE does not enter the dynamics of the forward SDE and coupled if it does. The corresponding parabolic PDE is semi-linear in case the FBSDE is uncoupled and quasi-linear if the FBSDE is coupled; see Peng [44, 46], Pardoux and Peng [40], Antonelli [1], Ma et al. [36, 37], Pardoux and Tang [42]. These connections between FBSDEs and PDEs have led to interesting stochastic representation results for solutions of semi-linear and quasi-linear parabolic PDEs, generalizing the Feynman–Kac representation of linear parabolic PDEs and opening the way to Monte Carlo methods for the numerical treatment of such PDEs, see for instance, Zhang [52], Bally and Pagès [3], Ma et al. [36, 37, 38], Bouchard and Touzi [9], Delarue and Menozzi [21]. However, PDEs corresponding to standard FBSDEs cannot be nonlinear in the second order derivatives because second order terms only arise linearly through Itô's formula from the quadratic variation of the underlying state process.

In this paper we introduce FBSDEs with second order dependence in the generator f . We call them second order backward stochastic differential equations (2BSDEs) and show how they are related to fully nonlinear parabolic PDEs. This extends the range of connections between stochastic equations and PDEs. In particular, it opens the way for the development

of Monte Carlo methods for the numerical solution of fully nonlinear parabolic PDEs. Our approach is motivated by results in Cheridito et al. [13, 14], which show how second order trading constraints lead to fully nonlinear parabolic Hamilton–Jacobi–Bellman equations for the super-replication cost of European contingent claims.

The structure of the paper is as follows: In Section 2, we explain the notation and introduce 2BSDEs together with their associated PDEs. In Section 3, we show that the existence of a sufficiently regular solution to the associated PDE implies the existence of a solution to the 2BSDE. Our main result, Theorem 4.10 in Section 4, shows the converse: If the PDE satisfies suitable Lipschitz and parabolicity conditions together with a comparison principle from the theory of viscosity solutions, then the existence of a solution to the 2BSDE implies that the PDE has a unique continuous viscosity solution v . Moreover, the solution of the 2BSDE can then be written in terms of v and the underlying state process. This implies that the solution of the 2BSDE is unique, and it provides a stochastic representation result for fully nonlinear parabolic PDEs. In Section 5 we discuss Monte Carlo schemes for the numerical solution of such PDEs. In Section 6 we show how the results of the paper can be adjusted to the case of PDEs with boundary conditions.

2 Notation and definitions

Let $d \geq 1$ be a natural number. By \mathcal{M}^d we denote the set of all $d \times d$ matrices with real components. B' is the transpose of a matrix $B \in \mathcal{M}^d$ and $\text{Tr}[B]$ its trace. By $\mathcal{M}_{\text{inv}}^d$ we denote the set of all invertible matrices in \mathcal{M}^d , by \mathcal{S}^d all symmetric matrices in \mathcal{M}^d , and by \mathcal{S}_+^d all positive semi-definite matrices in \mathcal{M}^d . For $B, C \in \mathcal{M}^d$, we write $B \geq C$ if $B - C \in \mathcal{S}_+^d$. For $x \in \mathbb{R}^d$, we set

$$|x| := \sqrt{x_1^2 + \cdots + x_d^2}$$

and for $B \in \mathcal{M}^d$,

$$|B| := \sqrt{\sum_{i,j=1}^d B_{ij}^2}.$$

Equalities and inequalities between random variables are always understood in the almost sure sense. We fix a finite time horizon $T \in (0, \infty)$ and let $(W_t)_{t \in [0, T]}$ be a d -dimensional Brownian motion on a complete probability space (Ω, \mathcal{F}, P) . For $s \in [0, T]$, we set $W_t^s := W_t - W_s$, $t \in [s, T]$ and denote by $\mathbb{F}^{s, T} = (\mathcal{F}_t^s)_{t \in [s, T]}$ the augmented filtration generated by $(W_t^s)_{t \in [s, T]}$.

$\mu : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \rightarrow \mathcal{M}_{\text{inv}}^d$ are assumed to be functions satisfying the following Lipschitz and growth conditions: There exists a constant K with

$$|\mu(x) - \mu(y)| + |\sigma(x) - \sigma(y)| \leq K|x - y| \quad \text{for all } x, y \in \mathbb{R}^d, \quad (2.1)$$

and a constant $p_1 \in [0, 1]$ such that

$$|\mu(x)| + |\sigma(x)| \leq K(1 + |x|^{p_1}) \quad \text{for all } x \in \mathbb{R}^d. \quad (2.2)$$

Then, for every initial condition $(s, x) \in [0, T] \times \mathbb{R}^d$, the forward SDE

$$\begin{aligned} dX_t &= \mu(X_t)dt + \sigma(X_t)dW_t, \quad t \in [s, T], \\ X_s &= x, \end{aligned} \tag{2.3}$$

has a unique strong solution $(X_t^{s,x})_{t \in [s, T]}$; see for instance, Theorem 5.2.9 in Karatzas and Shreve [30]. Note that for existence and uniqueness of a strong solution to the SDE (2.3), $p_1 = 1$ in condition (2.2) is enough. But for $p_1 \in [0, 1)$, we will get a better growth exponent p in Proposition 4.5 below. In any case, the process $(X_t^{s,x})_{t \in [s, T]}$ is adapted to the filtration $\mathbb{F}^{s, T}$, and by Itô's formula, we have for all $\varphi \in C^{1,2}([s, T] \times \mathbb{R}^d)$ and $t \in [s, T]$,

$$\varphi(t, X_t^{s,x}) = \varphi(s, x) + \int_s^t \mathcal{L}\varphi(r, X_r^{s,x}) dr + \int_s^t D\varphi(r, X_r^{s,x})' dX_r^{s,x}, \tag{2.4}$$

where

$$\mathcal{L}\varphi(t, x) = \varphi_t(t, x) + \frac{1}{2} \text{Tr}[D^2\varphi(t, x)\sigma(x)\sigma(x)'],$$

and $D\varphi$, $D^2\varphi$ are the gradient and the matrix of second derivatives of φ with respect to the x -variables.

In the whole paper, $f : [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{S}^d \rightarrow \mathbb{R}$ and $g : \mathbb{R}^d \rightarrow \mathbb{R}$ are continuous functions.

Definition 2.1 *Let $(s, x) \in [0, T] \times \mathbb{R}^d$ and $(Y_t, Z_t, \Gamma_t, A_t)_{t \in [s, T]}$ a quadruple of $\mathbb{F}^{s, T}$ -progressively measurable processes taking values in \mathbb{R} , \mathbb{R}^d , \mathcal{S}^d and \mathbb{R}^d , respectively. Then we call (Y, Z, Γ, A) a solution of the second order backward stochastic differential equation (2BSDE) corresponding to $(X^{s,x}, f, g)$ if*

$$dY_t = f(t, X_t^{s,x}, Y_t, Z_t, \Gamma_t) dt + Z_t' \circ dX_t^{s,x}, \quad t \in [s, T], \tag{2.5}$$

$$dZ_t = A_t dt + \Gamma_t dX_t^{s,x}, \quad t \in [s, T], \tag{2.6}$$

$$Y_T = g(X_T^{s,x}), \tag{2.7}$$

where $Z_t' \circ dX_t^{s,x}$ denotes Fisk–Stratonovich integration, which is related to Itô integration by

$$Z_t' \circ dX_t^{s,x} = Z_t' dX_t^{s,x} + \frac{1}{2} d\langle Z, X^{s,x} \rangle_t = Z_t' dX_t^{s,x} + \frac{1}{2} \text{Tr}[\Gamma_t \sigma(X_t^{s,x})\sigma(X_t^{s,x})'] dt. \tag{2.8}$$

The equations (2.5)–(2.7) can be viewed as a whole family of 2BSDEs indexed by $(s, x) \in [0, T] \times \mathbb{R}^d$. In the following sections, we will show relations between this family of 2BSDEs and the associated PDE

$$-v_t(t, x) + f(t, x, v(t, x), Dv(t, x), D^2v(t, x)) = 0 \quad \text{on } [0, T] \times \mathbb{R}^d, \tag{2.9}$$

with terminal condition

$$v(T, x) = g(x), \quad x \in \mathbb{R}^d. \tag{2.10}$$

Since Z is a semimartingale, the use of the Fisk–Stratonovich integral in (2.5) means no loss of generality, but it simplifies the notation in the PDE (2.9). Alternatively, (2.5) could be written in terms of the Itô integral as

$$dY_t = \tilde{f}(t, X_t^{s,x}, Y_t, Z_t, \Gamma_t) dt + Z_t' dX_t^{s,x} \quad (2.11)$$

for

$$\tilde{f}(t, x, y, z, \gamma) = f(t, x, y, z, \gamma) + \frac{1}{2} \text{Tr} [\gamma \sigma(x) \sigma(x)'] .$$

In terms of \tilde{f} , the PDE (2.9) reads as follows:

$$-v_t(t, x) + \tilde{f}(t, x, v(t, x), Dv(t, x), D^2v(t, x)) - \frac{1}{2} \text{Tr}[D^2v(t, x) \sigma(x) \sigma(x)'] = 0 .$$

Note that the form of the PDE (2.9) does not depend on the functions μ and σ determining the dynamics in (2.3). So, we could restrict our attention to the case where $\mu \equiv 0$ and $\sigma \equiv I_d$, the $d \times d$ identity matrix. But the freedom to choose μ and σ from a more general class provides additional flexibility in the design of the Monte Carlo schemes discussed in Section 5 below.

3 From a solution of the PDE to a solution of the 2BSDE

Assume $v : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ is a continuous function such that

$$v_t, Dv, D^2v, \mathcal{L}Dv \text{ exist and are continuous on } [0, T] \times \mathbb{R}^d ,$$

and v solves the PDE (2.9) with terminal condition (2.10). Then it follows directly from Itô's formula (2.4) that for each pair $(s, x) \in [0, T] \times \mathbb{R}^d$, the processes

$$\begin{aligned} Y_t &= v(t, X_t^{s,x}) , & t \in [s, T] , \\ Z_t &= Dv(t, X_t^{s,x}) , & t \in [s, T] , \\ \Gamma_t &= D^2v(t, X_t^{s,x}) , & t \in [s, T] , \\ A_t &= \mathcal{L}Dv(t, X_t^{s,x}) , & t \in [s, T] , \end{aligned}$$

solve the 2BSDE corresponding to $(X^{s,x}, f, g)$.

4 From a solution of the 2BSDE to a solution of the PDE

In all of Section 4 we assume that

$$f : [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{S}^d \rightarrow \mathbb{R} \quad \text{and} \quad g : \mathbb{R}^d \rightarrow \mathbb{R}$$

are continuous functions that satisfy the following Lipschitz and growth assumptions:

(A1) For every $N \geq 1$ there exists a constant F_N such that

$$|f(t, x, y, z, \gamma) - f(t, x, \tilde{y}, z, \gamma)| \leq F_N |y - \tilde{y}|$$

for all $(t, x, y, \tilde{y}, z, \gamma) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^2 \times \mathbb{R}^d \times \mathcal{S}^d$ with $\max\{|x|, |y|, |\tilde{y}|, |z|, |\gamma|\} \leq N$.

(A2) There exist constants F and $p_2 \geq 0$ such that

$$|f(t, x, y, z, \gamma)| \leq F(1 + |x|^{p_2} + |y| + |z|^{p_2} + |\gamma|^{p_2})$$

for all $(t, x, y, z, \gamma) \in [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{S}^d$.

(A3) There exist constants G and $p_3 \geq 0$ such that

$$|g(x)| \leq G(1 + |x|^{p_3}) \quad \text{for all } x \in \mathbb{R}^d.$$

4.1 Admissible strategies

We fix constants $p_4, p_5 \geq 0$ and denote for all $(s, x) \in [0, T] \times \mathbb{R}^d$ and $m \geq 0$ by $\mathcal{A}_m^{s,x}$ the class of all processes of the form

$$Z_t = z + \int_s^t A_r dr + \int_s^t \Gamma_r dX_r^{s,x}, \quad t \in [s, T],$$

where $z \in \mathbb{R}^d$, $(A_t)_{t \in [s, T]}$ is an \mathbb{R}^d -valued, $\mathbb{F}^{s, T}$ -progressively measurable process, $(\Gamma_t)_{t \in [s, T]}$ is an \mathcal{S}^d -valued, $\mathbb{F}^{s, T}$ -progressively measurable process such that

$$\max\{|Z_t|, |A_t|, |\Gamma_t|\} \leq m(1 + |X_t^{s,x}|^{p_4}) \quad \text{for all } t \in [s, T], \quad (4.1)$$

and

$$|\Gamma_r - \Gamma_t| \leq m(1 + |X_r^{s,x}|^{p_5} + |X_t^{s,x}|^{p_5})(|r - t| + |X_r^{s,x} - X_t^{s,x}|) \quad \text{for all } r, t \in [s, T]. \quad (4.2)$$

Set $\mathcal{A}^{s,x} := \bigcup_{m \geq 0} \mathcal{A}_m^{s,x}$. It follows from the assumptions (A1) and (A2) on f and condition (4.1) on Z that for all $y \in \mathbb{R}$ and $Z \in \mathcal{A}^{s,x}$, the forward SDE

$$dY_t = f(t, X_t^{s,x}, Y_t, Z_t, \Gamma_t) dt + Z'_t \circ dX_t^{s,x}, \quad t \in [s, T], \quad (4.3)$$

$$Y_s = y, \quad (4.4)$$

has a unique strong solution $(Y_t^{s,x,y,Z})_{t \in [s, T]}$ (this can, for instance, be shown with the arguments in the proofs of Theorems 2.3, 2.4 and 3.1 in Chapter IV of Ikeda and Watanabe [28]).

4.2 Auxiliary stochastic target problems

For every $m \geq 0$, we define the functions $V^m, U_m : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ as follows:

$$V^m(s, x) := \inf\{y \in \mathbb{R} : Y_T^{s,x,y,Z} \geq g(X_T^{s,x}) \text{ for some } Z \in \mathcal{A}_m^{s,x}\},$$

and

$$U_m(s, x) := \sup\{y \in \mathbb{R} : Y_T^{s,x,y,Z} \leq g(X_T^{s,x}) \text{ for some } Z \in \mathcal{A}_m^{s,x}\}.$$

Notice that these problems do not fit into the class of stochastic target problems studied by Soner and Touzi [49] and are more in the spirit of Cheridito et al. [13, 14].

Lemma 4.1 (Dynamic Programming Principle)

Let $(s, x, m) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}_+$ and $(y, Z) \in \mathbb{R} \times \mathcal{A}_m^{s,x}$ such that $Y_T^{s,x,y,Z} \geq g(X_T^{s,x})$. Then

$$Y_t^{s,x,y,Z} \geq V^m(t, X_t^{s,x}) \quad \text{for all } t \in (s, T).$$

Proof. Fix $t \in (s, T)$ and denote by $C^d[s, t]$ the set of all continuous functions from $[s, t]$ to \mathbb{R}^d . Since $X_t^{s,x}$ and $Y_t^{s,x,y,Z}$ are \mathcal{F}_t^s -measurable, there exist measurable functions

$$\xi : C^d[s, t] \rightarrow \mathbb{R}^d \quad \text{and} \quad \psi : C^d[s, t] \rightarrow \mathbb{R}$$

such that

$$X_t^{s,x} = \xi(W^{s,t}) \quad \text{and} \quad Y_t^{s,x,y,Z} = \psi(W^{s,t}),$$

where we denote $W^{s,t} := (W_r^s)_{r \in [s,t]}$. The process Z is of the form

$$Z_r = z + \int_s^r A_u du + \int_s^r \Gamma_u dX_u^{s,x}, \quad r \in [s, T],$$

for $z \in \mathbb{R}$, $(A_r)_{r \in [s,T]}$ an \mathbb{R}^d -valued, $\mathbb{F}^{s,T}$ -progressively measurable process, and $(\Gamma_r)_{r \in [s,T]}$ an \mathcal{S}^d -valued, $\mathbb{F}^{s,T}$ -progressively measurable process. Therefore, there exist progressively measurable functions (see Definition 3.5.15 in Karatzas and Shreve [30])

$$\begin{aligned} \zeta, \phi & : [s, T] \times C^d[s, T] \rightarrow \mathbb{R}^d \\ \chi & : [s, T] \times C^d[s, T] \rightarrow \mathcal{S}^d \end{aligned}$$

such that

$$Z_r = \zeta(r, W^{s,T}), \quad A_r = \phi(r, W^{s,T}) \quad \text{and} \quad \Gamma_r = \chi(r, W^{s,T}) \quad \text{for } r \in [s, T].$$

With obvious notation, we define for every $w \in C^d[s, t]$ the \mathbb{R}^d -valued, $\mathbb{F}^{t,T}$ -progressively measurable process $(Z_r^w)_{r \in [t,T]}$ by

$$Z_r^w = \zeta(t, w) + \int_t^r \phi(u, w + W^{t,T}) du + \int_t^r \chi(u, w + W^{t,T}) dX_u^{t,\xi(w)}, \quad r \in [t, T].$$

Let μ be the distribution of $W^{s,t}$ on $C^d[s, t]$. Then, $Z^w \in \mathcal{A}_m^{t,\xi(w)}$ for μ -almost all $w \in C^d[s, t]$, and

$$1 = P \left[Y_T^{s,x,y,Z} \geq g(X_T^{s,x}) \right] = \int_{C^d[s,t]} P \left[Y_T^{t,\xi(w),\psi(w),Z^w} \geq g(X_T^{t,\xi(w)}) \right] d\mu(w).$$

Hence, for μ -almost all $w \in C^d[s, t]$, the control Z^w satisfies

$$P \left[Y_T^{s,\xi(w),\psi(w),Z^w} \geq g(X_T^{s,\xi(w)}) \right] = 1.$$

This shows that $\psi(w) = Y_t^{t,\xi(w),\psi(w),Z^w} \geq V^m(t, \xi(w))$. In view of the definition of the functions ξ and ψ , this implies $Y_t^{s,x,y,Z} \geq V^m(t, X_t^{s,x})$. \square

Since we have no a priori knowledge of any regularity of the functions V^m and U_m , we introduce the semicontinuous envelopes as in Barles and Perthame [6]

$$V_*^m(t, x) := \liminf_{(\tilde{t}, \tilde{x}) \in [0, T], (\tilde{t}, \tilde{x}) \rightarrow (t, x)} V^m(\tilde{t}, \tilde{x}), \quad (t, x) \in [0, T] \times \mathbb{R}^d$$

and

$$U_m^*(t, x) := \limsup_{(\tilde{t}, \tilde{x}) \in [0, T], (\tilde{t}, \tilde{x}) \rightarrow (t, x)} U_m(\tilde{t}, \tilde{x}), \quad (t, x) \in [0, T] \times \mathbb{R}^d.$$

For $s \in [0, T]$, we also need the tighter semicontinuous envelopes

$$V_{*,s}^m(t, x) := \liminf_{(\tilde{t}, \tilde{x}) \in [s, T], (\tilde{t}, \tilde{x}) \rightarrow (t, x)} V^m(\tilde{t}, \tilde{x}), \quad (t, x) \in [s, T] \times \mathbb{R}^d,$$

and

$$U_m^{*,s}(t, x) := \limsup_{(\tilde{t}, \tilde{x}) \in [s, T], (\tilde{t}, \tilde{x}) \rightarrow (t, x)} U_m(\tilde{t}, \tilde{x}), \quad (t, x) \in [s, T] \times \mathbb{R}^d.$$

Note that

$$V_{*,s}^m = V_*^m \quad \text{and} \quad U_m^{*,s} = U_m^* \quad \text{on} \quad (s, T] \times \mathbb{R}^d,$$

whereas

$$V_{*,s}^m \geq V_*^m \quad \text{and} \quad U_m^{*,s} \leq U_m^* \quad \text{on} \quad \{s\} \times \mathbb{R}^d.$$

For the theory of viscosity solutions, we refer to Crandall et al. [16] and the book of Fleming and Soner [25].

Theorem 4.2 *Let $m \geq 0$ and assume there exists an $s \in [0, T]$ such that $V_{*,s}^m$ is \mathbb{R} -valued on $[s, T] \times \mathbb{R}^d$. Then $V_{*,s}^m$ is a viscosity supersolution of the PDE*

$$-v_t(t, x) + \sup_{\beta \in \mathcal{S}_+^d} f(t, x, v(t, x), Dv(t, x), D^2v(t, x) + \beta) = 0 \quad \text{on} \quad [s, T] \times \mathbb{R}^d.$$

Before turning to the proof of this result, let us state the corresponding claim for the value function U_m .

Corollary 4.3 *Let $m \geq 0$ and assume there exists an $s \in [0, T]$ such that $U_m^{*,s}$ is \mathbb{R} -valued on $[s, T] \times \mathbb{R}^d$. Then $U_m^{*,s}$ is a viscosity subsolution of the PDE*

$$-v_t(t, x) + \inf_{\beta \in \mathcal{S}_+^d} f(t, x, v(t, x), Dv(t, x), D^2v(t, x) - \beta) = 0 \quad \text{on} \quad [s, T] \times \mathbb{R}^d. \quad (4.5)$$

Proof. Observe that for all $(t, x) \in [s, T] \times \mathbb{R}^d$,

$$-U_m(t, x) = \inf \left\{ y \in \mathbb{R} : \hat{Y}_T^{t,x,y,Z} \geq -g(X_T^{t,x}) \text{ for some } Z \in \mathcal{A}_m^{t,x} \right\},$$

where for given $(y, Z) \in \mathbb{R} \times \mathcal{A}_m^{t,x}$, the process $\hat{Y}^{t,x,y,Z}$ is the unique strong solution of the SDE

$$\begin{aligned} dY_r &= -f(r, X_r^{t,x}, -Y_r, -Z_r, -\Gamma_r)dr + (Z_r)' \circ dX_r^{t,x}, \quad r \in [t, T], \\ Y_t &= y. \end{aligned}$$

Hence, it follows from Theorem 4.2 that $-U_m^{*,s} = (-U_m)_{*,s}$ is a viscosity supersolution of the PDE

$$-v_t(t, x) - \inf_{\beta \in \mathcal{S}_+^d} f(t, x, -v(t, x), -Dv(t, x), -D^2v(t, x) - \beta) = 0 \quad \text{on } [s, T] \times \mathbb{R}^d,$$

which shows that $U_m^{*,s}$ is a viscosity subsolution of the PDE (4.5) on $[s, T] \times \mathbb{R}^d$. \square

Proof of Theorem 4.2 Choose $(t_0, x_0) \in [s, T] \times \mathbb{R}^d$ and $\varphi \in C^\infty([s, T] \times \mathbb{R}^d)$ such that

$$0 = (V_{*,s}^m - \varphi)(t_0, x_0) = \min_{(t,x) \in [s,T] \times \mathbb{R}^d} (V_{*,s}^m - \varphi)(t, x).$$

Let $(t_n, x_n)_{n \geq 1}$ be a sequence in $[s, T] \times \mathbb{R}^d$ such that $(t_n, x_n) \rightarrow (t_0, x_0)$ and $V^m(t_n, x_n) \rightarrow V_{*,s}^m(t_0, x_0)$. There exist positive numbers $\varepsilon_n \rightarrow 0$ such that for $y_n = V^m(t_n, x_n) + \varepsilon_n$, there exists $Z^n \in \mathcal{A}_m^{t_n, x_n}$ with

$$Y_T^n \geq g(X_T^n),$$

where we denote $(X^n, Y^n) = (X^{t_n, x_n}, Y^{t_n, x_n, y_n, Z^n})$ and

$$Z_r^n = z_n + \int_{t_n}^r A_u^n du + \int_{t_n}^r \Gamma_u^n dX_u^n, \quad r \in [t_n, T].$$

Note that for all n , $\Gamma_{t_n}^n$ is almost surely constant, and $|z_n|, |\Gamma_{t_n}^n| \leq m(1 + |x_n|^{p_4})$ by assumption (4.1). Hence, by passing to a subsequence, we can assume that $z_n \rightarrow z_0 \in \mathbb{R}^d$ and $\Gamma_{t_n}^n \rightarrow \gamma_0 \in \mathcal{S}^d$. Observe that $\alpha_n := y_n - \varphi(t_n, x_n) \rightarrow 0$. We choose a decreasing sequence of numbers $\delta_n \in (0, T - t_n)$ such that $\delta_n \rightarrow 0$ and $\alpha_n/\delta_n \rightarrow 0$. By Lemma 4.1,

$$Y_{t_n+\delta_n}^n \geq V^m(t_n + \delta_n, X_{t_n+\delta_n}^n),$$

and therefore,

$$Y_{t_n+\delta_n}^n - y_n + \alpha_n \geq \varphi(t_n + \delta_n, X_{t_n+\delta_n}^n) - \varphi(t_n, x_n),$$

which, after two applications of Itô's formula, becomes

$$\begin{aligned} \alpha_n &+ \int_{t_n}^{t_n+\delta_n} [f(r, X_r^n, Y_r^n, Z_r^n, \Gamma_r^n) - \varphi_t(r, X_r^n)] dr \\ &+ [z_n - D\varphi(t_n, x_n)]' [X_{t_n+\delta_n}^n - x_n] \\ &+ \int_{t_n}^{t_n+\delta_n} \left(\int_{t_n}^r [A_u^n - \mathcal{L}D\varphi(u, X_u^n)] du \right)' \circ dX_r^n \\ &+ \int_{t_n}^{t_n+\delta_n} \left(\int_{t_n}^r [\Gamma_u^n - D^2\varphi(u, X_u^n)] dX_u^n \right)' \circ dX_r^n \geq 0 \end{aligned} \quad (4.6)$$

It is shown in Lemma 4.4 below that the sequence of random vectors

$$\left(\begin{array}{c} \delta_n^{-1} \int_{t_n}^{t_n+\delta_n} [f(r, X_r^n, Y_r^n, Z_r^n, \Gamma_r^n) - \varphi_t(r, X_r^n)] dr \\ \delta_n^{-1/2} [X_{t_n+\delta_n}^n - x_n] \\ \delta_n^{-1} \int_{t_n}^{t_n+\delta_n} \left(\int_{t_n}^r [A_u^n - \mathcal{L}D\varphi(u, X_u^n)] du \right)' \circ dX_r^n \\ \delta_n^{-1} \int_{t_n}^{t_n+\delta_n} \left(\int_{t_n}^r [\Gamma_u^n - D^2\varphi(u, X_u^n)] dX_u^n \right)' \circ dX_r^n \end{array} \right), \quad n \geq 1, \quad (4.7)$$

converges in distribution to

$$\begin{pmatrix} f(t_0, x_0, \varphi(t_0, x_0), z_0, \gamma_0) - \varphi_t(t_0, x_0) \\ \sigma(x_0)W_1 \\ 0 \\ \frac{1}{2}W_1'\sigma(x_0)'[\gamma_0 - D^2\varphi(t_0, x_0)]\sigma(x_0)W_1 \end{pmatrix}. \quad (4.8)$$

Set $\eta_n = |z_n - D\varphi(t_n, x_n)|$, and assume $\delta_n^{-1/2}\eta_n \rightarrow \infty$ along a subsequence. Then, along a further subsequence, $\eta_n^{-1}(z_n - D\varphi(t_n, x_n))$ converges to some $\eta_0 \in \mathbb{R}^d$ with

$$|\eta_0| = 1. \quad (4.9)$$

Multiplying inequality (4.6) with $\delta_n^{-1/2}\eta_n^{-1}$ and passing to the limit yields

$$\eta_0'\sigma(x_0)W_1 \geq 0,$$

which, since $\sigma(x_0)$ is invertible, contradicts (4.9). Hence, the sequence $(\delta_n^{-1/2}\eta_n)$ has to be bounded, and therefore, possibly after passing to a subsequence,

$$\delta_n^{-1/2}[z_n - D\varphi(t_n, x_n)] \text{ converges to some } \xi_0 \in \mathbb{R}^d.$$

It follows that $z_0 = D\varphi(t_0, x_0)$. Moreover, we can divide inequality (4.6) by δ_n and pass to the limit to get

$$\begin{aligned} & f(t_0, x_0, \varphi(t_0, x_0), D\varphi(t_0, x_0), \gamma_0) - \varphi_t(t_0, x_0) \\ & + \xi_0'\sigma(x_0)W_1 + \frac{1}{2}W_1'\sigma(x_0)'[\gamma_0 - D^2\varphi(t_0, x_0)]\sigma(x_0)W_1 \geq 0. \end{aligned} \quad (4.10)$$

Since the support of the random vector W_1 is \mathbb{R}^d , it follows from (4.10) that

$$\begin{aligned} & f(t_0, x_0, \varphi(t_0, x_0), D\varphi(t_0, x_0), \gamma_0) - \varphi_t(t_0, x_0) \\ & + \xi_0'\sigma(x_0)w + \frac{1}{2}w'\sigma(x_0)'[\gamma_0 - D^2\varphi(t_0, x_0)]\sigma(x_0)w \geq 0, \end{aligned}$$

for all $w \in \mathbb{R}^d$. This shows that

$$f(t_0, x_0, \varphi(t_0, x_0), D\varphi(t_0, x_0), \gamma_0) - \varphi_t(t_0, x_0) \geq 0 \quad \text{and} \quad \beta := \gamma_0 - D^2\varphi(t_0, x_0) \geq 0,$$

and therefore,

$$-\varphi_t(t_0, x_0) + \sup_{\beta \in \mathcal{S}_+^d} f(t_0, x_0, \varphi(t_0, x_0), D\varphi(t_0, x_0), D^2\varphi(t_0, x_0) + \beta) \geq 0.$$

□

Lemma 4.4 *The sequence of random vectors (4.7) converges in distribution to (4.8).*

Proof. With the methods used to solve Problem 5.3.15 in Karatzas and Shreve [30] it can also be shown that for all fixed $q > 0$ and $m \geq 0$, there exists a constant $C \geq 0$ such that for all $0 \leq t \leq r \leq T$, $x \in \mathbb{R}^d$, $y \in \mathbb{R}$ and $Z \in \mathcal{A}_m^{t,x}$,

$$\mathbb{E} \left[\max_{u \in [t, T]} |X_u^{t,x}|^q \right] \leq C(1 + |x|^q) \quad (4.11)$$

$$\mathbb{E} \left[\max_{u \in [t, r]} |X_u^{t,x} - x|^q \right] \leq C(1 + |x|^q)(r - t)^{q/2} \quad (4.12)$$

$$\mathbb{E} \left[\max_{u \in [t, T]} |Y_u^{t,x,y,Z}|^q \right] \leq C(1 + |y|^q + |x|^{\tilde{q}}) \quad (4.13)$$

$$\mathbb{E} \left[\max_{u \in [t, r]} |Y_u^{t,x,y,Z} - y|^q \right] \leq C(1 + |y|^q + |x|^{\tilde{q}})(r - t)^{q/2}, \quad (4.14)$$

where $\tilde{q} := \max \{p_2q, p_2p_4q, (p_4 + 2p_1)q\}$. For every $n \geq 1$, we introduce the $\mathbb{F}^{t_n, T}$ -stopping time

$$\tau_n := \inf \{r \geq t_n : X_r^n \notin B_1(x_0)\} \wedge (t_n + \delta_n),$$

where $B_1(x_0)$ denotes the open unit ball in \mathbb{R}^d around x_0 . It follows from the fact that $x_n \rightarrow x_0$ and (4.12) that

$$P[\tau_n < t_n + \delta_n] \rightarrow 0. \quad (4.15)$$

The difference

$$(X_{t_n + \delta_n}^n - x_n) - \sigma(x_0)(W_{t_n + \delta_n} - W_{t_n})$$

can be written as

$$\int_{t_n}^{t_n + \delta_n} \mu(X_r^n) dr + \int_{t_n}^{t_n + \delta_n} [\sigma(X_r^n) - \sigma(x_n)] dW_r + (\sigma(x_n) - \sigma(x_0))(W_{t_n + \delta_n} - W_{t_n}),$$

and obviously,

$$\frac{1}{\sqrt{\delta_n}} (\sigma(x_n) - \sigma(x_0))(W_{t_n + \delta_n} - W_{t_n}) \rightarrow 0 \quad \text{in } L^2.$$

Moreover, it can be deduced with standard arguments from (2.1), (2.2), (4.11) and (4.12) that

$$\frac{1}{\sqrt{\delta_n}} \int_{t_n}^{t_n + \delta_n} \mu(X_r^n) dr \rightarrow 0 \quad \text{and} \quad \frac{1}{\sqrt{\delta_n}} \int_{t_n}^{\tau_n} [\sigma(X_r^n) - \sigma(x_n)] dW_r \rightarrow 0 \quad \text{in } L^2.$$

This shows that

$$\frac{1}{\sqrt{\delta_n}} \{X_{t_n + \delta_n}^n - x_n - \sigma(x_0)(W_{t_n + \delta_n} - W_{t_n})\} \rightarrow 0 \quad \text{in probability.} \quad (4.16)$$

Similarly, it can be derived from assumption (4.1) on A^n that

$$\frac{1}{\delta_n} \int_{t_n}^{t_n + \delta_n} \left(\int_{t_n}^r [A_u^n - \mathcal{L}D\varphi(u, X_u^n)] du \right)' \circ dX_r^n \rightarrow 0 \quad \text{in probability.} \quad (4.17)$$

By the continuity assumption (4.2) on Γ^n ,

$$\frac{1}{\delta_n} \int_{t_n}^{t_n + \delta_n} \left(\int_{t_n}^r [\Gamma_u^n - \Gamma_{t_n}^n] dX_u^n \right)' \circ dX_r^n \rightarrow 0 \quad \text{in } L^2,$$

and

$$\frac{1}{\delta_n} \left\{ \int_{t_n}^{t_n+\delta_n} \left(\int_{t_n}^r \Gamma_{t_n}^n dX_u^n \right)' \circ dX_r^n - \frac{1}{2} (W_{t_n+\delta_n} - W_{t_n})' \sigma(x_0)' \gamma_0 \sigma(x_0) (W_{t_n+\delta_n} - W_{t_n}) \right\} \rightarrow 0 \quad \text{in probability.}$$

Hence,

$$\frac{1}{\delta_n} \left\{ \int_{t_n}^{t_n+\delta_n} \left(\int_{t_n}^r \Gamma_u^n dX_u^n \right)' \circ dX_r^n - \frac{1}{2} (W_{t_n+\delta_n} - W_{t_n})' \sigma(x_0)' \gamma_0 \sigma(x_0) (W_{t_n+\delta_n} - W_{t_n}) \right\} \rightarrow 0 \quad \text{in probability.} \quad (4.18)$$

Similarly, it can be shown that

$$\frac{1}{\delta_n} \left\{ \int_{t_n}^{t_n+\delta_n} \left(\int_{t_n}^r D^2 \varphi(u, X_u^n) dX_u^n \right)' \circ dX_r^n - \frac{1}{2} (W_{t_n+\delta_n} - W_{t_n})' \sigma(x_0)' D^2 \varphi(t_0, x_0) \sigma(x_0) (W_{t_n+\delta_n} - W_{t_n}) \right\} \rightarrow 0 \quad \text{in probability.} \quad (4.19)$$

Finally, it follows from the continuity of f and φ_t as well as (4.1), (4.2), (4.11), (4.12) and (4.14) that

$$\frac{1}{\delta_n} \int_{t_n}^{t_n+\delta_n} [f(X_r^n, Y_r^n, Z_r^n, \Gamma_r^n) - \varphi_t(r, X_r^n)] dr \rightarrow f(x_0, D\varphi(t_0, x_0), z_0, \gamma_0) - \varphi_t(t_0, x_0) \quad (4.20)$$

in probability. Now, the lemma follows from (4.16)–(4.20) and the simple fact that for each n , the random vector

$$\begin{pmatrix} f(x_0, \varphi(t_0, x_0), z_0, \gamma_0) - \varphi_t(t_0, x_0) \\ \delta_n^{-1/2} \sigma(x_0) (W_{t_n+\delta_n} - W_{t_n}) \\ 0 \\ \delta_n^{-1/2} (W_{t_n+\delta_n} - W_{t_n})' \sigma(x_0)' [\gamma_0 - D^2 \varphi(t_0, x_0)] \sigma(x_0) (W_{t_n+\delta_n} - W_{t_n}) \end{pmatrix}$$

has the same distribution as

$$\begin{pmatrix} f(x_0, \varphi(t_0, x_0), z_0, \gamma_0) - \varphi_t(t_0, x_0) \\ \sigma(x_0) W_1 \\ 0 \\ \frac{1}{2} W_1' \sigma(x_0)' [\gamma_0 - D^2 \varphi(t_0, x_0)] \sigma(x_0) W_1 \end{pmatrix}.$$

□

We conclude this subsection with the following estimates on the growth of the value functions V^m and U_m .

Proposition 4.5 *Let $p = \max\{p_2, p_3, p_2 p_4, p_4 + 2p_1\}$. Then there exists for every $m \geq 0$ a constant $C_m \geq 0$ such that*

$$V_*^m(t, x) \geq -C_m(1 + |x|^p) \quad \text{and} \quad (4.21)$$

$$U_m^*(t, x) \leq C_m(1 + |x|^p) \quad (4.22)$$

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

$$V_*^m(T, x) \geq g(x) \quad \text{and} \quad (4.23)$$

$$U_m^*(T, x) \leq g(x) \quad (4.24)$$

for all $x \in \mathbb{R}$.

Proof. We show (4.22) and (4.24). The proofs of (4.21) and (4.23) are completely analogous. To prove (4.22) it is enough to show that for fixed $m \geq 0$, there exists a constant $C_m \geq 0$ such that for all $(t, x) \in [0, T] \times \mathbb{R}^d$ and $(y, Z) \in \mathbb{R} \times \mathcal{A}_m^{t,x}$ satisfying $Y_T^{t,x,y,Z} \leq g(X_T^{t,x})$, we have

$$y \leq C_m(1 + |x|^p).$$

For $y \leq 0$ there is nothing to show. So, we assume $y > 0$ and introduce the stopping time

$$\tau := \inf \{r \geq t \mid Y_r^{t,x,y,Z} = 0\} \wedge T.$$

Then, we have for all $r \in [t, T]$,

$$\begin{aligned} & Y_{r \wedge \tau}^{t,x,y,Z} + \int_r^{r \vee \tau} f(u, X_u^{t,x}, Y_u^{t,x,y,Z}, Z_u, \Gamma_u) du + \int_r^{r \vee \tau} Z'_u \mu(X_u^{t,x}) du \\ & + \int_r^{r \vee \tau} Z'_u \sigma(X_u^{t,x}) dW_u + \frac{1}{2} \int_r^{r \vee \tau} \text{Tr}[\Gamma_u \sigma(X_u^{t,x}) \sigma(X_u^{t,x})'] du \\ & = Y_{r \wedge \tau}^{t,x,y,Z} + \int_r^{r \vee \tau} f(u, X_u^{t,x}, Y_u^{t,x,y,Z}, Z_u, \Gamma_u) du + \int_r^{r \vee \tau} Z'_u \circ dX_u^{t,x} \\ & = Y_\tau^{t,x,y,Z} \\ & \leq g(X_T^{t,x}) \vee 0. \end{aligned}$$

Hence, it follows from (A2), (A3), (2.2) and (4.1) that for $\tilde{p} = \max\{p_2, p_2 p_4, p_4 + 2p_1\}$,

$$\begin{aligned} & h(r) := \mathbb{E} \left[Y_{r \wedge \tau}^{t,x,y,Z} \right] \\ & \leq \mathbb{E} \left[|g(X_T^{t,x})| \right] + \mathbb{E} \left[\int_r^{r \vee \tau} |f(u, X_u^{t,x}, Y_u^{t,x,y,Z}, Z_u, \Gamma_u)| du \right] \\ & \quad + \mathbb{E} \left[\int_r^{r \vee \tau} |Z'_u \mu(X_u^{t,x})| du \right] + \mathbb{E} \left[\frac{1}{2} \int_r^{r \vee \tau} |\text{Tr}[\Gamma_u \sigma(X_u^{t,x}) \sigma(X_u^{t,x})']| du \right] \\ & \leq G \mathbb{E} \left[1 + |X_T^{t,x}|^{p_3} \right] + F \int_r^T h(u) du + L \int_r^T (1 + |X_u^{t,x}|^{\tilde{p}}) du \\ & \leq \tilde{L}(1 + |x|^p) + F \int_r^T h(u) du, \end{aligned}$$

for constants L and \tilde{L} independent of t, x, y and Z . It follows from Gronwall's lemma that

$$h(r) \leq \tilde{L}(1 + |x|^p) e^{F(T-r)} \quad \text{for all } r \in [t, T].$$

In particular, $y = h(t) \leq C_m(1 + |x|^p)$ for some constant C_m independent of t, x, y and Z .

To prove (4.24), we assume by way of contradiction that there exists an $x \in \mathbb{R}^d$ such that $U_m^*(T, x) \geq g(x) + 3\varepsilon$ for some $\varepsilon > 0$. Then, there exists a sequence $(t_n, x_n)_{n \geq 1}$ in $[0, T) \times \mathbb{R}^d$ converging to (T, x) such that $U_m(t_n, x_n) \geq g(x) + 2\varepsilon$ for all $n \geq 1$. Hence, for every integer $n \geq 1$, there exists a real number $y_n \in [g(x) + \varepsilon, g(x) + 2\varepsilon]$ and a process $Z^n \in \mathcal{A}_m^{t_n, x_n}$ of the form $Z_r^n = z_n + \int_{t_n}^r A_u^n du + \int_{t_n}^r \Gamma_u^n dX_u^{t_n, x_n}$ such that

$$y_n \leq g(X_T^{t_n, x_n}) - \int_{t_n}^T f(r, X_r^{t_n, x_n}, Y_r^{t_n, x_n, y_n, Z^n}, Z_r^n, \Gamma_r^n) dr - \int_{t_n}^T (Z_r^n)' \circ dX_r^{t_n, x_n}. \quad (4.25)$$

By (4.1), (4.11), (4.12) and (4.13), the right-hand side of (4.25) converges to $g(x)$ in probability. Therefore, it follows from (4.25) that $g(x) + \varepsilon \leq g(x)$. But this is absurd, and hence, we must have $U_m^*(T, x) \leq g(x)$ for all $x \in \mathbb{R}^d$. \square

4.3 Main result

For our main result, Theorem 4.10, and the ensuing Corollary 4.11 below we need two more assumptions on the functions f and g , the first of which is

(A4) For all $(t, x, y, z) \in [0, T) \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d$ and $\gamma, \tilde{\gamma} \in \mathcal{S}^d$,

$$f(t, x, y, z, \gamma) \geq f(t, x, y, z, \tilde{\gamma}) \quad \text{whenever} \quad \gamma \leq \tilde{\gamma}.$$

Remark 4.6 Assume (A1)–(A4) and that there exists an $s \in [0, T)$ such that $V_{*,s}^m$ is \mathbb{R} -valued on $[s, T] \times \mathbb{R}^d$. Then it immediately follows from Theorem 4.2 that $V_{*,s}^m$ is a viscosity supersolution of the PDE (2.9) on $[s, T) \times \mathbb{R}^d$. Analogously, if (A1)–(A4) hold and there exists an $s \in [0, T)$ such that $U_m^{*,s}$ is \mathbb{R} -valued on $[s, T] \times \mathbb{R}^d$, Corollary 4.3 implies that $U_m^{*,s}$ is a viscosity subsolution of the PDE (2.9) on $[s, T) \times \mathbb{R}^d$.

For our last assumption and the statement of Theorem 4.10 we need the following

Definition 4.7 Let $s \in [0, T)$ and $q \geq 0$.

1. We call a function $v : [s, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ a viscosity solution with growth exponent q of the PDE (2.9) with terminal condition (2.10) if v is a viscosity solution of (2.9) on $[s, T) \times \mathbb{R}^d$ with $v^*(T, x) = v_*(T, x) = g(x)$ for all $x \in \mathbb{R}^d$ and there exists a constant C such that $|v(t, x)| \leq C(1 + |x|^q)$ for all $(t, x) \in [s, T] \times \mathbb{R}^d$.

2. We say that the PDE (2.9) with terminal condition (2.10) satisfies the comparison principle on $[s, T] \times \mathbb{R}^d$ with growth exponent q if the following holds:

If $w : [s, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ is lower semicontinuous and a viscosity supersolution of (2.9) on $[s, T) \times \mathbb{R}^d$ and $u : [s, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ upper semicontinuous and a viscosity subsolution of (2.9) on $[s, T) \times \mathbb{R}^d$ such that

$$w(T, x) \geq g(x) \geq u(T, x) \quad \text{for all } x \in \mathbb{R}^d$$

and there exists a constant $C \geq 0$ such that

$$w(t, x) \geq -C(1 + |x|^q) \quad \text{and} \quad u(t, x) \leq C(1 + |x|^q) \quad \text{for all } (t, x) \in [s, T) \times \mathbb{R}^d,$$

then $w \geq u$ on $[s, T] \times \mathbb{R}^d$.

With this definition our last assumption on f and g is

(A5) For all $s \in [0, T)$, the PDE (2.9) with terminal condition (2.10) satisfies the comparison principle on $[s, T] \times \mathbb{R}^d$ with growth exponent $p = \max \{p_2, p_3, p_2 p_4, p_4 + 2p_1\}$.

Remarks 4.8

1. The monotonicity assumption (A4) is natural from the PDE viewpoint. It implies that f is elliptic and the PDE (2.9) parabolic. If f satisfies the following stronger version of (A4): there exists a constant $C > 0$ such that

$$f(t, x, y, z, \gamma - B) \geq f(t, x, y, z, \gamma) + C \operatorname{Tr}[B] \tag{4.26}$$

for all $(t, x, y, z, \gamma) \in [0, T) \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{S}^d$ and $B \in \mathcal{S}_+^d$, then the PDE (2.9) is called uniformly parabolic, and there exist general results on existence, uniqueness and smoothness of solutions; see for instance Krylov [31] or Evans [24]. When f is linear in the γ variable (in particular, for the semi- and quasi-linear equations discussed in Subsections 5.2 and 5.3 below), the condition (4.26) essentially guarantees existence, uniqueness and smoothness of solutions to the PDE (2.9)–(2.10); see for instance, Section 5.4 in Ladyzenskaya et al. [32]. Without parabolicity there are no comparison results for PDEs of the form (2.9)–(2.10).

2. Condition (A5) is an implicit assumption on the functions f and g . But we find it more convenient to assume comparison directly in the form (A5) instead of putting technical assumptions on f and g which guarantee that the PDE (2.9) with terminal condition (2.10) satisfies (A5). Several comparison results for nonlinear PDEs are available in the literature; see for example, Crandall et al. [16], Fleming and Soner [25], Caffarelli and Cabré [10]. However, most results are stated for equations in bounded domains. For equations in the whole space, the critical issue is the interplay between the growth of solutions at infinity and the growth of the nonlinearity. We list some typical situations where the comparison principle holds:

a) *Comparison principle with growth exponent 1:* Assume (A1)–(A4) and that there exists a function $h : [0, \infty] \rightarrow [0, \infty]$ with $\lim_{x \rightarrow 0} h(x) = 0$ such that

$$|f(t, x, y, \alpha(x - \tilde{x}), A) - f(t, \tilde{x}, y, \alpha(x - \tilde{x}), B)| \leq h(\alpha|x - \tilde{x}|^2 + |x - \tilde{x}|),$$

for all (t, x, \tilde{x}, y) , $\alpha > 0$ and A, B satisfying

$$-\alpha \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \leq \begin{bmatrix} A & 0 \\ 0 & -B \end{bmatrix} \leq \alpha \begin{bmatrix} I & -I \\ -I & I \end{bmatrix}.$$

Then it follows from Theorem 8.2 in Crandall et al. [16] that equations of the form (2.9)–(2.10) satisfy the comparison principle with growth exponent 0 if the domain is bounded. If the domain is unbounded, it follows from the modifications outlined in Section 5.D of [16] that the PDE (2.9)–(2.10) satisfies the comparison principle with growth exponent 1.

b) *Stochastic control problems.* When the equation (2.9) is a dynamic programming equation related to a stochastic optimal control problem, then a comparison theorem for bounded solutions is given in Fleming and Soner [25], Section V.9, Theorem V.9.1. In this case, f is of the form

$$f(t, x, y, z, \gamma) = \sup_{u \in U} \{ \alpha(t, x, u) + \beta(t, x, u)y + b(t, x, u)'z - \operatorname{Tr}[c(t, x, u)\gamma] \},$$

see Subsection 5.5 below. Theorem V.9.1 in [25] is proved under the assumption that $\beta \equiv 0$, U is a compact set and α, b, c are uniformly Lipschitz and growing at most linearly (see IV (2.1) in [25]). This result can be extended directly to the case where β satisfies a similar condition and to equations related to differential games, that is, when

$$f(t, x, y, z, \gamma) = \sup_{u \in U} \inf_{\tilde{u} \in \tilde{U}} \{ \alpha(t, x, u, \tilde{u}) + \beta(t, x, u, \tilde{u})y + b(t, x, u, \tilde{u})'z - \text{Tr}[c(t, x, u, \tilde{u})\gamma] \} .$$

c) *Unbounded solutions.* Many techniques in dealing with unbounded solutions were developed by Ishii [29] for first order equations (that is, when f is independent of γ). These techniques can be extended to second order equations. Some related results can be found in Barles et al. [5, 4]. In [4], in addition to comparison results for PDEs, one can also find BSDEs based on Markov processes with jumps.

The following simple support result is needed in the proof of Theorem 4.10 and Corollary 4.11.

Lemma 4.9 *Let $(s, x) \in [0, T] \times \mathbb{R}^d$. Then for all $t \in (s, T]$, the random variable $X_t^{s,x}$ has full support in \mathbb{R}^d .*

Proof. By assumption, $\sigma(x)$ is non-degenerate for all $x \in \mathbb{R}^d$. Therefore, the distribution of the process $(X_t^{s,x})_{t \in [s, T]}$ is equivalent to the distribution of the unique strong solution of the SDE

$$X_t = x + \int_s^t \sigma(X_r) dW_r, \quad t \in [s, T];$$

see for instance, result 7.6.4 in Liptser and Shiryaev [34] or Theorem 2.4 in Cheridito et al. [12]. Now the lemma follows from the arguments in the proof of Theorem 3.1 in Stroock and Varadhan [50]. \square

Theorem 4.10 (Uniqueness for 2BSDE)

Assume (A1)–(A5) and that there exists $(s, x) \in [0, T] \times \mathbb{R}^d$ such that the 2BSDE corresponding to $(X^{s,x}, f, g)$ has a solution $(Y^{s,x}, Z^{s,x}, \Gamma^{s,x}, A^{s,x})$ with $Z^{s,x} \in \mathcal{A}^{s,x}$. Then the associated PDE (2.9) with terminal condition (2.10) has a unique viscosity solution v on $[s, T] \times \mathbb{R}^d$ with growth exponent $p = \max\{p_2, p_3, p_2p_4, p_4 + 2p_1\}$, v is continuous on $[s, T] \times \mathbb{R}^d$, and the process $Y^{s,x}$ is of the form

$$Y_t^{s,x} = v(t, X_t^{s,x}), \quad t \in [s, T], \quad (4.27)$$

In particular, $(Y^{s,x}, Z^{s,x}, \Gamma^{s,x}, A^{s,x})$ is the only solution of the 2BSDE corresponding to $(X^{s,x}, f, g)$ with $Z^{s,x} \in \mathcal{A}^{s,x}$.

Proof. Let $(s, x) \in [0, T] \times \mathbb{R}^d$. If $(X^{s,x}, Y^{s,x}, \Gamma^{s,x}, A^{s,x})$ is a solution of the 2BSDE corresponding to $(X^{s,x}, f, g)$ with $Z^{s,x} \in \mathcal{A}_m^{s,x}$ for some $m \geq 0$, then it follows from Lemma 4.1 that

$$Y_t^{s,x} \geq V^m(t, X_t^{s,x}) \quad \text{for all } t \in (s, T), \quad (4.28)$$

and by symmetry,

$$Y_t^{s,x} \leq U_m(t, X_t^{s,x}) \quad \text{for all } t \in (s, T). \quad (4.29)$$

Recall that the inequalities (4.28) and (4.29) are understood in the P -almost sure sense. But, by Lemma 4.9, $X_t^{s,x}$ has full support in \mathbb{R}^d for all $t \in (s, T]$. Therefore, we get from (4.28) and (4.29) that

$$V^m(t, x) \leq U_m(t, x) \quad \text{for all } (t, x) \text{ in a dense subset of } [s, T] \times \mathbb{R}^d.$$

It follows that

$$V_{*,s}^m \leq U_m^{*,s} \quad \text{on } [s, T] \times \mathbb{R}^d.$$

Together with Proposition 4.5, this shows that $V_{*,s}^m$ and $U_m^{*,s}$ are \mathbb{R} -valued on $[s, T] \times \mathbb{R}^d$. By Remark 4.6, $V_{*,s}^m$ is a viscosity supersolution and $U_m^{*,s}$ a viscosity subsolution of the PDE (2.9) on $[s, T] \times \mathbb{R}^d$. Therefore, it follows from Proposition 4.5 and assumption (A5) that

$$V_{*,s}^m \geq U_m^{*,s} \quad \text{on } [s, T] \times \mathbb{R}^d.$$

Hence, the function $v = V_{*,s}^m = V^m = U_m = U_m^{*,s}$ is continuous on $[s, T] \times \mathbb{R}^d$ and a viscosity solution with growth exponent p of the PDE (2.9)–(2.10). By (A5), v is the only viscosity solution of the PDE (2.9)–(2.10) on $[s, T] \times \mathbb{R}^d$ with growth exponent p . From (4.28) and (4.29) we get

$$Y_t^{s,x} = v(t, X_t^{s,x}) \quad \text{for all } t \in [s, T].$$

Now, $Z^{s,x}$ is uniquely determined by

$$Y_t^{s,x} = Y_s^{s,x} + \int_s^t f(r, X_r^{s,x}, Y_r^{s,x}, Z_r^{s,x}, \Gamma_r^{s,x}) dr + \int_s^t (Z_r^{s,x})' \circ dX_r^{s,x}, \quad t \in [s, T],$$

and $\Gamma^{s,x}$ and $A^{s,x}$ are uniquely determined by

$$Z_t^{s,x} = Z_s^{s,x} + \int_s^t A_r^{s,x} dr + \int_s^t \Gamma_r^{s,x} dX_r^{s,x}, \quad t \in [s, T].$$

□

In the subsequent corollary, we use the following notation: $\mathbb{H}^2(\mathbb{R})$, $\mathbb{H}^2(\mathbb{R}^d)$ and $\mathbb{H}^2(\mathcal{M}^d)$ denote the spaces of all $\mathbb{F}^{0,T}$ -progressively measurable processes $(H_t)_{t \in [0, T]}$ with values in \mathbb{R} , \mathbb{R}^d , and \mathcal{M}^d , respectively, such that

$$\|H\|_{\mathbb{H}^2} := \mathbb{E} \left[\int_0^T |H_t|^2 dt \right] < \infty.$$

For $s \in (0, T]$, we extend $\mathbb{F}^{s,T}$ -progressively measurable processes $(H_t)_{t \in [s, T]}$ to the whole time interval $[0, T]$ by setting

$$H_t := H_s \quad \text{for } t \in [0, s).$$

Corollary 4.11 *Assume (A1)–(A5) and there exists $x_0 \in \mathbb{R}^d$ such that the 2BSDE corresponding to (X^{0,x_0}, f, g) has a solution $(Y^{0,x_0}, Z^{0,x_0}, \Gamma^{0,x_0}, A^{0,x_0})$ with $Z^{0,x_0} \in \mathcal{A}^{0,x_0}$. Then for all $(s, x) \in [0, T] \times \mathbb{R}^d$, there exists exactly one solution $(Y^{s,x}, Z^{s,x}, \Gamma^{s,x}, A^{s,x})$ to the 2BSDE corresponding to $(X^{s,x}, f, g)$ such that $Z^{s,x} \in \mathcal{A}^{s,x}$. Furthermore, the mapping*

$$(s, x) \mapsto (Y^{s,x}, Z^{s,x}, \Gamma^{s,x}, A^{s,x})$$

is continuous from $[0, T] \times \mathbb{R}^d$ to $\mathbb{H}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R}^d) \times \mathbb{H}^2(\mathcal{M}^d) \times \mathbb{H}^2(\mathbb{R}^d)$.

Proof. Assume there exists $x_0 \in \mathbb{R}^d$ such that the 2BSDE corresponding to $(X^{0,x}, f, g)$ has a solution $(Y^{0,x_0}, Z^{0,x_0}, \Gamma^{0,x_0}, A^{0,x_0})$ with $Z^{0,x_0} \in \mathcal{A}_m^{0,x_0}$ for some $m \geq 0$. Then, by Theorem 4.10, $v = V^m = U_m$ is a continuous function on $[0, T] \times \mathbb{R}^d$ and the process Y^{0,x_0} is of the form

$$Y_t^{0,x_0} = v(t, X_t^{0,x_0}) \quad \text{for all } t \in [0, T].$$

By Lemma 4.9, X_t^{0,x_0} has full support in \mathbb{R}^d for all $t \in (0, T]$. Hence, it follows from the disintegration argument in the proof of Lemma 4.1 that for all (s, x) in a dense subset \mathcal{D} of $[0, T] \times \mathbb{R}^d$, there exists a solution $(Y^{s,x}, Z^{s,x}, \Gamma^{s,x}, A^{s,x})$ to the 2BSDE corresponding to $(X^{s,x}, f, g)$ such that $Y_t^{s,x} = v(t, X_t^{s,x})$ and $Z^{s,x} \in \mathcal{A}_m^{s,x}$. For arbitrary $(s, x) \in [0, T] \times \mathbb{R}^d$, there exists a sequence $(s_n, x_n)_{n \geq 1}$ in \mathcal{D} converging to (s, x) . To simplify the notation we write X^n for the process X^{s_n, x_n} and $(Y^n, Z^n, \Gamma^n, A^n)$ for $(Y^{s_n, x_n}, Z^{s_n, x_n}, \Gamma^{s_n, x_n}, A^{s_n, x_n})$. In the following we are going to show that

$$(Y^n, Z^n, \Gamma^n, A^n) \rightarrow (Y, Z, \Gamma, A) \quad \text{in } \mathbb{H}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R}^d) \times \mathbb{H}^2(\mathcal{M}^d) \times \mathbb{H}^2(\mathbb{R}^d), \quad (4.30)$$

where $Y_t = v(t, X_t^{s,x})$, $t \in [s, T]$, and (Y, Z, Γ, A) is a solution to the BSDE corresponding to $(X^{s,x}, f, g)$ with $Z \in \mathcal{A}_m^{s,x}$. Then, by Theorem 4.10, (Y, Z, Γ, A) is the only solution of the BSDE corresponding to $(X^{s,x}, f, g)$ with $Z \in \mathcal{A}^{s,x}$, and the corollary follows.

To prove (4.30), we first notice that with the arguments used in the solution of Problem 5.3.15 in Karatzas and Shreve [30] it can be derived from the Lipschitz assumption (2.1) on the coefficients μ and σ that for every $q > 0$, there exists a constant $C_q \geq 0$ such that

$$\mathbb{E} \left[\sup_{t \in [0, T]} |X_t^n - X_t^{s,x}|^q \right] \leq C_q (1 + |x_n|^q + |x|^q) (|s_n - s|^{q/2} + |x_n - x|^q) \quad \text{for all } n \geq 1. \quad (4.31)$$

Since v is continuous and polynomially bounded of order $p = \max\{p_2, p_3, p_2 p_4, p_4 + 2p_1\}$, it follows from (4.31) that for all $q > 0$,

$$\mathbb{E} \left[\sup_{t \in [0, T]} |Y_t^n - Y_t|^q \right] \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

and

$$\mathbb{E} \left[\sup_{t \in [0, T]} |Y_t^n - Y_t^k|^q \right] \rightarrow 0 \quad \text{as } n, k \rightarrow \infty. \quad (4.32)$$

In particular,

$$Y_0^n - Y_0^k \rightarrow 0 \quad \text{and} \quad Y_T^n - Y_T^k \rightarrow 0 \quad \text{in } L^2(\Omega, \mathcal{F}, P) \quad \text{as } n, k \rightarrow \infty.$$

Note that

$$dY_t^n = G_t^n dt + (H_t^n)' dW_t, \quad t \in [0, T], \quad (4.33)$$

for

$$G_t^n = 1_{\{t \geq s_n\}} [f(t, X_t^n, Y_t^n, Z_t^n, \Gamma_t^n) + (Z_t^n)' \mu(X_t^n)]$$

and

$$H_t^n = 1_{\{t \geq s_n\}} \sigma(X_t^{s_n, x_n})' Z_t^{s_n, x_n}.$$

From the growth assumptions (2.2) and (A2) on μ , σ and f , the estimates (4.11) and (4.13) for X^n and Y^n , and assumption (4.1) on Z^n and Γ^n we obtain

$$\sup_{n \geq 1} \mathbb{E} \left[\sup_{t \in [0, T]} \{|G_t^n|^q + |H_t^n|^q\} \right] < \infty \quad \text{for all } q > 0. \quad (4.34)$$

(4.32)–(4.34) yield

$$\int_0^T (Y_t^n - Y_t^k) d(Y_t^n - Y_t^k) \rightarrow 0 \quad \text{in } L^2(\Omega, \mathcal{F}, P) \quad \text{as } k, n \rightarrow \infty,$$

and therefore, by Itô's formula,

$$\int_0^T |H_t^n - H_t^k|^2 dt = (Y_T^n - Y_T^k)^2 - (Y_0^n - Y_0^k)^2 - 2 \int_0^T (Y_t^n - Y_t^k) d(Y_t^n - Y_t^k) \rightarrow 0$$

in $L^1(\Omega, \mathcal{F}, P)$. Since $\mathbb{H}^2(\mathbb{R}^d)$ is a complete metric space, it contains a process $(H_t)_{t \in [0, T]}$ such that $\|H^n - H\|_{\mathbb{H}^2} \rightarrow 0$. Define

$$Z_t := \begin{cases} (\sigma(X_t^{s,x}))^{-1} H_t & \text{for } t \in [s, T] \\ (\sigma(X_s^{s,x}))^{-1} H_s & \text{for } t \in [0, s) \end{cases}.$$

Then, $Z^n \rightarrow Z$ in measure with respect to $P \times dt$ on $\Omega \times [0, T]$. In view of the estimate (4.11) for X^n and assumption (4.1) on Z^n , this implies that for all $q > 0$,

$$\mathbb{E} \left[\int_0^T |Z_t^n - Z_t|^q dt \right] \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

In particular, $\|Z^n - Z\|_{\mathbb{H}^2} \rightarrow 0$ for $n \rightarrow \infty$, and

$$\mathbb{E} \left[\int_0^T |Z_t^n - Z_t^k|^q dt \right] \rightarrow 0 \quad \text{for } n, k \rightarrow \infty. \quad (4.35)$$

Since

$$dZ_t^n = I_t^n dt + J_t^n dW_t, \quad t \in [0, T], \quad (4.36)$$

for

$$I_t^n = 1_{\{t \geq s_n\}} [A_t^n + \Gamma_t^n \mu(X_t^n)] \quad \text{and} \quad J_t^n = 1_{\{t \geq s_n\}} \Gamma_t^n \sigma(X_t^n),$$

it follows from the growth bounds (2.2) and (4.1) on μ , σ , A^n and Γ^n that there exists a constant $L \geq 0$ such that

$$\mathbb{E} \left[|Z_r^n - Z_t^n|^2 \right] \leq L(r - t) \quad \text{for all } n \geq 1 \text{ and } r, t \in [0, T].$$

This and (4.35) show that

$$\left| Z_0^n - Z_0^k \right| \rightarrow 0 \quad \text{and} \quad \left| Z_T^n - Z_T^k \right| \rightarrow 0 \quad \text{in } L^2(\Omega, \mathcal{F}, P) \quad \text{as } n, k \rightarrow \infty. \quad (4.37)$$

By (4.11), (2.2) and (4.1), we also have

$$\sup_{n \geq 1} \mathbb{E} \left[\sup_{t \in [0, T]} \{|I_t^n|^q + |J_t^n|^q\} \right] < \infty \quad \text{for all } q > 0.$$

Thus, it follows from (4.35)–(4.37) that

$$\int_0^T |J_t^n - J_t^k|^2 dt = |Z_T^n - Z_T^k|^2 - |Z_0^n - Z_0^k|^2 - 2 \int_0^T (Z_t^n - Z_t^k)' d(Z_t^n - Z_t^k) \rightarrow 0$$

in $L^1(\Omega, \mathcal{F}, P)$, which implies that $\|J^n - J\|_{\mathbb{H}^2} \rightarrow 0$ for a process $(J_t)_{t \in [0, T]}$ in $\mathbb{H}^2(\mathcal{M}^d)$. Define

$$\Gamma_t := \begin{cases} J_t \sigma(X_t^{s,x})^{-1} & \text{for } t \in [s, T] \\ J_s \sigma(X_s^{s,x})^{-1} & \text{for } t \in [0, s) \end{cases}.$$

Then, $(\Gamma^n)_{n \geq 1}$ converges to Γ in measure with respect to $P \times dt$ on $\Omega \times [0, T]$. By assumption (4.1) on Γ^n , we also have $\|\Gamma^n - \Gamma\|_{\mathbb{H}^2} \rightarrow 0$ for $n \rightarrow \infty$. Now, it can easily be checked that there exists a process $(A_t)_{t \in [0, T]}$ in $\mathbb{H}^2(\mathbb{R}^d)$ such that $\|A^n - A\|_{\mathbb{H}^2} \rightarrow 0$ as $n \rightarrow \infty$. Since all the triples (Z^n, Γ^n, A^n) , $n \geq 1$, satisfy the conditions (4.1) and (4.2), so does (Z, Γ, A) , and it readily follows from all the convergence results in this proof that

$$Z_t = Z_s + \int_s^t A_r dr + \int_s^t \Gamma_r dX_r^{s,x}, \quad t \in [s, T],$$

and

$$Y_t^{s,x} = v(s, x) + \int_s^t f(r, X_r^{s,x}, Y_r^{s,x}, Z_r, \Gamma_r) dr + \int_s^t Z_r' \circ dX_r^{s,x}, \quad t \in [s, T].$$

□

Remark 4.12 If the assumptions of Corollary 4.11 are fulfilled, it follows from (4.27) that $v(s, x) = Y_s^{s,x}$ for all $(s, x) \in [0, T] \times \mathbb{R}^d$. Hence, $v(s, x)$ can be approximated by backward simulation of the process $(Y_t^{s,x})_{t \in [s, T]}$. If v is $C^{1,2}$, it follows from Itô's lemma that $Z_t^{s,x} = Dv(t, X_t^{s,x})$, $t \in [s, T]$. Then, $Dv(s, x)$ can be approximated by backward simulation of $(Z_t^{s,x})_{t \in [s, T]}$. If v and all the components of Dv are $C^{1,2}$, then we also have $\Gamma_t^{s,x} = D^2v(t, X_t^{s,x})$, $t \in [s, T]$, and $D^2v(s, x)$ can be approximated by backward simulation of $(\Gamma_t^{s,x})_{t \in [s, T]}$. A formal discussion of a potential numerical scheme for the backward simulation of the processes $Y^{s,x}$, $Z^{s,x}$ and $\Gamma^{s,x}$ is provided in Subsection 5.4 below.

Remark 4.13 Assume there exists a classical solution v of the PDE (2.9) such that

$$v_t, Dv, D^2v, \mathcal{L}Dv \text{ exist and are continuous on } [0, T] \times \mathbb{R}^d,$$

there exists a constant $m \geq 0$ such that

$$\left. \begin{array}{l} |Dv(t, x)| \\ |D^2v(t, x)| \\ |\mathcal{L}Dv(t, x)| \end{array} \right\} \leq m(1 + |x|^{p_4}) \quad \text{for all } t \in [0, T] \text{ and } x \in \mathbb{R}^d, \quad (4.38)$$

and

$$|D^2v(t, x) - D^2v(s, y)| \leq m(1 + |x|^{p_5} + |y|^{p_5})(|t - s| + |x - y|) \quad (4.39)$$

for all $0 \leq t, s \leq T$ and $x, y \in \mathbb{R}^d$.

Note that (4.39) follows, for instance, if $\frac{\partial}{\partial t}D^2v$ and D^3v exist and

$$\left. \begin{array}{l} \max_{ij} \left| \frac{\partial}{\partial t} (D_{ij}^2 v(t, x)) \right| \\ \max_{ij} |D(D_{ij}^2 v(t, x))| \end{array} \right\} \leq \frac{m}{d} (1 + |x|^{p_5}) \quad \text{for all } 0 \leq t \leq T \text{ and } x \in \mathbb{R}^d.$$

Fix $(s, x) \in [0, T) \times \mathbb{R}^d$. By Section 3, the processes

$$\begin{aligned} Y_t &= v(t, X_t^{s,x}), \quad t \in [s, T], \\ Z_t &= Dv(t, X_t^{s,x}), \quad t \in [s, T], \\ \Gamma_t &= D^2v(t, X_t^{s,x}), \quad t \in [s, T], \\ A_t &= \mathcal{L}Dv(t, X_t^{s,x}), \quad t \in [s, T], \end{aligned}$$

solve the 2BSDE corresponding to $(X^{s,x}, f, g)$. By (4.38) and (4.39), Z is in $\mathcal{A}_m^{s,x}$ (see (4.1) and (4.2)). Hence, if the assumptions of Theorem 4.10 are fulfilled, (Y, Z, Γ, A) is the only solution of the 2BSDE corresponding to $(X^{s,x}, f, g)$ with $Z \in \mathcal{A}^{s,x}$.

5 Monte Carlo methods for the solution of parabolic PDEs

In this section, we provide a formal discussion of the numerical implications of our representation results. We start by recalling some well-known facts in the linear case. Then we review recent advances in the semi- and quasi-linear cases and conclude with the fully nonlinear case related to Theorem 4.10 and Corollary 4.11.

5.1 The linear case

In this subsection, we assume that the function f is of the form

$$f(t, x, y, z, \gamma) = -\alpha(t, x) - \beta(t, x)y - \mu(x)'z - \frac{1}{2} \text{Tr} [\sigma(x)\sigma(x)'\gamma]$$

Then, (2.9) is a linear parabolic PDE. Under standard conditions, it has a smooth solution v , and the Feynman–Kac representation theorem states that for all $(s, x) \in [0, T] \times \mathbb{R}^d$,

$$v(s, x) = \mathbb{E} \left[\int_s^T B_{s,t} \alpha(t, X_t^{s,x}) dt + B_{s,T} g(X_T^{s,x}) \right],$$

where

$$B_{s,t} := \exp \left(\int_s^t \beta(r, X_r^{s,x}) dr \right)$$

(see, for instance, Theorem 5.7.6 in Karatzas and Shreve [30]). This representation suggests a numerical approximation of the function v by means of the so-called Monte Carlo method:

- (i) Given J independent copies $\{X^j, 1 \leq j \leq J\}$ of the process $X^{s,x}$, set

$$\hat{v}^{(J)}(s, x) := \frac{1}{J} \sum_{j=1}^J \int_s^T B_{s,t}^j \alpha(t, X_t^j) dt + B_{s,T}^j g(X_T^j),$$

where $B_{s,t}^j := \exp\left(\int_s^t \beta(r, X_r^j) dr\right)$. Then, it follows from the law of large numbers and the central limit theorem that

$$\hat{v}^{(J)}(s, x) \rightarrow v(s, x) \quad \text{a.s.} \quad \text{and} \quad \sqrt{J} \left(\hat{v}^{(J)}(s, x) - v(s, x) \right) \rightarrow \mathbf{N}(0, \rho) \quad \text{in distribution,}$$

where \mathbf{N} denotes the normal distribution and ρ is the variance of the random variable $\int_s^T B_{s,t} \alpha(t, X_t^{s,x}) dt + B_{s,T} g(X_T^{s,x})$. Hence, $\hat{v}^{(J)}(s, x)$ is a consistent approximation of $v(s, x)$, and in contrast to finite differences or finite elements methods, the error estimate is of order $J^{-1/2}$, independently of the dimension d .

(ii) In practice, it is not possible to produce independent copies $\{X^j, 1 \leq j \leq J\}$ of the process $X^{s,x}$, except in trivial cases. In most cases, the above Monte Carlo approximation is performed by replacing the process $X^{s,x}$ by a suitable discrete-time approximation X^N with time step of order N^{-1} for which independent copies $\{X^{N,j}, 1 \leq j \leq J\}$ can be produced. The simplest discrete-time approximation is the following discrete Euler scheme: Set $X_s^N = x$ and for $1 \leq n \leq N$,

$$X_{t_n}^N = X_{t_{n-1}}^N + \mu(X_{t_{n-1}}^N)(t_n - t_{n-1}) + \sigma(X_{t_{n-1}}^N)(W_{t_n} - W_{t_{n-1}}),$$

where $t_n := s + n(T - s)/N$. We refer to Talay [51] for a survey of the main results in this area.

5.2 The semi-linear case

We next consider the case where f is given by

$$f(t, x, y, z, \gamma) = \varphi(t, x, y, z) - \mu(x)'z - \frac{1}{2} \text{Tr} [\sigma(x)\sigma(x)'\gamma].$$

Then the PDE (2.9) is called semi-linear. We assume that the assumptions of Corollary 4.11 are satisfied. In view of the connection (2.8) between Fisk–Stratonovich and Itô integration, the 2BSDE (2.5)–(2.7) reduces to an uncoupled FBSDE of the form

$$\begin{aligned} dY_t &= \varphi(t, X_t^{s,x}, Y_t, Z_t)dt + Z_t' \sigma(X_t^{s,x})dW_t, \quad t \in [s, T), \\ Y_T &= g(X_T^{s,x}); \end{aligned}$$

compare to Peng [44, 46] and Pardoux and Peng [40]. For $N \geq 1$, we denote $t_n := s + n(T - s)/N$, $n = 0, \dots, N$, and we define the discrete-time approximation Y^N of Y by the backward scheme

$$Y_T^N := g(X_T^{s,x}),$$

and, for $n = 1, \dots, N$,

$$Y_{t_{n-1}}^N := \mathbb{E} \left[Y_{t_n}^N \mid X_{t_{n-1}}^{s,x} \right] - \varphi \left(t_{n-1}, X_{t_{n-1}}^{s,x}, Y_{t_{n-1}}^N, Z_{t_{n-1}}^N \right) (t_n - t_{n-1}) \quad (5.1)$$

$$Z_{t_{n-1}}^N := \frac{1}{t_n - t_{n-1}} (\sigma(X_{t_{n-1}}^{s,x})')^{-1} \mathbb{E} \left[(W_{t_n} - W_{t_{n-1}}) Y_{t_n}^N \mid X_{t_{n-1}}^{s,x} \right]. \quad (5.2)$$

Then, we have

$$\limsup_{N \rightarrow \infty} \sqrt{N} |Y_s^N - v(s, x)| < \infty;$$

see for instance, Zhang [52], Bally and Pagès [3], Bouchard and Touzi [9], Gobet et al. [27]. The practical implementation of this backward scheme requires the computation of the conditional expectations in (5.1) and (5.2). This suggests the use of a Monte Carlo approximation, as in the linear case. But at every time step, one needs to compute conditional expectations based on J independent copies $\{X^j, 1 \leq j \leq J\}$ of the process $X^{s,x}$. Recently, several approaches to this problem have been developed. We refer to Carrière [11], Longstaff and Schwartz [35], Lions and Regnier [33], Bally and Pagès [3], Glasserman [26], Bouchard and Touzi [9].

5.3 The quasi-linear case

It is shown in Antonelli [1] and Ma et al. [36] that coupled FBSDEs of the form

$$\begin{aligned} dX_t &= \mu(t, X_t, Y_t, Z_t)dt + \sigma(t, X_t, Y_t)dW_t, \quad t \in [s, T], \\ X_s &= x \\ dY_t &= \varphi(t, X_t, Y_t, Z_t)dt + Z_t' \sigma(t, X_t, Y_t)dW_t, \quad t \in [s, T], \\ Y_T &= g(X_T), \end{aligned}$$

are related to quasi-linear PDEs of the form (2.9)–(2.10) with

$$f(t, x, y, z, \gamma) = \varphi(t, x, y, z) - \mu(t, x, y, z)'z - \frac{1}{2} \text{Tr} [\sigma(t, x, y)\sigma(t, x, y)'\gamma],$$

see also, Pardoux and Tang [42]. Delarue and Menozzi [21] have used this relation to build a Monte Carlo scheme for the numerical solution of quasi-linear parabolic PDEs.

5.4 The fully nonlinear case

We now discuss the case of a general f as in Section 4. Set

$$\varphi(t, x, y, z, \gamma) := f(t, x, y, z, \gamma) + \mu(x)'z + \frac{1}{2} \text{Tr} [\sigma(x)\sigma(x)'\gamma]. \quad (5.3)$$

Then for all $(s, x) \in [0, T] \times \mathbb{R}^d$ the 2BSDE corresponding to $(X^{s,x}, f, g)$ can be written as

$$dY_t = \varphi(t, X_t^{s,x}, Y_t, Z_t, \Gamma_t)dt + Z_t' \sigma(X_t^{s,x})dW_t, \quad t \in [s, T], \quad (5.4)$$

$$dZ_t = A_t dt + \Gamma_t dX_t^{s,x}, \quad t \in [s, T], \quad (5.5)$$

$$Y_T = g(X_T^{s,x}). \quad (5.6)$$

We assume that the conditions of Corollary 4.11 hold true, so that the PDE (2.9)–(2.10) has a unique viscosity solution v with growth exponent $p = \max\{p_2, p_3, p_2 p_4, p_4 + 2p_1\}$, and for all $(s, x) \in [0, T] \times \mathbb{R}^d$, there exists a unique solution $(Y^{s,x}, Z^{s,x}, \Gamma^{s,x}, A^{s,x})$ to the 2BSDE (5.4)–(5.6) with $Z^{s,x} \in \mathcal{A}^{s,x}$.

Comparing with the backward scheme (5.1)–(5.2) in the semi-linear case, we suggest the following discrete-time approximation of the processes $Y^{s,x}$, $Z^{s,x}$ and $\Gamma^{s,x}$:

$$Y_T^N := g(X_T^{s,x}), \quad Z_T^N := Dg(X_T^{s,x}),$$

and, for $n = 1, \dots, N$,

$$\begin{aligned} Y_{t_{n-1}}^N &:= \mathbb{E} \left[Y_{t_n}^N | X_{t_{n-1}}^{s,x} \right] - \varphi(t_{n-1}, X_{t_{n-1}}^{s,x}, Y_{t_{n-1}}^N, Z_{t_{n-1}}^N, \Gamma_{t_{n-1}}^N)(t_n - t_{n-1}) \\ Z_{t_{n-1}}^N &:= \frac{1}{t_n - t_{n-1}} (\sigma(X_{t_{n-1}}^{s,x}))^{-1} \mathbb{E} \left[(W_{t_n} - W_{t_{n-1}}) Y_{t_n}^N | X_{t_{n-1}}^{s,x} \right], \\ \Gamma_{t_{n-1}}^N &:= \frac{1}{t_n - t_{n-1}} \mathbb{E} \left[Z_{t_n}^N (W_{t_n} - W_{t_{n-1}})' | X_{t_{n-1}}^{s,x} \right] \sigma(X_{t_{n-1}}^{s,x})^{-1}, \end{aligned}$$

where $t_n := s + n(T - s)/n$. A precise analysis of this backward scheme is left for future research. We conjecture that

$$\begin{aligned} Y_s^N &\rightarrow v(s, x) \quad \text{as } N \rightarrow \infty, \\ Z_s^N &\rightarrow Dv(s, x) \quad \text{as } N \rightarrow \infty \quad \text{if } v \text{ is } C^{1,2}, \end{aligned}$$

and

$$\Gamma_s^N \rightarrow D^2v(s, x) \quad \text{as } N \rightarrow \infty$$

if v and all components of Dv are $C^{1,2}$.

5.5 Connections with standard stochastic control

For $s \in [0, T)$, let $\tilde{\mathcal{U}}^s$ be the collection of all $\mathbb{F}^{s,T}$ -progressively measurable processes $(\nu_t)_{t \in [s, T]}$ with values in a given bounded subset $U \subset \mathbb{R}^k$. Let $b : [0, T] \times \mathbb{R}^d \times U \rightarrow \mathbb{R}^d$ and $a : [0, T] \times \mathbb{R}^d \times U \rightarrow \mathcal{S}^d$ be continuous functions, Lipschitz in x uniformly in (t, u) . We call a process $\nu \in \tilde{\mathcal{U}}^s$ an admissible control if

$$\mathbb{E} \left[\int_s^T (|b(t, x, \nu_t)| + |a(t, x, \nu_t)|^2) dt \right] < \infty$$

for all $x \in \mathbb{R}^d$ and denote the class of all admissible controls by \mathcal{U}^s . For every pair of initial conditions $(s, x) \in [0, T] \times \mathbb{R}^d$ and each admissible control process $\nu \in \mathcal{U}^s$, the SDE

$$\begin{aligned} dX_t &= b(t, X_t, \nu_t) dt + a(t, X_t, \nu_t) dW_t, \quad t \in [s, T], \\ X_s &= x, \end{aligned} \tag{5.7}$$

has a unique strong solution, which we denote by $(X_t^{s,x,\nu})_{t \in [s, T]}$. Let $\alpha, \beta : [0, T] \times \mathbb{R}^d \times U \rightarrow \mathbb{R}$ and $g : \mathbb{R}^d \rightarrow \mathbb{R}$ be continuous functions with $\beta \leq 0$, and assume that α and g have quadratic growth in x uniformly in (t, u) . Consider the stochastic control problem

$$v(s, x) := \sup_{\nu \in \mathcal{U}^s} \mathbb{E} \left[\int_s^T B_{s,t}^\nu \alpha(t, X_t^{s,x,\nu}, \nu_t) dt + B_{s,T}^\nu g(X_T^{s,x,\nu}) \right],$$

where

$$B_{s,t}^\nu := \exp \left(\int_s^t \beta(r, X_r^{s,x,\nu}, \nu_r) dr \right), \quad 0 \leq s \leq t \leq T.$$

By the classical method of dynamic programming, the function v can be shown to solve the Hamilton–Jacobi–Bellman equation

$$\begin{aligned} -v_t(t, x) + f(t, x, v(t, x), Dv(t, x), D^2v(t, x)) &= 0 \\ v(T, x) &= g(x), \end{aligned} \tag{5.8}$$

where

$$f(t, x, y, z, \gamma) := \sup_{u \in U} \left\{ \alpha(t, x, u) + \beta(t, x, u)y + b(t, x, u)'z - \frac{1}{2} \text{Tr}[aa'(t, x, u) \gamma] \right\}.$$

This is a fully nonlinear parabolic PDE covered by the class (2.9)–(2.10). Note that f is convex in the triple (y, z, γ) . The semi-linear case is obtained when there is no control on the diffusion part, that is, when $a(t, x, u) = a(t, x)$ is independent of u .

If the value function v has a stochastic representation in terms of a 2BSDE satisfying the assumptions of Theorem 4.10, then the Monte Carlo scheme of the previous subsection can be applied to approximate v .

Under suitable regularity assumptions, the optimal control at time t is known to be of the form

$$\hat{u}(t, x, v(t, x), Dv(t, x), D^2v(t, x)),$$

where \hat{u} is a maximizer of the expression

$$\sup_{u \in U} \left\{ \alpha(t, x, u) + \beta(t, x, u)v(t, x) + b(t, x, u)'Dv(t, x) - \frac{1}{2} \text{Tr}[aa'(t, x, u)D^2v(t, x)] \right\}.$$

Notice that the numerical scheme suggested in Subsection 5.4 calculates the values of the processes $X^{s,x}$, $Y^{s,x}$, $Z^{s,x}$ and $\Gamma^{s,x}$ at each time step. Therefore, it also provides the optimal control by

$$\hat{v}_t = \hat{u}(t, X_t^{s,x}, Y_t^{s,x}, Z_t^{s,x}, \Gamma_t^{s,x}).$$

6 Boundary value problems

In this section, we briefly outline how the results of this paper can be adjusted to boundary value problems. Namely, let $O \subset \mathbb{R}^d$ be an open set. For $(s, x) \in [0, T] \times O$, the process $X^{s,x}$ is given as in (2.3), but we stop it at the boundary of O . Then we extend the terminal condition (2.7) in the 2BSDE to a boundary condition. In other words, we introduce the exit time

$$\theta := \inf \{t \geq s \mid X_t^{s,x} \notin O\}$$

and modify the 2BSDE (2.5)–(2.7) to

$$\begin{aligned} Y_{t \wedge \theta} &= g(X_{T \wedge \theta}^{s,x}) - \int_{t \wedge \theta}^{T \wedge \theta} f(r, X_r^{s,x}, Y_r, \Gamma_r, A_r) dr - \int_{t \wedge \theta}^{T \wedge \theta} Z_r' \circ dX_r^{s,x}, \quad t \in [s, T], \\ Z_{t \wedge \theta} &= Z_{T \wedge \theta} - \int_{t \wedge \theta}^{T \wedge \theta} A_r dr - \int_{t \wedge \theta}^{T \wedge \theta} \Gamma_r dX_r^{s,x}, \quad t \in [s, T]. \end{aligned}$$

Then the corresponding PDE is the same as (2.9)–(2.10), but it only holds in $[0, T] \times O$. Also, the terminal condition $v(T, x) = g(x)$ only holds in O . In addition, the following lateral boundary condition holds

$$v(t, x) = g(x), \quad \text{for all } (t, x) \in [0, T] \times \partial O.$$

All the results of the previous sections can easily be adapted to this case. Moreover, if we assume that O is bounded, most of the technicalities related to the growth of solutions are avoided as the solutions are expected to be bounded.

Acknowledgements. Parts of this paper were completed during a visit of Soner and Touzi to the Department of Operations Research and Financial Engineering at Princeton University. They would like to thank Princeton University, Erhan Çinlar and Patrick Cheridito for the hospitality. We also thank Nicole El Karoui, Arturo Kohatsu-Higa and Shige Peng for stimulating and helpful discussions.

References

- [1] Antonelli, F. (1993). Backward-forward stochastic differential equations. *Ann. Appl. Probab.* 3(3), 777–793.
- [2] Arkin, V. and Saksonov, M. (1979). Necessary optimality conditions in problems of the control of stochastic differential equations. *Dokl. Akad. Nauk SSSR* 244(1), 11–15.
- [3] Bally, V., Pagès, G. (2003). Error analysis of the quantization algorithm for obstacle problems. *Stoch. Proc. Appl.* 106(1), 1–40.
- [4] Barles, G., Biton, S., Bourgoing, M., Ley, O. (2003). Uniqueness results for quasilinear parabolic equations through viscosity solutions’ methods. *Calc. Var. PDEs*, 18(2), 159–179.
- [5] Barles, G., Buckdahn, R., and Pardoux, E. (1997). Backward stochastic differential equations and integral-partial differential equations, *Stochastics Stochastics Rep.*, 60(1-2), 57–83.
- [6] Barles, G. and Perthame, B. (1988). Exit time problems in optimal control and vanishing viscosity method. *SIAM J. Cont. Opt.* 26(5), 1133–1148.
- [7] Bismut, J.M. (1973). Conjugate convex functions in optimal stochastic control. *J. Math. Anal. Appl.* 44, 384–404.
- [8] Bismut, J.M. (1978). Contrôle des système linéaire quadratiques: applications de l’intégrale stochastique, *Sém. Probab. XII, Lect. Notes. Math.* vol. 649. Springer-Verlag, 180–264.
- [9] Bouchard, B., Touzi, N. (2004). Discrete-time approximation and Monte Carlo simulation of backward stochastic differential equations. *Stoch. Proc. Appl.* 111(2), 175–206.
- [10] Caffarelli, L., Cabré, X. (1995). *Fully Nonlinear Elliptic Equations*. AMS, Providence, RI.
- [11] Carrière, J.F. (1996). Valuation of the early-exercise price for options using simulations and nonparametric regression. *Insurance Math. Econom.* 19(1), 19–30.

- [12] Cheridito, P., Filipović, D., Yor, M. (2005). Equivalent and absolutely continuous measure changes for jump-diffusion processes. *Ann. Appl. Probab.* 15(3), 1713–1732.
- [13] Cheridito, P., Soner, H.M., Touzi, N. (2005). Small time path behavior of double stochastic integrals and applications to stochastic control, *Ann. Appl. Probab.* 15(4), 2472–2495.
- [14] Cheridito, P., Soner, H.M., Touzi, N. (2005). The multi-dimensional super-replication problem under gamma constraints, *Annales de l'Institut Henri Poincaré (C) Non Linear Analysis* 22(5), 633–666.
- [15] Chevance, D. (1997). Numerical Methods for Backward Stochastic Differential Equations. *Publ. Newton Inst.*, Cambridge University Press.
- [16] Crandall, M.G., Ishii, H., Lions, P.L. (1992). User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc.* 27(1), 1–67.
- [17] Cvitanović, J., Ma, J. (1996). Hedging options for a large investor and forward-backward SDE's, *Ann. Appl. Probab.* 6(2), 370–398.
- [18] Cvitanović, J., Karatzas, I. (1996). Backward stochastic differential equations with reflection and Dynkin games. *Ann. Probab.*, 24(4), 2024–2056.
- [19] Cvitanović, J., Karatzas, I., Soner, H.M. (1998). Backward SDEs with constraints on the gains-process, *Ann. Probab.* 26(4), 1522–1551.
- [20] Delarue, F. (2002). On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case, *Stoch. Proc. Appl.* 99(2), 209–286.
- [21] Delarue, F., Menozzi, F. (2006). A forward-backward stochastic algorithm for quasi-linear PDEs. *Ann. Appl. Probab.* 16(1), 140–184.
- [22] Douglas, J., Ma, J., Protter, P. (1996). Numerical methods for forward-backward stochastic differential equations. *Ann. Appl. Probab.* 6(3), 940–968.
- [23] El Karoui, N., Peng, S., Quenez, M.C. (1997). Backward stochastic differential equations in finance. *Mathematical Finance* 7(1), 1–71.
- [24] Evans, L.C. (1998). *Partial Differential Equations*. AMS, Providence, RI.
- [25] Fleming, W.H., Soner, H.M. (1993). *Controlled Markov Processes and Viscosity Solutions*. Applications of Mathematics 25. Springer-Verlag, New York.
- [26] Glasserman, P. (2003). *Monte Carlo Methods in Financial Engineering*. Applications of Mathematics 53. Springer-Verlag, New York.
- [27] Gobet, E., Lemor, J.P, Warin, X. (2005). A regression-based Monte Carlo method to solve backward stochastic differential equations. *Ann. Appl. Probab.* 15(3), 2172–2202.

- [28] Ikeda, N., Watanabe, S. (1989). *Stochastic Differential Equations and Diffusion Processes*, Second Edition. North-Holland Publishing Company.
- [29] Ishii, H. (1984). Uniqueness of unbounded viscosity solutions of Hamilton–Jacobi equations. *Indiana Univ. Math. J.* 33(5), 721–748.
- [30] Karatzas, I., Shreve, S. (1991). *Brownian Motion and Stochastic Calculus*, Second Edition. Springer-Verlag.
- [31] Krylov, N. (1987). *Nonlinear Elliptic and Parabolic Partial Differential Equations of Second Order*. Mathematics and its Applications 7, D. Reidel Publishing Co., Dordrecht.
- [32] Ladyzenskaya, O.A., Solonnikov, V.A., Uralceva, N.N. (1967). *Linear and Quasilinear Equations of Parabolic Type*. AMS, Providence, RI.
- [33] Lions, P.L., Regnier, H. (2001). Calcul du prix et des sensibilités d’une option américaine par une méthode de Monte Carlo. Preprint.
- [34] Liptser, R.S., Shiryaev, A.N. (2001). *Statistics of Random Processes. I. General Theory*, Second Edition. Applications of Mathematics 5. Springer-Verlag, Berlin.
- [35] Longstaff, F.A., Schwartz E.S. (2001). Valuing American options by simulation: a simple least-squares approach. *Rev. Fin. Stud.* 14(1), 113–147.
- [36] Ma J., Protter, P., Yong, J. (1994). Solving forward-backward stochastic differential equations explicitly – a four step scheme. *Prob. Theory Rel. Fields* 98(3), 339–359.
- [37] Ma, J., Yong, J. (1999). *Forward-Backward Stochastic Differential Equations and their Applications*. Lect. Notes Math. 1702, Springer-Verlag.
- [38] Ma, J., Protter, P., San Martín, J., Torres, S. (2002). Numerical methods for backward stochastic differential equations, *Ann. Appl. Probab.* 12(1), 302–316.
- [39] Pardoux, E., Peng, S. (1990). Adapted solution of a backward stochastic differential equation. *Systems Control Lett.* 14(1), 55–61.
- [40] Pardoux, E., Peng, S. (1992). *Backward stochastic differential equations and quasilinear parabolic partial differential equations*. Lect. Notes Contr. Inf. Sci. 176. Springer-Verlag, 200–217.
- [41] Pardoux, E., Peng, S. (1994). Backward doubly stochastic differential equations and systems of quasilinear SPDEs, *Probab. Theory Rel. Fields* 98(2), 209–227.
- [42] Pardoux, E., Tang, S. (1999). Forward-backward stochastic differential equations and quasilinear parabolic PDEs. *Prob. Th. Rel. Fields* 114(2), 123–150.
- [43] Peng, S. (1990). A general stochastic maximum principle for optimal control problems. *SIAM J. Control Optim.* 28(4), 966–979.

- [44] Peng, S. (1991). Probabilistic interpretation for systems of quasilinear parabolic partial differential equations. *Stochastics Stochastics Rep.* 37(1-2), 61–74.
- [45] Peng, S. (1992) Stochastic Hamilton–Jacobi–Bellman equations. *SIAM J. Control Optim.* 30(2), 284–304.
- [46] Peng, S. (1992). A generalized dynamic programming principle and Hamilton–Jacobi–Bellman equation. *Stochastics Stochastics Rep.* 38(2), 119–134.
- [47] Peng, S. (1992). A nonlinear Feynman–Kac formula and applications. *Control Theory, Stochastic Analysis and Applications*. World Sci. Publishing, River Edge, NJ. 173–184.
- [48] Peng, S. (1993). Backward stochastic differential equations and application to optimal control. *Appl. Math. Optim.* 27(2), 125–144.
- [49] Soner, H.M., Touzi, N. (2002). Stochastic target problems, dynamic programming, and viscosity solutions. *SIAM J. on Control Optim.* 41(2), 404–424.
- [50] Stroock, D.W., Varadhan, S.R.S (1972). On the support of diffusion processes with applications to the strong maximum principle. Proc. Sixth Berkeley Symp. Math. Stat. Prob. III. Univ. California Press, Berkeley, 333–359.
- [51] Talay, D. (1996). Probabilistic numerical methods for partial differential equations: elements of analysis, in *Probabilistic Models for Nonlinear Partial Differential Equations*, D. Talay and L. Tubaro, editors, Lect. Notes Math. 1627, 48–196.
- [52] Zhang, J. (2001). Some Fine Properties of Backward Stochastic Differential Equations. PhD Thesis, Purdue University.