Chapter 2
Introduction to Forward-Backward Stochastic Differential Equations

Forward-Backward Stochastic Differential Equations (FBSDEs) provide a powerful modelling tool that has been intensively used in various areas of stochastic control and, in particular, in mathematical finance. They were first introduced by Bismut [3, 4] and then studied in a general way by Pardoux and Peng [41]. Since then, FBSDEs have attracted a lot of interest.[1] Although the basic theory is now well understood, new questions or applications arise every day, making it a very active field of research. Their link with a class of non-linear PDEs is also very fruitful and has led to the design of probabilistic methods for solving such PDEs, as discussed in the next chapter.

We present here a short self-contained introduction to FBSDEs which should be enough to grasp the main concepts presented in the subsequent chapters of this brief. Suggested lectures, on top of the main research articles, are [25, 38, 43].

Throughout this chapter, we let $(\Omega, \mathscr{A}, \mathbb{P})$ be a complete probability space supporting a d-dimensional Brownian Motion W. We shall denote by $\mathbb{F} = \{\mathscr{F}_t\}_{t \geq 0}$ the natural (augmented) filtration of W. Adaptedness and other measurability properties of processes have to be understood with respect to $\mathbb{F}$. When this is not the case (and when it matters), it will be clearly pointed out in the text.

[1] Pardoux and Peng counts more than 2000 citations as of March 2017.

J.-F. Chassagneux et al., *A Forward-Backward SDEs Approach to Pricing in Carbon Markets*, SpringerBriefs in Mathematics of Planet Earth,
DOI 10.1007/978-3-319-63115-8_2

2.1 Backward Stochastic Differential Equations

For a prescribed terminal time $T > 0$, the solution of a backward stochastic differential equation is a pair (Y, Z) satisfying on $[0, T]$

$$\begin{cases} dY_t &= -f(t, Y_t, Z_t)dt + Z_t dW_t, \\ Y_T &= \xi, \end{cases}$$

for some progressively measurable random function f, called the *driver*, and a terminal condition ξ which is a $\mathscr{F}_T$-measurable random variable. It is reasonable to assume that a solution satisfies some conditions so that the various integrals appearing above make sense, and this will be discussed below.

The first peculiarity of BSDEs is that contrary to (forward) SDEs, the solution is not known at the initial time 0 but at the terminal time T. The second difference with forward SDEs comes from the fact that the solution is a pair (Y, Z). Before giving some general existence and uniqueness results and stating precisely some assumptions on the coefficients, we will comment on the shape of the equation and give some hints on the extra process Z.

The simplest example is $f \equiv 0$ and $\xi \in L^2(\mathscr{F}_T)$, where for $t \in [0, T]$, $L^2(\mathscr{F}_t)$ stands for the set of square integrable $\mathscr{F}_t$ measurable random variables. Then, the natural solution to the differential equation $\frac{dY_t}{dt} = 0$ and $Y_T = \xi$ is $Y_t = \xi$, which is generally not adapted (unless ξ is deterministic). The best approximation—say in L^2—is given by the martingale $Y_t = \mathbb{E}[\xi \mid \mathscr{F}_t]$. Using the martingale representation theorem, we introduce a Z-process which is square integrable

$$Y_t = \mathbb{E}[\xi \mid \mathscr{F}_t] = \mathbb{E}[\xi] + \int_0^t Z_s \, dW_s,$$

leading to

$$Y_t = \xi - \int_t^T Z_s \, dW_s, \qquad \text{i.e.} \quad -dY_t = -Z_t \, dW_t, \quad \text{with} \quad Y_T = \xi.$$

We observe in this example that the role of Z is to guarantee the adaptedness of Y. In particular, a deterministic terminal condition will lead to $Z \equiv 0$. This basic example can already be linked to pricing in complete financial markets: The Y represents the price of a contingent claim with random terminal payoff ξ and the Z is linked to the replication portfolio. We discuss this financial application in detail in Sect. 2.1.2.

We now present the basic well-posedness results for BSDEs.

2.1.1 Well-Posedness of BSDEs

Here we present results in the Lipschitz setting, which were first studied in [41], see also [25]. This setting allows us to present the theory in a quite advanced and useful form without encountering too many complications. Moreover, this is the main framework generally adopted for numerical studies.

In order to state precisely the main existence and uniqueness result for BSDEs in the Lipschitz framework, we have to introduce some notation and assumptions.

- We denote by $\mathscr{S}^2(\mathbb{R}^k)$ the vector space of RCLL[2] adapted processes Y, with values in $\mathbb{R}^k$, and such that:

$$\|Y\|^2_{\mathscr{S}^2} := \mathbb{E}\left[\sup_{0\le t\le T} |Y_t|^2\right] < \infty,$$

  and $\mathscr{S}^2_{\mathrm{c}}(\mathbb{R}^k)$ is the subspace of continuous processes.
- The set $\mathscr{H}^2(\mathbb{R}^{k\times d})$ is the set of $\mathbb{R}^{k\times d}$-valued progressively measurable Z-processes such that

$$\|Z\|^2_{\mathscr{H}^2} := \mathbb{E}\left[\int_0^T |Z_t|^2 dt\right] < \infty,$$

  where for $z \in \mathbb{R}^{k\times d}$, $|z|^2 = \mathrm{Tr}(zz^\dagger)$.

We shall often omit $\mathbb{R}^k$ and $\mathbb{R}^{k\times d}$; the spaces $\mathscr{S}^2$, $\mathscr{S}^2_{\mathrm{c}}$ and $\mathscr{H}^2$ are Banach spaces.

A random $\mathbb{R}^k$-valued function f defined on $[0, T] \times \Omega \times \mathbb{R}^k \times \mathbb{R}^{k\times d}$ is such that for all $(y, z) \in \mathbb{R}^k \times \mathbb{R}^{k\times d}$, the process $\{f(t, y, z)\}_{0\le t\le T}$ is progressively measurable. We also assume that

$(H1)$: There exists a positive constant L such that $\mathbb{P}$ a.s.:

1. Lipschitz continuity in (y, z): for all t, y, y', z, z',

$$\left|f(t, y, z) - f(t, y', z')\right| \le L\left(|y - y'| + \|z - z'\|\right);$$

2. Integrability condition:

$$\mathbb{E}\left[|\xi|^2 + \int_0^T |f(r, 0, 0)|^2 dr\right] < \infty.$$

Theorem 2.1 *Under $(H1)$, there exists a unique solution $(Y, Z) \in \mathscr{S}^2_c \times \mathscr{H}^2$ to*

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)\mathrm{d}s - \int_t^T Z_s \mathrm{d}W_s\,, \quad 0 \le t \le T. \tag{2.1}$$

---

[2] Right Continuous with Left Limits.

We skip the proof of the above theorem, which is based on a contraction mapping argument, see the proof of Theorem 2.1 in [25]. We will give a proof, using essentially the same arguments, for an existence and uniqueness result in a (slightly) more involved setting below, see Theorem 2.6.

2.1.1.1 Linear BSDEs

We first study linear BSDEs for which we can give an almost explicit solution. For this section, we set $k = 1$: Y is then real-valued and Z a d-dimensional row vector.

Proposition 2.1 *Let $\{(a_t, b_t)\}_{t\in[0,T]}$ be progressively measurable and bounded processes with values in $\mathbb{R}\times\mathbb{R}^d$. Let $\{c_t\}_{t\in[0,T]}$ be an element of $\mathscr{H}^2(\mathbb{R})$ and ξ a square integrable $\mathscr{F}_T$-measurable random variable.*

The linear BSDE

$$Y_t = \xi + \int_t^T \{a_r Y_r + Z_r b_r + c_r\}\, dr - \int_t^T Z_r\, dW_r \tag{2.2}$$

has a unique solution given by:

$$\forall t \in [0, T], \qquad Y_t = \Gamma_t^{-1}\, \mathbb{E}\left[\xi \Gamma_T + \int_t^T c_r \Gamma_r\, dr \,\middle|\, \mathscr{F}_t\right], \tag{2.3}$$

where for all $t \in [0, T]$,

$$\Gamma_t = \exp\left\{\int_0^t b_r dW_r - \frac{1}{2}\int_0^t |b_r|^2\, dr + \int_0^t a_r\, dr\right\}.$$

Proof We first write down the dynamics of Γ:

$$d\Gamma_t = \Gamma_t\left(a_t\, dt + b_t dW_t\right), \qquad \Gamma_0 = 1.$$

Using Doob's inequality, we easily see that $\Gamma \in \mathscr{S}_c^2$, as b is bounded. It is also clear that there is a unique solution to (2.1): define $f(t, y, z) = a_t y + z b_t + c_t$, which obviously satisfies $(H1)$, and we know that $Y \in \mathscr{S}_c^2$.
Using the product formula, we compute

$$d(\Gamma_t Y_t) = \Gamma_t\, dY_t + Y_t\, d\Gamma_t + d\langle \Gamma, Y\rangle_t = -\Gamma_t c_t\, dt + \Gamma_t Z_t\, dW_t + \Gamma_t Y_t b_t dW_t,$$

showing that $\Gamma_t Y_t + \int_0^t c_r \Gamma_r \, dr$ is a local martingale, which is, in fact, a martingale as $c \in \mathscr{H}^2$ and Γ, Y are in $\mathscr{S}^2$. Then,

$$\Gamma_t Y_t + \int_0^t c_r \Gamma_r \, dr = \mathbb{E}\Big[\Gamma_T Y_T + \int_0^T c_r \Gamma_r \, dr \, \Big| \, \mathscr{F}_t \Big],$$

which concludes the proof. □

Linear BSDEs play an important role in the theory as outlined by the results below, especially via the Comparison Theorem. They were also the first BSDEs to be introduced by Bismut in [3] to study the quadratic-linear stochastic control problem. They appear there as the adjoint process in the variational characterisation of an optimal control.

Remark 2.1 Observe that $\xi \geq 0$ and $c_t \geq 0$ leads to $Y_t \geq 0$.

2.1.1.2 The Comparison Theorem

This section presents the "comparison theorem", which allows us to compare two solutions of two BSDEs (in $\mathbb{R}$) as soon as we can compare the terminal conditions and the drivers of the BSDEs.

Theorem 2.2 *Let $k = 1$ and assume that (ξ', f') satisfies $(H1)$, the solution to the associated BSDE is denoted (Y', Z'). Let (Y, Z) be a solution of a BSDE with parameters (ξ, f) and satisfying $\int_0^T f(t, Y_t, Z_t)\mathrm{d}t \in L^2(\mathscr{F}_T)$. We also assume that $\mathbb{P}$ a.s. $\xi \leq \xi'$ and $f(t, Y_t, Z_t) \leq f'(t, Y_t, Z_t)$ $\lambda \otimes \mathbb{P}$-a.e. ($\lambda$ denoting the Lebesgue measure). Then,*

$$\mathbb{P}\, a.s., \qquad \forall t \in [0, T], \qquad Y_t \leq Y'_t.$$

If, moreover, $Y_0 = Y'_0$, then $\mathbb{P}\, a.s.$, $Y_t = Y'_t$, $0 \leq t \leq T$ and $f(t, Y_t, Z_t) = f'(t, Y_t, Z_t)$ $\lambda \otimes \mathbb{P}$-a.e. In particular, as soon as $\mathbb{P}\left(\xi < \xi'\right) > 0$ or $f(t, Y_t, Z_t) < f'(t, Y_t, Z_t)$ on a set with positive $\lambda \otimes \mathbb{P}$-measure then $Y_0 < Y'_0$.

Proof The proof uses a linearisation argument. Defining $U = Y' - Y$; $V = Z' - Z$ and $\zeta = \xi' - \xi$, we have

$$U_t = \zeta + \int_t^T \big(f'(r, Y'_r, Z'_r) - f(r, Y_r, Z_r) \big) \, dr - \int_t^T V_r \, dW_r.$$

We observe that

$$\begin{aligned} f'(r, Y'_r, Z'_r) - f(r, Y_r, Z_r) = \; & f'(r, Y'_r, Z'_r) - f'(r, Y_r, Z'_r) + f'(r, Y_r, Z'_r) - f'(r, Y_r, Z_r) \\ & + f'(r, Y_r, Z_r) - f(r, Y_r, Z_r) \;\; \text{(non-negative)}. \end{aligned}$$

We introduce a and b: a is $\mathbb{R}$-valued and b a d-dimensional vector. We set

$$a_r := \frac{f'(r, Y'_r, Z'_r) - f'(r, Y_r, Z'_r)}{U_r} \mathbf{1}_{\{U_r \neq 0\}}.$$

For $0 \le i \le d$, we consider the vector $Z_r^{(i)}$ whose last $d - i$ components are those of Z'_r and the first i components are those of Z_r. For $1 \le i \le d$, we set

$$b_r^i = \frac{f'\left(r, Y_r, Z_r^{(i-1)}\right) - f'\left(r, Y_r, Z_r^{(i)}\right)}{V_r^i} \mathbf{1}_{\{V_r^i \neq 0\}}.$$

Importantly, as f' is Lipschitz, the two processes are bounded and progressively measurable. We then observe that

$$U_t = \zeta + \int_t^T (a_r U_r + V_r b_r + c_r)\, dr - \int_t^T V_r\, dW_r,$$

where $c_r = f'(r, Y_r, Z_r) - f(r, Y_r, Z_r)$. By assumption, we have $\zeta \ge 0$ and $c_r \ge 0$. Using the formula given in Proposition 2.1, we have, for all $t \in [0, T]$,

$$U_t = \Gamma_t^{-1} \mathbb{E}\left[\zeta \Gamma_T + \int_t^T c_r \Gamma_r\, dr \,\middle|\, \mathscr{F}_t\right],$$

with, for $0 \le r \le T$,

$$\Gamma_r = \exp\left\{\int_0^r b_u dW_u - \frac{1}{2}\int_0^r |b_u|^2\, du + \int_0^r a_u\, du\right\}.$$

Following Remark 2.1, we get that $U_t \ge 0$, which proves the first statement of the theorem.

Moreover, if $U_0 = 0$ then we have

$$0 = \mathbb{E}\left[\zeta \Gamma_T + \int_0^T c_r \Gamma_r\, dr\right],$$

and the random variable is non-negative. Then, it is equal to zero $\mathbb{P}$ a.s., which implies $\zeta = 0$ and $c_r = 0$, concluding the proof of the theorem. □

2.1.2 Application to Non-linear Pricing

In this section, we study an application of BSDEs in Mathematical Finance, namely the pricing of European contingent claims. We first present the framework of the linear pricing rule in a perfect market and the corresponding linear BSDE. We then introduce some imperfection in the market and show that the option price is still given by a BSDE but with a non-linear Lipschitz driver [25].

2.1.2.1 Super-Replication in a Perfect Market

The market, in its simplest setting, is consists of two assets: a non-risky asset (bank account) delivering an interest rate r, which is a deterministic quantity, and a risky asset (a stock) whose price at any time t is given by S_t. The stochastic process S has the following Black–Scholes type dynamics:

$$S_t = S_0 + \int_0^t r S_s ds + \int_0^t S_s \sigma_s dW_s \ , \ \ S_0 \in (0, \infty).$$

The random coefficient σ is essentially bounded and satisfies $\sigma_s \geq \varepsilon > 0$ for all $s \in [0, T]$. The fact that the drift is rS implies that the dynamics of asset price is already written under the risk neutral probability.

We study the price of a contingent claim that has maturity T and random payoff ξ, with ξ belonging to $L^2(\mathscr{F}_T)$. The goal is to construct an asset portfolio that will perfectly replicate the random payoff ξ. In our setting, a portfolio is described by a stochastic process (α, ϕ) where

- α is the amount of money in the bank account;
- ϕ is the amount invested in the risky asset.

At time t, its value is given by

$$V_t = \alpha_t + \phi_t \ . \tag{2.4}$$

On an infinitesimal time interval dt, the variation in value of the bank account is given by $\alpha_t r \mathrm{d}t$ and the variation in value due to the risky asset is given by $\frac{\phi_t}{S_t}\mathrm{d}S_t$ (due to price change). The stochastic process (α, β) is a strategy that controls the value of the portfolio, but not all strategies can be used. For modelling purposes, one restricts the set of strategies to self-financing strategies, i.e. strategies such that the change in value of the portfolio is given by

$$\mathrm{d}V_t = r\alpha_t \mathrm{d}t + \frac{\phi_t}{S_t}\mathrm{d}S_t \ . \tag{2.5}$$

In other words, the change in value of the portfolio is only due to a change in value of the assets. We then compute, using (2.4), that

$$V_t = V_0 + \int_0^t r V_s \mathrm{d}s + \int_0^t \phi_s \sigma_s \mathrm{d}W_s.$$

We observe that the value of V only depends on ϕ and V_0. We also need to impose some technical conditions on ϕ and we will assume that $\mathbb{E}\Big[\int_0^t |\phi_s|^2 \mathrm{d}s\Big] < \infty$ so that the stochastic integral is a martingale.

The super-replication problem is to find a strategy that will hedge the terminal payoff with the minimal initial cost

$$p := \inf \mathscr{G}_0 \quad \text{with} \quad \mathscr{G}_0 = \{v \in \mathbb{R} | \exists \phi \in \mathscr{H}^2 \ , \ V_T^{v,\phi} \geq g(S_T)\} \,. \tag{2.6}$$

Proposition 2.2 *Consider $(\mathscr{Y}, \mathscr{Z})$, the solution to the following linear BSDE*

$$\mathscr{Y}_t = \xi - \int_t^T r\mathscr{Y}_s \mathrm{d}s - \int_t^T \mathscr{Z}_s dW_s \ . \tag{2.7}$$

Then the super-replication price is a replication price and is given by $p := \mathscr{Y}_0$. The replication strategy is $\phi^* = \frac{\mathscr{Z}}{\sigma}$.

Proof 1. The existence and uniqueness of the solution to (2.7) has already been discussed above. It is straightforwardly seen that $\mathscr{Y}_t = V_t^{\mathscr{Y}_0,\phi^*}$ for all $t \in [0, T]$ and then that $V_T^{\mathscr{Y}_0,\phi^*} = \xi$. This proves that $\mathscr{Y}_0 \geq p$ and that $\mathscr{G}_0$ is non-empty.

2. Now, let $v \in \mathscr{G}_0$ and $\phi \in \mathscr{H}^2$ such that $V_T^{v,\phi} \geq \xi$. A simple application of Itô's Formula shows that $(e^{-rt} V_t^{v,\phi})_{t \in [0,T]}$ is a martingale. In particular, we have that

$$v = \mathbb{E}\Big[e^{-rT} V_T^{v,\phi}\Big] \geq \mathbb{E}\big[e^{-rT}\xi\big] = \mathscr{Y}_0,$$

the last equality coming from (2.3). This yields that for all $v \in \mathscr{G}_0$, $v \geq \mathscr{Y}_0$. The proof is then concluded by taking the infimum on $\mathscr{G}_0$. □

Remark 2.2 The price at any date $t \in [0, T]$ is given by $\mathscr{Y}_t = \mathbb{E}\big[e^{r(T-t)}\xi \mid \mathscr{F}_t\big]$.

2.1.2.2 A Non-linear Market

We now consider a case of market imperfection: We work with two different rates for borrowing (R) and lending (r) with $R > r$.

We want to price a European contingent claim in this market following a hedging strategy. The main difference now is that the cash dynamics is given by

$$\begin{aligned} \mathrm{d}\alpha_t &= r\alpha_t \mathbf{1}_{\{\alpha_t \geq 0\}} \mathrm{d}t + R\alpha_t \mathbf{1}_{\{\alpha_t < 0\}} \mathrm{d}t \\ &= (r\alpha_t + (r - R)[\alpha_t]_-)\mathrm{d}t \,. \end{aligned} \tag{2.8}$$

The self-financing condition (2.5) rewrites in this context as

$$\mathrm{d}V_t = (r\alpha_t + (r - R)[\alpha_t]_-)\mathrm{d}t + \phi_t\sigma_t\mathrm{d}W_t.$$

Recalling that $V_t = \alpha_t + \phi_t$, we then compute

$$\mathrm{d}V_t = \{rV_t + (r - R)[V_t - \phi_t]_-\}\,\mathrm{d}t + \phi_t\sigma_t\mathrm{d}W_t.$$

The super hedging problem is still given by (2.6), only the dynamics of V has changed.

Theorem 2.3 *Let (Y, Z) be the solution to the following non-linear BSDE*

$$Y_s = \xi - \int_s^T \left\{rY_t + (r - R)[Y_t - \frac{Z_t}{\sigma_t}]_-\right\}\mathrm{d}t - \int_s^T Z_t\mathrm{d}W_t\ . \tag{2.9}$$

Then the super-replication price is a replication price and is given by $p := Y_0$. The replication strategy is $\phi^* = \frac{Z}{\sigma}$.

Proof 1. From the dynamics of V, we observe that $\mathscr{Y}_t = V_t^{\mathscr{Y}_0,\phi^*}$ for all $t \in [0, T]$ and then that $V_T^{\mathscr{Y}_0,\phi^*} = \xi$. This proves that $Y_0 \geq p$ and that $\mathscr{G}_0$ is non-empty.
2. Now, let $v \in \mathscr{G}_0$ and $\phi \in \mathscr{H}^2$ such that $V_T^{v,\phi} \geq \xi$. We can interpret $(V_t^{v,\phi})_{t\in[0,T]}$ as a BSDE with terminal condition $V_T^{v,\phi}$ and the same driver as (2.9). Then a direct application of the Comparison Theorem 2.2 leads to $v = V_0^{v,\phi} \geq Y_0$. The proof is concluded by taking the infimum on $\mathscr{G}_0$. □

The link between mathematical finance and BSDE theory is very fruitful, see e.g. [25]. For example, recently, Crepey [16, 17] has studied counterparty risk in the framework of BSDEs.

2.1.3 Applications to Stochastic Control

We illustrate in this section how BSDEs can be used to solve stochastic control problems. We present here a direct approach and we refer to Sect. 2.3.1 for a variational approach related to the stochastic maximum principle.

Let us consider an $\mathbb{R}$-valued process X which is given under $\mathbb{P}$ as the solution of the following differential equation:

$$dX_t = \sigma(X_t)dW_t\ , \tag{2.10}$$

where $\sigma : \mathbb{R} \to \mathbb{R}$ is a Lipschitz function satisfying $\Lambda \geq \sigma(x) \geq \frac{1}{\Lambda}$ for some $\Lambda > 0$ and W is a one-dimensional standard Brownian Motion.

We denote by $\mathscr{U}$ the set of progressively measurable processes α with values in a compact interval U. For a given $\alpha \in \mathscr{U} \subset \mathbb{R}$, we define

$$W_t^\alpha = W_t - \int_0^t \sigma^{-1}(X_t) b(X_t, \alpha_t) \mathrm{d}t ,$$

where $b : \mathbb{R}^2 \to \mathbb{R}$ is a Lipschitz function with bounded support. Applying Girsanov's Theorem, we thus have that W^α is a Brownian Motion under a new probability $\mathbb{P}^\alpha$ (which is absolutely continuous with respect to $\mathbb{P}$). The dynamics of X under $\mathbb{P}^\alpha$ reads as

$$dX_t = b(X_t, \alpha_t)dt + \sigma(X_t) dW_t^\alpha . \tag{2.11}$$

The control problem is classically given by the following optimisation

$$\min_{\alpha \in \mathscr{U}} J(\alpha) \quad \text{with} \quad J(\alpha) := \mathbb{E}^{\mathbb{P}^\alpha}\left[g(X_T) + \int_0^T h(X_s, \alpha_s) \mathrm{d}s \right] , \tag{2.12}$$

where $g : \mathbb{R} \to \mathbb{R}$ and $h : \mathbb{R}^2 \to \mathbb{R}$ are Lipschitz functions.

Let us assume that

$$a^*(x, z) := \mathrm{argmin}_{a \in U} H(x, z, a) \quad \text{where} \quad H(x, z, a) := h(x, a) + z\sigma^{-1}(x) b(x, a) \tag{2.13}$$

is well defined as a Lipschitz continuous function of (x, z). Let us then introduce

$$H^*(x, z) := h(x, a^*(x, z)) + b(x, a^*(x, z))z . \tag{2.14}$$

The function H above is called the Hamiltonian of the system, and H^* is its optimal value. We then have the following result.

Theorem 2.4 *In the above setting, the control problem (2.12) has a solution Y_0^* given by the initial value of the following BSDE*

$$Y_t^* = g(X_T) + \int_t^T H^*(X_s, Z_s^*) ds - \int_t^T Z_s^* dW_s , \tag{2.15}$$

and an optimal control is $\alpha_t^* = a^*(X_t, Z_t^*)$, $t \in [0, T]$.

Proof In this restrictive setting, H^* is Lipschitz continuous and (2.15) has a unique solution. For $\alpha \in \mathscr{U}$, we consider the solution (Y^α, Z^α) of the following BSDE

$$Y_t^\alpha = g(X_T) + \int_t^T h(X_s, \alpha_s) ds - \int_t^T Z_s^\alpha dW_s^\alpha , \tag{2.16}$$

observing that $Y_0^\alpha = J(\alpha)$. Then, rewriting the above dynamics under $\mathbb{P}$, we obtain

$$Y_t^\alpha = g(X_T) + \int_t^T H(X_s, Z_s^\alpha, \alpha_s)ds - \int_t^T Z_s^\alpha dW_s\,. \tag{2.17}$$

We now use the Comparison Theorem 2.2, recalling the definition of H^* in (2.14), to obtain that

$$Y_0^\alpha \geq Y_0^* = Y^{\alpha^*}\,,$$

which concludes the proof. □

This—by now classical—approach to solving stochastic control problems using BSDEs can be extended to various settings, in particular for non-zero sum games [28], where the existence of the representing BSDE is quite difficult to obtain.

2.1.4 Extensions

The theory of BSDEs is rich and powerful. It has attracted a lot of interest in the past 25 years. In this section, we report briefly on some extensions to the Lipschitz setting and the basic shape of Eq. (2.1). Note that we still present the case of Brownian filtrations but, of course, BSDEs have been studied in relation to jump processes as well, see e.g. [2]. Nor are we going to delve into the study of second-order BSDEs [48]. As already remarked, BSDE theory is still an active field of research and we do not aim to be exhaustive in the list we give below.

2.1.4.1 Constrained BSDEs

For modelling purposes, the processes Y and Z sometimes need to be constrained to belong to some possibly random sets. Generally, Eq. (2.1) no longer holds and one has to add a finite variation process as part of the solution, which then reads as

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)\mathrm{d}s - \int_t^T Z_s \mathrm{d}W_s + \int_t^T dK_s\,, \quad 0 \leq t \leq T. \tag{2.18}$$

The solution is now a triple (Y, Z, K) and, obviously, some other conditions are needed to guarantee uniqueness, depending on the applications.

Reflected BSDEs (RBSDEs)

In the one-dimensional setting, RBSDEs are linked, in their simplest form, to optimal stopping problems and the pricing of American options in non-linear markets [24]. If the exercise price of the option is given by a process $(L_t)_{0\leq t\leq T}$, then $(Y_t)_{0\leq t\leq T}$, representing the option price, has to satisfy $Y_t \geq L_t$ for all $0 \leq t \leq T$. Thus, Y is forced to belong to the random set $[L, \infty)$. The process K, which forces Y above S, is a continuous increasing process in this setting. Uniqueness for K is obtained thanks to the condition $\int_0^T (Y_t - L_t) dK_t = 0$: This simply states that K is active only when Y touches the boundary L.

It is also possible to add an upper obstacle U for Y leading to doubly reflected BSDEs [18]. In this case, Y is forced to belong to $[L, U]$ and the main applications are Dynkin games and the pricing of Game options, which are callable American Options, e.g. convertible bonds.

More generally, in the multi-dimensional setting, Y can be constrained to a closed convex domain $\mathscr{D}$, possibly random. The question is then the direction of reflection at the boundary of the convex domain. The case of normal reflection is treated in full generality in [27]. The case of oblique reflection is more involved, see [14] and the references therein for an account of RBSDEs linked to optimal switching problems.

BSDEs with Constraints on Z

BSDEs with constraints on the Z-process have been introduced in [19]. The minimal solution to (2.18) is found such that $Z \in \mathscr{D}$. In this case, BSDEs are linked to the pricing of European Options when some investment constraints are present on the market.

2.1.4.2 The Non-lipschitz Setting

The Lipschitz setting has been extended in various ways, but then existence and uniqueness results are much more difficult to obtain, when available. The first extension concerns coefficients with the monotonic property in y only, see [20], where the case of random terminal time is also treated. Let us mention the application to stochastic homogenisation [40]. The notion of generalised BSDEs has been introduced in [45], where the driver involves integration with respect to a finite variation process. This allows us to represent solutions of non-linear PDEs with generalised Neumann boundary condition. See Sect. 2.2.2 for an account of the link between PDEs and FBSDEs.

An important generalisation for applications comes from the introduction of quadratic growth in the component Z. This has been introduced for utility maximisation problems in [30, 47] and recently to principal-agent problems [23]. The one-dimensional case (for Y) is now well understood, see e.g. [30, 32]. The article [49] is the first to give an existence and uniqueness result for multi-dimensional BSDEs with quadratic growth. The general case is known to be difficult [26] albeit with recent progress [29, 50].

BSDEs with only continuous coefficients but with linear growth are considered in a multi-dimensional setting in [28] with an important application to non-zero sum games.

2.1.4.3 McKean–Vlasov FBSDEs

Recently, BSDEs have been introduced to study large population stochastic control problems. These are control problems of the type (2.11) and (2.12) but involving many interacting agents. A classical example is the following. Consider N agents, whose personal state is given by (for player i)

$$dX_t^i = b(X_t^i, \mu_t^n, \alpha_t^i)dt + \sigma dW_t^i,$$

(W^i) are independent Brownian motions, $\mu_t^n = \frac{1}{n}\sum_i \delta_{X_t^i}$ is the statistical distribution of the system, and α^i the control of the player. The cost to minimise for each player (given here for player i) is

$$J^i(\alpha) = \mathbb{E}\left[g(X_T^i, \mu_T^n) + \int_0^T f(t, X_t^i, \mu_t^n, \alpha_t^i)\mathrm{d}t \right] .$$

Note that the players interact via μ^n only. Games with a large number of player are difficult to solve. The hope here is to obtain an asymptotic ($N \to \infty$) description of the equilibrium, hopefully "easier" to handle.

The notion of equilibrium is then fundamental as it yields to different limiting equations. Individualistic, i.e. Nash-like, equilibria were first considered by Lasry and Lions [35], who coined them Mean Field Games. They were introduced at the same time by Caines, Huang and Malhamé [31]. Cooperative equilibrium leads to the control of McKean–Vlasov SDEs [11]. For a comparison between the two approaches, we mention [12]. The probabilistic approach to studying such problems has been developed by Carmona and Delarue [6, 10, 11], see the references therein for early works, and leads to the study of the following system

$$\begin{cases} X_t = \xi + \int_0^t b(X_s, Y_s, Z_s, \mathbb{P}_{(X_s,Y_s,)})ds + \int_0^t \sigma(X_s, Y_s, \mathbb{P}_{(X_s,Y_s,)})dW_s, \\ Y_t = g(X_T, \mathbb{P}_{X_T}) + \int_t^T f(X_s, Y_s, Z_s, \mathbb{P}_{(X_s,Y_s,)})ds - \int_t^T Z_s dW_s. \end{cases}$$

The main peculiarity, besides the coupling between the two equations, is the fact that the coefficients depend upon the law of the solution. These McKean–Vlasov FBSDEs have been linked to some non-linear PDEs written on the Wassertein space [8, 13]. Let us mention finally that the very difficult question of the convergence of the controlled particle system to the mean-field limit has been studied recently in [5].

2.2 Markovian BSDEs

We present in this section an important class of BSDEs for practical applications, namely the class of Markovian BSDEs. The main reference is [42]. A key point of our study will be the numerical approximation of such BSDEs, which will be studied in detail in Chap. 4.

2.2.1 First Definition and Markov Property

In this section, we consider the situation where all the randomness in the coefficients of the BSDEs (driver and terminal condition) comes from the value of a forward SDEs. In this special setting, the solution (Y, Z) can be linked to the solution of a parabolic PDE, as discussed below in Sect. 2.2.2. We first introduce some definitions and describe the Markovian property of the BSDE in this framework. In particular, we study the solution (X, Y, Z) to the following system

$$\begin{cases} X_t &= X_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s, \\ Y_t &= g(X_T) + \int_t^T f(X_s, Y_s, Z_s)ds - \int_t^T Z_s dW_s, \end{cases} \tag{2.19}$$

where $X_0 \in \mathbb{R}^n$ and we assume that the coefficient functions satisfy
- (HL): $b : \mathbb{R}^n \to \mathbb{R}^n$, $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times d}$, $g : \mathbb{R}^n \to \mathbb{R}^k$, $f : \mathbb{R}^n \times \mathbb{R}^k \times \mathbb{R}^{k\times d} \to \mathbb{R}^d$ are L-Lipschitz continuous functions for some constant $L > 0$.

In the above system, X is called the forward component and (Y, Z) the backward component. The existence and uniqueness of a strong solution to the forward SDE satisfied by X is quite classical in this Lipschitz setting, see e.g. [33]. It satisfies, for all $p \geq 1$,

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X_t|^p\right] < \infty\ . \tag{2.20}$$

For the backward component, one can apply Theorem 2.1, observing that

$$\mathbb{E}\left[|g(X_T)|^2 + \int_0^T |f(X_s, 0, 0)|^2 ds\right] < +\infty\,,$$

from the integrability property of X given in (2.20).
This discussion leads us to the following result.

Corollary 2.1 *Under (HL), there exists a unique solution $(X, Y, Z) \in \mathscr{S}_c^2 \times \mathscr{S}_c^2 \times \mathscr{H}^2$ to Eq. (2.19).*

At this stage, it is important to notice that the solution of the backward component does not appear in the coefficients of the forward SDE: This system is generally coined a "decoupled" FBSDE. We will see in Sect. 2.3 a more involved setting.

In order to describe the Markovian property of (Y, Z) rigorously, we introduce the flow of the above system, i.e. we consider (2.19) with $(t, x) \in [0, T] \times \mathbb{R}^n$ as the initial condition instead of $(0, X^0)$. We thus define

$$\begin{cases} X_s^{t,x} = x + \int_t^s b(X_r^{t,x})dr + \int_t^s \sigma(X_r^{t,x})dW_r, \\ Y_s^{t,x} = g(X_T^{t,x}) + \int_s^T f(X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x})dr - \int_s^T Z_r^{t,x} dW_r. \end{cases} \tag{2.21}$$

For $t \leq u$, we denote by $\mathscr{F}_u^t$ the augmentation of the filtration $\sigma\big(W_r - W_t,\ t \leq r \leq u\big)$ and we observe that $\big\{\big(X_u^{t,x}, Y_u^{t,x}, \big)\big\}_{t \leq u \leq T}$ is adapted to the filtration $\big\{\mathscr{F}_u^t\big\}_{t \leq u \leq T}$. We can work with a version of $\big\{Z_u^{t,x}\big\}_{t \leq u \leq T}$ adapted to $\big\{\mathscr{F}_u^t\big\}_{t \leq u \leq T}$.

In particular, $Y_t^{t,x}$ is deterministic and this leads naturally to the following definition.

Definition 2.1 For all $(t, x) \in [0, T] \times \mathbb{R}^n$, we set $u(t, x) := Y_t^{t,x}$.

Let us quote the following basic but key property for u [42].

Proposition 2.3 *Under (HL), the following holds, for all x, y in $\mathbb{R}^n$, t, t' in* $[0, T]$,

$$|u(t, x) - u(t', y)| \leq \Lambda(|x - y| + (1 + |x|)\sqrt{|t - t'|}) ,$$

for some positive constant Λ depending only on L and T.

The main result of this section is the following. We skip its proof as we prove below a similar result in a more complicated setting, see Proposition 2.9.

Proposition 2.4 *Let $t \in [0, T]$ and θ be a square integrable random variable independent of W. Then, $\mathbb{P}$ a.s.,*

$$\forall s \in [t, T], \qquad Y_s^{t,\theta} = u\left(s, X_s^{t,\theta}\right).$$

In particular, for the system (2.19), we have $Y_t = u(t, X_t)$, $t \in [0, T]$. Note that in the case $f = 0$, the previous equality expresses simply the Markov property of the process X.

2.2.2 The Link with PDEs

In this section, we characterise u as the (unique) solution of a parabolic PDE. This result is often referred to as a non-linear Feynman–Kac formula.

Let us consider the following PDE:

$$\begin{cases} \partial_t v(t,x) + \mathscr{L} v(t,x) + f\,(x, v(t,x), \partial_x v(t,x)\sigma(t,x)) = 0, & (t,x) \in [0,T[\times\mathbb{R}^n, \\ v(T,\cdot) = g, \end{cases} \tag{2.22}$$

where $\mathscr{L}$ is the Dynkin operator associated to the diffusion X, namely,

$$\mathscr{L}\varphi(x) = \partial_x\varphi(x)b(x) + \frac{1}{2}\mathrm{Tr}\left[\partial^2_{xx}\varphi\sigma\sigma^\dagger\right](x)\ ,$$

for φ a $\mathscr{C}^2$-function.

Proposition 2.5 *Assume that the PDE (2.22) has a solution v, belonging to* $\mathscr{C}^{1,2}\big([0,T]\times\mathbb{R}^n\big)$*, and such that* $(t,x)\in[0,T]\times\mathbb{R}^n$*,* $|\nabla v(t,x)| \le C\,(1+|x|^q)$*.*

Then for all $(t,x)\in[0,T]\times\mathbb{R}^n$*, the solution of the BSDE (2.21),* $\big\{Y_r^{t,x}, Z_r^{t,x}\big\}_{t\le r\le T}$*, is given by* $\big\{v\big(r, X_r^{t,x}\big), \partial_x v\sigma\big(r, X_r^{t,x}\big)\big\}_{t\le r\le T}$*. In particular, we have* $u(t,x) = Y_t^{t,x} = v(t,x)$*.*

Proof This follows directly from the application of Itô's formula to $t \mapsto v(t, X_t)$. □

There is a reciprocal result which is slightly more difficult to obtain, as we do not a priori have a better regularity property for u than the one given in Proposition 2.3. One can assume more regularity conditions to prove that u is a classical solution of (2.22), see e.g. [42, Theorem 3.2]. In the Lipschitz setting, one has to use a weaker notion of solution. When Y is one-dimensional, a natural notion to consider is that of a viscosity solution, see e.g. [15].

Definition 2.2 Let v be a continuous function on $[0,T]\times\mathbb{R}^n$ satisfying $v(T,x) = g(x)$. The function v is a sub-solution (resp. super-solution) in the viscosity sense to the PDE (2.22) if, for all functions φ belonging $\mathscr{C}^{1,2}\big([0,T]\times\mathbb{R}^n\big)$, we have,

$$-\partial_t\varphi(t,x) - \mathscr{L}\varphi(t,x) - f\,(t,x,u(t,x),\partial_x\varphi\sigma(t,x)) \le 0, \qquad (\text{resp. } \ge 0),$$

for all points $(t,x)\in]0,T[\times\mathbb{R}^n$ at which $v-\varphi$ has a local maximum (resp. local minimum): The function v is a viscosity solution if it is a viscosity sub- and super-solution.

For a presentation of viscosity solution theory, we refer to the seminal paper [15].

Remark 2.3 In the above definition, one can assume that the maximum (or minimum) is 0, simply by replacing φ by $\varphi - \varphi(t,x) + u(t,x)$.

Theorem 2.5 *The function* $u(t,x) := Y_t^{t,x}$ *is a viscosity solution of the PDE (2.3).*

Proof We already know that u is continuous and satisfies $u(T,\cdot) = g$. We show that u is a sub-solution (the proof of the super-solution property follows from similar arguments).

Let φ in $\mathscr{C}^{1,2}$ such that $u-\varphi$ has at (t_0,x_0) a local maximum $(0<t_0<T)$ equal to 0. We want to show that

$$\partial_t\varphi(t_0,x_0)+\mathscr{L}\varphi(t_0,x_0)+f\big(x_0,u(t_0,x_0),\partial_x\varphi\sigma(t_0,x_0)\big)\geq 0.$$

We will assume that there exists a $\delta>0$ such that

$$\partial_t\varphi(t_0,x_0)+\mathscr{L}\varphi(t_0,x_0)+f\big(x_0,u(t_0,x_0),\partial_x\varphi\sigma(t_0,x_0)\big)=-\delta<0$$

and work toward a contradiction. The function $u-\varphi$ has a local maximum at (t_0,x_0) equal to 0. By continuity there exists $0<\alpha\leq T-t_0$ such that $t_0\leq t\leq t_0+\alpha$ and $|x-x_0|\leq\alpha$,

$$u(t,x)\leq\varphi(t,x)\quad\text{and}\quad\partial_t\varphi(t,x)+\mathscr{L}\varphi(t,x)+f\big(t,x,u(t,x),\partial_x\varphi\sigma(t,x)\big)\leq-\delta/2.$$

Let us consider the stopping time $\tau=\inf\big\{u\geq t_0\,;\,\big|X_u^{t_0,x_0}-x_0\big|>\alpha\big\}\wedge t_0+\alpha$. Since X^{t_0,x_0} is a continuous process, we have $\big|X_\tau^{t_0,x_0}-x_0\big|\leq\alpha$.
We apply Itô's formula to $\varphi(r,X_r^{t_0,x_0})$ between $u\wedge\tau$ and $(t_0+\alpha)\wedge\tau=\tau$ and obtain for $t_0\leq u\leq t_0+\alpha$,

$$\begin{aligned}\varphi\big(u\wedge\tau,X_{u\wedge\tau}^{t_0,x_0}\big)=\varphi\big(\tau,X_\tau^{t_0,x_0}\big)&-\int_{u\wedge\tau}^{\tau}\{\partial_t\varphi+\mathscr{L}\varphi\}\big(r,X_r^{t_0,x_0}\big)\,dr\\&-\int_{u\wedge\tau}^{\tau}\partial_x\varphi\sigma\big(r,X_r^{t_0,x_0}\big)\,dW_r\,;\end{aligned}$$

defining for $t_0\leq u\leq t_0+\alpha$, $Y'_u=\varphi\big(u\wedge\tau,X_{u\wedge\tau}^{t_0,x_0}\big)$ and $Z'_u=\mathbf{1}_{\{u\leq\tau\}}\partial_x\varphi\sigma\big(r,X_r^{t_0,x_0}\big)$, the previous equality reads, for $t_0\leq u\leq t_0+\alpha$,

$$Y'_u=\varphi\big(\tau,X_\tau^{t_0,x_0}\big)+\int_u^{t_0+\alpha}-\mathbf{1}_{\{r\leq\tau\}}\big\{\varphi'+\mathscr{L}\varphi\big\}\big(r,X_r^{t_0,x_0}\big)\,dr-\int_u^{t_0+\alpha}Z'_r\,dW_r.$$

Similarly, we set for $t_0\leq u\leq t_0+\alpha$, $Y_u=Y_{u\wedge\tau}^{t_0,x_0}$ and $Z_u=\mathbf{1}_{\{u\leq\tau\}}Z_u^{t_0,x_0}$, then

$$Y_u=Y_{t_0+\alpha}+\int_u^{t_0+\alpha}\mathbf{1}_{\{r\leq\tau\}}\,f\big(r,X_r^{t_0,x_0},Y_r,Z_r\big)\,dr-\int_u^{t_0+\alpha}Z_r\,dW_r,\qquad t_0\leq u\leq t_0+\alpha.$$

The Markov property, see Proposition 2.4, implies that $\mathbb{P}$ a.s. for all $t_0\leq r\leq t_0+\alpha$, $Y_r^{t_0,x_0}=u(r,X_r^{t_0,x_0})$ and thus $Y_{t_0+\alpha}=Y_\tau^{t_0,x_0}=u(\tau,X_\tau^{t_0,x_0})$. The previous equality reads as

$$Y_u=u\big(\tau,X_\tau^{t_0,x_0}\big)+\int_u^{t_0+\alpha}\mathbf{1}_{r\leq\tau}\,f\big(r,X_r^{t_0,x_0},u\big(r,X_r^{t_0,x_0}\big),Z_r\big)\,dr-\int_u^{t_0+\alpha}Z_r\,dW_r,$$

for $t_0 \le u \le t_0 + \alpha$. We now apply the Comparison Theorem, see Theorem 2.2, to $\left(Y'_u, Z'_u\right)_u$ and $(Y_u, Z_u)_u$. From the definition of τ, we have $u\left(\tau, X_\tau^{t_0,x_0}\right) \le \varphi\left(\tau, X_\tau^{t_0,x_0}\right)$ and

$$\begin{aligned}\mathbf{1}_{\{r\le\tau\}} f\left(r, X_r^{t_0,x_0}, u(r, X_r^{t_0,x_0}), Z'_r\right) &= \mathbf{1}_{r\le\tau} f\left(r, X_r^{t_0,x_0}, u(r, X_r^{t_0,x_0}), \partial_x\varphi\sigma(r, X_r^{t_0,x_0})\right) \\ &\le -\mathbf{1}_{\{r\le\tau\}} \left\{\varphi' + \mathscr{L}\varphi\right\} \left(r, X_r^{t_0,x_0}\right).\end{aligned}$$

Moreover,

$$\begin{aligned}\mathbb{E}\left[\int_{t_0}^{t_0+\alpha} -\mathbf{1}_{r\le\tau} \left(\varphi' + \mathscr{L}\varphi + f\right)\left(r, X_r^{t_0,x_0}, u(r, X_r^{t_0,x_0}), \partial_x\varphi\sigma(r, X_r^{t_0,x_0})\right) dr\right] \\ \ge \\ \mathbb{E}[\tau - t_0]\,\delta/2 > 0.\end{aligned}$$

Indeed, $\delta > 0$ and $\tau > t_0$ since $|X_{t_0}^{t_0,x_0} - x_0| = 0 < \alpha$. We thus apply the strict version of the Comparison Theorem, see Theorem 2.2, to obtain $u(t_0, x_0) = Y_{t_0} < Y'_{t_0} = \varphi(t_0, x_0)$. This is absurd since $u(t_0, x_0) = \varphi(t_0, x_0)$. □

It turns out that u is the unique viscosity solution (in the class of functions with polynomial growth) of (2.3), see [39].

In the Markovian setting, the solution to the BSDE (2.19) appears as "stochastic characteristics" for the parabolic PDE (2.22). This representation of the solution u allows one to design probabilistic numerical methods for the approximation of the PDE. This point will be discussed in Chap. 4.

2.3 Coupled Forward-Backward SDEs

For the purpose of our application to carbon emissions markets, we need to study a more complicated form of Forward-Backward SDEs, namely:

$$\begin{cases} X_t &= X_0 + \int_0^t b(X_s, Y_s, Z_s)ds + \int_0^t \sigma(X_s, Y_s)dW_s, \\ Y_t &= g(X_T) + \int_t^T f(X_s, Y_s, Z_s)ds - \int_t^T Z_s dW_s. \end{cases} \tag{2.23}$$

Compared to (2.19), we observe that the solution of the backward component appears in the coefficients of the forward component, which renders the question of existence and uniqueness much more intricate.

This section will present some results from the theory of fully coupled BSDEs (2.23). Classical references on the subject are [1, 21, 36, 44] and [38], see the references therein as well.

We first present, as a motivation, an application of (2.23) to a stochastic control problem of the type presented in Sect. 2.1.3.

2.3.1 *The Pontryagin Approach to Stochastic Control Problems*

We present the stochastic maximum principle for a convex optimisation problem. It contains as a special case linear-quadratic optimisation problems, see [3, 4] and [38].

In the one-dimensional setting, we consider the following control process

$$X_t^\alpha = x + \int_0^t (aX_s^\alpha + b\alpha_s)ds + W_t \ , \tag{2.24}$$

where the control $\alpha \in \mathscr{H}^2$. The cost functional is given by

$$J(\alpha) = \mathbb{E}\Big[\int_0^T [h(X_s^\alpha) + |\alpha_s|^2]ds + g(X_T^\alpha)\Big] \tag{2.25}$$

where $g : \mathbb{R} \to \mathbb{R}$ and $h : \mathbb{R} \to \mathbb{R}$ are $\mathscr{C}^2$ convex functions with bounded second derivatives. The optimisation problem is then

$$\inf_{\alpha \in \mathscr{H}^2} J(\alpha) \ .$$

Remark 2.4 We observe that $\alpha \mapsto J(\alpha)$ is strictly convex and $J(\alpha) \to +\infty$ when $|\alpha| \to +\infty$. From convex analysis results, one can show that there is a unique solution to this optimisation problem.

In order to study the solvability of the problem, we compute the Gâteaux derivative of the functional $\alpha \mapsto J(\alpha)$.

Lemma 2.1 *Let* $\alpha, v \in \mathscr{H}^2$. *Then*

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon}(J(\alpha + \varepsilon v) - J(\alpha)) = \mathbb{E}\Big[\int_0^T [h'(X_s^\alpha)\mathscr{X}_s + \alpha_s v_s]ds + g'(X_T^\alpha)\mathscr{X}_T\Big],$$

where $\mathscr{X}_t = \int_0^t (a\mathscr{X}_s + bv_s)ds$ *and* g' *(resp.* h'*) stands for the first derivative of* g *(resp.* h*).*

Proof We compute

$$\begin{aligned}\frac{1}{\varepsilon}(J(\alpha + \varepsilon v) - J(u)) =& \frac{1}{\varepsilon}\mathbb{E}\Big[\int_0^T \big([h(X_s^{\alpha+\varepsilon v}) - h(X_s^\alpha)] + \{(\alpha_s + \varepsilon v_s)^2 - \alpha_s^2\}\big)\, ds\Big] \\ &+ \frac{1}{\varepsilon}\mathbb{E}\big[g(X_T^{\alpha+\varepsilon v}) - g(X_T^\alpha)\big] \ .\end{aligned}$$

We observe that

$$\frac{1}{\varepsilon}\mathbb{E}\big[|X_T^{\alpha+\varepsilon v}|^2 - |X_T^\alpha|^2\big] = \mathbb{E}\left[\delta_\varepsilon X \frac{X_T^{\alpha+\varepsilon v} + X_T^\alpha}{2}\right],$$

with $\delta_\varepsilon X := \frac{1}{\varepsilon}(X^{\alpha+\varepsilon v} - X^\alpha)$. Identifying the dynamics, we have that $\delta_\varepsilon X = \mathscr{X}$. We then compute

$$\begin{aligned}\mathscr{E}(\varepsilon) &:= \frac{1}{\varepsilon}\mathbb{E}\big[g(X_T^{\alpha+\varepsilon v}) - g(X_T^\alpha) - \mathscr{X}_T g'(X_T^\alpha)\big] \\ &= \mathbb{E}\left[\mathscr{X}_T \int_0^{\mathbf{1}} [g'(X_t^\alpha + \varepsilon\lambda\mathscr{X}_T) - g'(X_T^\alpha)]d\lambda\right] \le C\varepsilon\mathbb{E}\big[|\mathscr{X}_T|^2\big].\end{aligned}$$

We conclude that $\lim_{\varepsilon\downarrow 0}\mathscr{E}(\varepsilon) = 0$. □

We reformulate the result of the previous lemma in a more efficient way by introducing the adjoint equation

$$\mathrm{d}Y_t = -[aY_t + h'(X_t^\alpha)]\mathrm{d}t + Z_t\mathrm{d}W_t \text{ and } Y_T = g'(X_T^\alpha),$$

which is a linear BSDE.

Lemma 2.2 *Let $\alpha, v \in \mathscr{H}^2$. Then,*

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}(J(\alpha+\varepsilon v) - J(\alpha)) = \mathbb{E}\left[\int_0^T \{bY_s + \alpha_s\}v_s\mathrm{d}s\right].$$

Proof Applying the product rule, we compute

$$\mathbb{E}\big[g'(X_T^\alpha)\mathscr{X}_T\big] = \mathbb{E}[Y_T\mathscr{X}_T] = \mathbb{E}\left[\int_0^t [bY_s v_s - \mathscr{X}_s h'(X_s^\alpha)]\mathrm{d}s\right].$$

□

Proposition 2.6 (Necessary condition) *If α is an optimal control, then*

$$\mathbb{E}\left[\int_0^T (bY_s + \alpha_s)v_s\mathrm{d}s\right] = 0, \quad \forall v \in \mathscr{H}^2.$$

It is then given by $\alpha_s = -bY_s$, where Y satisfies the following optimality system

$$\begin{cases} \mathrm{d}Y_t = -[aY_t + h'(X_t)]\mathrm{d}t + Z_t\mathrm{d}W_t \textit{ and } Y_T = g'(X_T), \\ \mathrm{d}X_t = (aX_t - b^2Y_t)\mathrm{d}t + \mathrm{d}W_t \textit{ and } X_0 = x. \end{cases} \tag{2.26}$$

Proof This is a direct application of Lemma 2.2, noticing that $J(\alpha+\varepsilon v) - J(\alpha) \ge 0$ if u is an optimum. □

Proposition 2.7 *Let $\hat{Y}$ be a solution to (2.26), we set $\hat{\alpha} = -b\hat{Y}$. Then $\hat{\alpha}$ is a solution to the minimisation problem.*

Proof By convexity of $x \mapsto g(x)$, for any control α, we have

$$\mathbb{E}\Big[g(X_T^\alpha) - g(X_T^{\hat{\alpha}})\Big] \geq \mathbb{E}\Big[g'(X_T^{\hat{\alpha}})(X_T^u - X_T^{\hat{\alpha}})\Big] = \mathbb{E}\Big[\hat{Y}_T(X_T^\alpha - X_T^{\hat{\alpha}})\Big].$$

Applying the product rule, we compute

$$\mathbb{E}\Big[\hat{Y}_T(X_T^\alpha - X_T^{\hat{\alpha}})\Big] = \mathbb{E}\Big[\int_0^T \{b\hat{Y}_s(\alpha_s - \hat{\alpha}_s) - h'(X_s^{\hat{\alpha}})(X_s^\alpha - X_s^{\hat{\alpha}})\}\mathrm{d}s\Big].$$

We also observe that

$$\begin{gathered}\mathbb{E}\Big[\int_0^T \{(h(X_s^\alpha) + |\alpha_s|^2) - (h(X_s^{\hat{\alpha}}) + |\hat{\alpha}_s|^2)\}\mathrm{d}s\Big] \\ \geq \\ \mathbb{E}\Big[\int_0^T \{h'(X_s^{\hat{\alpha}})(X_s^\alpha - X_s^{\hat{\alpha}}) + (\alpha_s - \hat{\alpha}_s)\hat{\alpha}_s\}\mathrm{d}s\Big].\end{gathered}$$

Summing the two previous inequalities, we get

$$J(\alpha) - J(\hat{\alpha}) \geq 0,$$

which concludes the proof by arbitrariness of α. □

Remark 2.5 From Remark 2.4, one obtains existence and uniqueness of solutions to (2.26), a fully coupled FBSDE.

2.3.2 Well-Posedness of FBSDEs in Small Time Duration

We start with an example of non-solvable FBSDEs, which is a special case of Proposition 3.1, Chap. 1 in [38].

Example 2.1 We consider the following system of FBSDEs

$$\begin{cases} \mathrm{d}Y_t = -X_t\mathrm{d}t + Z_t\mathrm{d}W_t \quad \text{and} \quad Y_T = -X_T, \\ \mathrm{d}X_t = Y_t\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t \quad \text{and} \quad X_0 = x, \end{cases} \tag{2.27}$$

where $T = \frac{3\pi}{4}$ and σ is a Lipschitz function.

If $x \neq 0$, there is no solution to the above equation in $\mathscr{S}^2 \times \mathscr{S}^2 \times \mathscr{H}^2$.

Proof Indeed, assume there is one. Then, we observe that $x(t) := \mathbb{E}[X_t]$ and $y(t) := \mathbb{E}[Y_t]$ satisfy the system of ODEs:

$$\begin{cases} dy(t) = -x(t)dt \quad \text{and} \quad y(T) = -x(T)\,, \\ dx(t) = y(t)dt \quad \text{and} \quad x(0) = x\,. \end{cases} \tag{2.28}$$

The proof then follows from direct computations: first we have that $x(t) = x\cos(t) + \mu\sin(t)$, for $\mu \in \mathbb{R}$, and in particular $x(T) = -x\sqrt{2}/2 + \mu\sqrt{2}/2$. Also, we get $y(t) = -x\sin(t) + \mu\cos(t)$ and $y(T) = -x\sqrt{2}/2 - \mu\sqrt{2}/2$. From this we observe that $y(T) + x(T) = -x\sqrt{2}$, which allows us to conclude the proof. □

The difficulty encountered above can be overcome by assuming a small coupling between the backward and forward equation, see [44], or similarly working with a small terminal time T, as described in the next section.

2.3.2.1 Existence and Uniqueness

For this section, we simply assume that the functions b, f, σ, g are L-Lipschitz continuous. The most restrictive assumption comes from the fact that we can no longer consider arbitrary terminal time.

Theorem 2.6 *For all $T \le T^* := \gamma(L)$, there exists a unique solution to the following Forward-Backward SDE:*

$$\begin{cases} X_t = \theta + \int_0^t b(X_s, Y_s, Z_s)\mathrm{d}s + \int_0^t \sigma(X_s, Y_s)\mathrm{d}W_s, \\ Y_t = g(X_T) + \int_t^T f(X_s, Y_s, Z_s)\mathrm{d}s - \int_t^T Z_s\mathrm{d}W_s, \end{cases} \tag{2.29}$$

where θ is a square integrable random variable independent of W.

Proof For the reader's convenience, for a process $\Theta \in \mathscr{S}_c^2 \times \mathscr{S}_c^2 \times \mathscr{H}^2$ we define

$$\|\Theta\|_2 = \|\Theta^1\|_{\mathscr{S}^2} + \|\Theta^2\|_{\mathscr{S}^2} + \|\Theta^3\|_{\mathscr{H}^2}\,.$$

We consider the function $\Phi : \mathscr{S}^2 \to \mathscr{S}^2$ mapping a process Γ to the first component X of (X, Y, Z), the solution to

$$X_t = \xi + \int_0^t b(X_s, Y_s, Z_s)\mathrm{d}s + \int_0^t \sigma(X_s, Y_s)\mathrm{d}W_s, \tag{2.30}$$

$$Y_t = g(\Gamma_T) + \int_t^T f(\Gamma_s, Y_s, Z_s)\mathrm{d}s - \int_0^t Z_s\mathrm{d}W_s, \tag{2.31}$$

which is well defined, by using Corollary 2.1.

We now show that for T small enough, Φ is a contraction.
1. Consider two processes Γ, Γ' and $\Theta = (X, Y, Z)$, $\Theta' = (X', Y', Z')$ the associated solution to (2.30)–(2.31). We define $\delta\Gamma = \Gamma - \Gamma'$, $\delta X = X - X'$, $\delta Y = Y - Y'$, $\delta Z = Z - Z'$, $\delta\Theta = \Theta - \Theta'$ and $\delta b = b(\Theta) - b(\Theta')$, $\delta\sigma = \sigma(X, Y) - \sigma(X', Y')$, $\delta f = f(\Gamma, Y, Z) - f(\Gamma', Y', Z')$, $\delta g = g(\Gamma) - g(\Gamma')$.
Applying Itôs formula to $t \mapsto |\delta X_t|^2$, we get

$$|\delta X_t|^2 = 2\int_0^t \delta X_s \delta b_s \mathrm{d}s + 2\int_0^t \delta X_s \delta\sigma_s \mathrm{d}W_s + \int_0^t |\delta\sigma_s|^2 \mathrm{d}s. \tag{2.32}$$

We compute

$$\mathbb{E}\Big[\int_0^T |\delta X_s \delta b_s| \mathrm{d}s\Big] \le C_L \mathbb{E}\Big[\int_0^T (|\delta X_s|^2 + |\delta Y_s|^2)\mathrm{d}s + \int_0^T \delta X_s \delta Z_s \mathrm{d}s\Big] \tag{2.33}$$

and using Young's inequality,

$$\mathbb{E}\Big[\int_0^T \delta X_s \delta Z_s \mathrm{d}s\Big] \le \frac{1}{\sqrt{T}}\mathbb{E}\Big[\int_0^T |\delta X_s|^2 \mathrm{d}s\Big] + \sqrt{T}\mathbb{E}\Big[\int_0^T |\delta Z_s|^2 \mathrm{d}s\Big].$$

Inserting the previous inequality into (2.33), we obtain

$$\mathbb{E}\Big[\int_0^T |\delta X_s \delta b_s| \mathrm{d}s\Big] \le C_L\sqrt{T}\left(\mathbb{E}\Big[\sup_s |\delta X_s|^2 + \int_0^T |\delta Z_s|^2 \mathrm{d}s\Big] + \sup_s \mathbb{E}\big[|\delta Y_s|^2\big]\right). \tag{2.34}$$

We also easily get

$$\mathbb{E}\Big[\int_0^T |\delta\sigma_s|^2 \mathrm{d}s\Big] \le C_L T\left(\mathbb{E}\Big[\sup_s |\delta X_s|^2\Big] + \sup_s \mathbb{E}\big[|\delta Y_s|^2\big]\right). \tag{2.35}$$

Now we compute, using the Burkholder–Davis–Gundy inequality,

$$\begin{aligned}\mathbb{E}\Big[\sup_t \Big|\int_0^t \delta X_s \delta\sigma_s \mathrm{d}W_s\Big|\Big] &\le C\mathbb{E}\Big[\Big|\int_0^t |\delta X_s \delta\sigma_s|^2 \mathrm{d}s\Big|^{\frac{1}{2}}\Big] \\ &\le \frac{1}{2}\mathbb{E}\Big[\sup_s |\delta X_s|^2\Big] + C_L\mathbb{E}\Big[\int_0^T |\delta Y_s|^2 \mathrm{d}s\Big].\end{aligned} \tag{2.36}$$

Combining (2.34)–(2.36) with (2.32), we get

$$\mathbb{E}\Big[\sup_s |\delta X_s|^2\Big] \le \frac{C_L\sqrt{T}}{\frac{1}{2} - C_L\sqrt{T}}\left(\sup_s \mathbb{E}\big[|\delta Y_s|^2\big] + \mathbb{E}\Big[\int_0^T |\delta Z_s|^2 \mathrm{d}s\Big]\right), \tag{2.37}$$

for T small enough.

2. Applying Itô's formula to $t \mapsto |\delta Y_t|^2$, we get

$$|\delta Y_t|^2 + \int_t^T |\delta Z_s|^2 \mathrm{d}s = |\delta g_T|^2 + 2\int_t^T \delta Y_s \delta f_s \mathrm{d}s + 2\int_t^T \delta Y_s \delta Z_s \mathrm{d}W_s .$$

Applying the Burkholder–Davis–Gundy inequality, we compute

$$\begin{aligned}
\mathbb{E}\Big[\sup_{t\in[0,T]} |\int_0^t \delta Y_s \delta Z_s dW_s|\Big] &\le C\mathbb{E}\Big[|\int_0^t (\delta Y_s \delta Z_s)^2 ds|^{\frac{1}{2}}\Big] \\
&\le C\mathbb{E}\Big[\sup_{s\in[0,t]} |\delta Y_s| \left(\int_0^t |\delta Z_s|^2 ds\right)^{\frac{1}{2}}\Big] \\
&\le C\left(\mathbb{E}\Big[\sup_{s\in[0,t]} |\delta Y_s|^2\Big] + \mathbb{E}\Big[\int_0^t |\delta Z_s|^2 ds\Big]\right) < \infty,
\end{aligned}$$

where we use the Cauchy–Schwarz inequality for the last step. Observing, moreover, that

$$2\delta Y_s \delta f_s \le C_L(|\delta Y_s|^2 + |\delta \Gamma_s|^2) + \frac{1}{2}|\delta Z_s|^2 ,$$

we get

$$\sup_t \mathbb{E}\big[|\delta Y_t|^2\big] + \frac{1}{2}\mathbb{E}\Big[\int_0^T |\delta Z_s|^2 \mathrm{d}s\Big] \le C_L \mathbb{E}\Big[\sup_s |\delta \Gamma_s|^2\Big] + C_L T \sup_s \mathbb{E}\big[|\delta Y_s|^2\big] .$$

This leads to

$$\sup_s \mathbb{E}\Big[|\delta Y_s|^2\Big] \le \frac{C_L}{1 - C_L T}\mathbb{E}\Big[\sup_s |\delta \Gamma_s|^2\Big] \ \text{ and } \ \mathbb{E}\Big[\int_0^T |\delta Z_s|^2 \mathrm{d}s\Big] \le C_L \mathbb{E}\Big[\sup_s |\delta \Gamma_s|^2\Big] , \tag{2.38}$$

for T small enough.
The proof is concluded by inserting (2.38) into (2.37). □

2.3.2.2 The Decoupling Field and a Quasilinear PDE

As in the decoupled case, we consider for all $(t, x) \in [0, T] \times \mathbb{R}^d$, the system

$$\begin{cases} X_s^{t,x} = x + \int_t^s b(X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x})dr + \int_t^s \sigma(X_r^{t,x}, Y_r^{t,x})dW_r, \\ Y_s^{t,x} = g(X_T^{t,x}) + \int_s^T f(X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x})dr - \int_s^T Z_r^{t,x} dW_r, \end{cases} \tag{2.39}$$

which has a unique solution, from Theorem 2.6, as soon as $T \le T^*$. We observe that $Y_t^{t,x}$ is a deterministic quantity which leads us naturally to the following definition

Definition 2.3 (*decoupling field*) For $(t, x) \in [0, T] \times \mathbb{R}^n$, we set $u(t, x) := Y_t^{t,x}$.

Using similar arguments as in the proof of Theorem 2.6, one proves the following properties for the decoupling field.

Proposition 2.8 (basic properties) *For $T \leq T^*$ and all $(t, t', x, x') \in [0, T]^2 \times \mathbb{R}^{2n}$,*

$$|u(t, x)| \leq C(1 + |x|) \text{ and } |u(t, x) - u(t', x')| \leq C\left(|x - x'| + (1 + |x|)|t - t'|^{\frac{1}{2}}\right).$$

We now prove the key property of the decoupling field, which explicitly describes Y as a function of X: $Y_s^{t,x} = u(s, X_s^{t,x})$, $t \leq s \leq T$. We will see later that, whenever u is smooth, $Z_s^{t,x} = \partial_x u\sigma(X_s^{t,x}, u(s, X_s^{t,x}))$. Thus, the same property as in the case of Markovian BSDEs presented in Sect. 2.2 holds true in this coupled setting. Note that the function u could then be used in the coefficient of $X^{t,x}$ instead of (Y, Z): In this sense, it decouples the backward and forward equations.

Proposition 2.9 *We have that $\mathbb{P}$ a.s. for all $s \leq T$*

$$Y_s^{t,\theta} = u(s, X_s^{t,\theta}),$$

where θ is a square integrable random variable independent of W.

Proof The proof follows the proof of Corollary 1.5 in [21].

1. We first show that $u(t, \theta) = Y_t^{t,\theta}$. Indeed, we compute, for $\varepsilon > 0$,

$$\begin{aligned}\mathbb{E}\big[\mathbf{1}_{|\theta - x| < \varepsilon}|u(t, \theta) - Y_t^{t,\theta}|^2\big] &\leq 2\mathbb{E}\big[\mathbf{1}_{|\theta - x| < \varepsilon}(|\theta - x|^2 + |Y_t^{t,x} - Y_t^{t,\theta}|^2)\big] \\ &\leq C_L\mathbb{E}\big[\mathbf{1}_{|\theta - x| < \varepsilon}|\theta - x|^2\big]\end{aligned}$$

from the above results.

We then have

$$\begin{aligned}\mathbb{E}\big[|u(t, \theta) - Y_t^{t,\theta}|^2\big] &\leq C_L\mathbb{E}\left[\sum_{k \in \mathbb{Z}} \mathbf{1}_{\{|\theta - k/N| < \frac{1}{N}\}}|\theta - k/N|^2\right] \\ &\leq C_L/N^2,\end{aligned}$$

which concludes the proof of this step.

2. Consider the following FBSDE

$$\begin{cases} X_v = X_s^{t,\theta} + \int_s^v b(\Theta_r)\mathrm{d}r + \int_t^v \sigma(X_r, Y_r)\mathrm{d}s, \\ Y_v = g(X_T) + \int_v^T f(\Theta_r)\mathrm{d}r - \int_v^T Z_r\mathrm{d}W_s, \end{cases} \tag{2.40}$$

for $t \leq s \leq T$. We conclude the proof by observing that $\Theta^{t,\theta}$ is a solution on $[s, T]$ to the above FBSDEs, leading to $u(s, X_s^{t,\theta}) = Y_s^{t,\theta}$, by the previous step.

□

If the function u is smooth enough—which can be proven if the coefficients are themselves smooth enough [21]—then it is a solution to the quasilinear PDE

$$\begin{cases} \partial_t v + b(\cdot, v, \partial_x v\sigma(\cdot, v))\partial_x v + \frac{1}{2}\mathrm{Tr}[\partial^2_{xx} v\sigma(\cdot, v)\sigma(\cdot, v)^\dagger] + f(\cdot, v, \partial_x v\sigma) = 0\,, \\ v(T, \cdot) = g(\cdot)\,. \end{cases} \tag{2.41}$$

In the following proposition, we show the converse result.

Proposition 2.10 *Assume that v is a classical bounded solution to (2.41) with bounded first and second derivatives. Then, for $t \le s \le T$, $Y^{t,\theta}_s = v(s, X^{s,\xi}_s)$ and $Z_s = \partial_x v(s, X^{s,\theta}_s)\sigma(X^{s,\theta}_s, v(s, X^{s,\theta}_s))$, where $\Theta^{t,\theta} := (X^{t,\theta}, Y^{t,\theta}, Z^{t,\theta})$ is a solution to*

$$\begin{cases} X^{t,\theta}_s = \theta + \int_t^s b(\Theta^{t,\theta}_s)\mathrm{d}s + \int_t^s \sigma(X^{t,\theta}_s, Y^{t,\theta}_s)\mathrm{d}W_s\,, \\ Y^{t,\theta}_s = g(X^{t,\theta}_T) + \int_t^T f(\Theta^{t,\theta}_s)\mathrm{d}s - \int_t^T Z^{t,\theta}_s \mathrm{d}W_s\,. \end{cases} \tag{2.42}$$

If, moreover, σ is bounded, then uniqueness holds for the FBSDE.

Proof 1. Existence: We first consider $X^{t,\theta}$ to be the solution of

$$X^{t,\theta}_s = \xi + \int_t^s b(X^{t,\theta}_r, v(r, X^{t,\theta}_r), \partial_x v(r, X^{t,\theta}_r)\sigma(X^{t,\theta}_r, v(r, X^{t,\theta}_r)))\mathrm{d}r$$
$$+ \int_t^s \sigma(X^{t,\theta}_r, v(r, X^{t,\theta}_r))\mathrm{d}W_r\,.$$

By the property of the function v, it has a unique solution. To conclude the proof, one only has to apply Ito's formula to $s \mapsto v(s, X^{t,\theta}_s)$ and identify the term.

2. Uniqueness when σ is bounded. Consider any other solution (X, Y, Z) to (2.42). Define $\bar{Y}_s = v(s, X_s)$ and $\bar{Z}_s = \partial_x v(s, X_s)\sigma(X_s, \bar{Y}_s)$. Applying Itô's formula, we obtain

$$\bar{Y}_t = g(X_T) - \int_t^T (\partial_t v(s, X_s) + b(X_s, Y_s, Z_s)\partial_x v(s, X_s)$$
$$+ \frac{1}{2}\sigma(X_s, Y_s)^2 \partial^2_{xx} v(s, X_s))\mathrm{d}s - \int_t^T \tilde{Z}_s \mathrm{d}W_s\,,$$

with $\tilde{Z}_s = \partial_x v(s, X_s)\sigma(X_s, Y_s)$. This leads to

$$\bar{Y}_t = g(X_T) + \int_t^T (f(X_s, \bar{Y}_s, \tilde{Z}_s) + R_s)\mathrm{d}s - \int_t^T \tilde{Z}_s \mathrm{d}W_s\,,$$

with

$$
\begin{aligned}
R_s &= (f(X_s, \bar{Y}_s, \bar{Z}_s) - f(X_s, \bar{Y}_s, \tilde{Z}_s)) + \{b(X_s, \bar{Y}_s, \bar{Z}_s) - b(X_s, Y_s, Z_s)\}\partial_x v(s, X_s) \\
&\quad + \frac{1}{2}\{\sigma(X_s, \bar{Y}_s)^2 - \sigma(X_s, Y_s)^2\}\partial^2_{xx} v(s, X_s)).
\end{aligned}
$$

We first compute, using the property of σ, v, f, that

$$
\begin{aligned}
|R_s| &\le C(|\bar{Y}_s - Y_s| + |\bar{Z}_s - \tilde{Z}_s| + |\tilde{Z}_s - Z_s|) \\
&\le C(|\bar{Y}_s - Y_s| + |\tilde{Z}_s - Z_s|).
\end{aligned}
$$

For the last inequality we observe that

$$
|\partial_x v(s, X_s)\sigma(X_s, \bar{Y}_s) - \partial_x v(s, X_s)\sigma(X_s, Y_s)| \le |\bar{Y}_s - Y_s| .
$$

We then apply Itô's formula to $t \mapsto |\delta Y_t|^2$ with $\delta Y = \bar{Y} - Y$ and use classical arguments to obtain that $Y = \bar{Y}$ and $Z = \tilde{Z} = \bar{Z}$.

Then X and $X^{t,\theta}$ satisfy the same SDE, concluding the proof of this step. □

2.3.3 Existence and Uniqueness for Arbitrary Terminal Time

The question of existence and uniqueness of solutions to coupled Forward-Backward SDEs for an arbitrary time duration is difficult. In the Lipschitz setting, one has to assume some structural conditions on the coefficients to hope to prove such results. There are various methods that can be used to obtain them and they are complementary. The main ones are the *four step scheme* [36], the *method of continuation* [46, 51] and more recently a method based on *characteristic BSDEs* [37]. The last two allow us to work in non-Markovian settings. It is now well understood that a key object used to obtain the existence and uniqueness of solutions to FBSDEs is the *decoupling field* u and in particular the possibility of proving a uniform Lipschitz property in its spatial variable.

As clearly stated in the following chapters, in this book we need to work with Markovian FBSDEs, so we will focus on this class of coupled equations only. In this case, one can profit from the link between the decoupling field and the quasilinear PDE introduced above. The main method then to obtain existence and uniqueness is related to the *four step scheme*. In the sequel, we present a variation on this method, detailed in [21], which is certainly the main result in the non-degenerate diffusion coefficient case. We conclude the chapter with the presentation of an extension to the case of degenerate diffusion, which turns out to be the most important one for our application to carbon emission pricing, see Chap. 3.

2.3.3.1 Non-degenerate Diffusion Coefficient

In this section, we present the existence and uniqueness result for FBSDEs obtained by Delarue in [21, Theorem 2.6] for an arbitrary terminal time.

As already pointed out in Example 2.1, one needs to assume further conditions on the coefficient to obtain such result. The first key assumption is the non-degeneracy condition on the diffusion coefficient σ:

$$\forall \upsilon \in \mathbb{R}^d , \quad \upsilon^\dagger \sigma \sigma^\dagger (x, y) \upsilon \geq \lambda |\upsilon|^2 , \tag{2.43}$$

for some $\lambda > 0$. We also impose a boundedness assumption,

$$\begin{cases} |f(x, y, z)| + |b(x, y, z)| + |\sigma(x, y)| \leq \Lambda(1 + |y| + |z|) , \\ |g(x)| \leq \Lambda , \end{cases} \tag{2.44}$$

typically ruling out Example 2.1. The main result reads as follows.

Theorem 2.7 *Under (HL) and conditions (2.43)–(2.44), there exists a unique solution to (2.23) for all terminal times T.*

The proof of this theorem is quite involved, see [21]. Let us simply describe the strategy to obtain such a result. A natural scheme to go from a small time to an arbitrary time duration is the following:

- One first works on the interval $[T - \delta_1, T]$, where $\delta_1 \leq \gamma(L)$ is given by Theorem 2.6 and from which follows the existence and uniqueness of a solution (X, Y, Z) and of a decoupling field u.
- One then works on the interval $[T - (\delta_1 + \delta_2), T - \delta_1]$ for some δ_2 to be determined. The terminal condition is no longer given by g, but by $u(T - \delta_1, \cdot)$, whose Lipschitz constant is denoted by C_L^1 (and it is not L any more!). We can apply the result of the previous section and then we get that δ_2 must satisfy $\delta_2 \leq \gamma(C_L^1)$.
- We can obtain by induction a sequence of (δ_n) and consider existence and uniqueness on $[T - \sum_{k=1}^n \delta_k, T]$, etc. The problem is that the δ_n may become smaller and smaller, not reaching 0 at the limit, i.e. failing to obtain a solution on $[0, T]$.

The key to making the above scheme work is then to control the Lipschitz constant of u, uniformly in time. This is done using gradient estimates on the quasilinear PDE satisfied by u in a smooth setting. This difficult estimate can be obtained by an analytical argument, see [34], or by a probabilistic approach, see [22]. The Lipschitz case is then obtained by a regularisation argument.

2.3.3.2 FBSDEs with Singular Coefficients

In this final section, we would like to address the existence and uniqueness of solutions to FBSDEs (2.23) when the coefficient σ is possibly degenerate and the terminal condition is non-Lipschitz. The need for such an extension comes from the application to carbon emissions markets, as mentioned in [9]. However, it is outside the scope of this book to give a complete account of this difficult problem. To fix our ideas, we consider the following system

$$\begin{cases} dX_t^1 &= b^1(X_t^1)dt + \sigma^1(X_t^1)dW_t\,, \\ dX_t^2 &= b^2(X_t^1, Y_t)dt\,, \\ dY_t &= Z_t dW_t\,, \end{cases} \tag{2.45}$$

where $X = (X^1, X^2)$ is the forward component with a degenerate diffusion coefficient and the backward part is linear, f being set to 0. This system is studied in detail in [7] and is related to the one used in Chap. 3. The terminal condition g is assumed here to be a non-decreasing bounded function.

With this model, to obtain existence and uniqueness for an arbitrary terminal time, some specific conditions on the drift coefficient b^2 are required. One needs to impose that b^2 is strictly increasing and satisfies, for ℓ_1, ℓ_2 two positive constants,

$$\ell_1|y-y'|^2 \le (y-y')[b^2(x,y) - b^2(x,y')] \le \ell_2|y-y'|^2\,, \quad y, y' \in \mathbb{R}^2\,. \tag{2.46}$$

As explained in [7], the behavior of the system at the terminal condition is quite peculiar. This comes from both the degeneracy of the diffusion coefficient and the irregularity of the terminal condition. In particular, if $\sigma^1(x) \ge \varepsilon > 0$, the law of X_T^2 exhibits Dirac masses. Moreover, the authors prove that the Markovian structure is lost at time T (Y_T is no longer a function of X_T^2). This means we should relax the terminal condition requirement. The general result reads as follows.

Theorem 2.8 (Theorem 2.2 in [7]) *Let* (HL) *and condition* (2.46) *hold. Then there exists a unique process* $(X, Y, Z) \in \mathscr{S}_c^2 \times \mathscr{S}_c^2 \times \mathscr{H}^2$ *satisfying* (2.45) *with the terminal condition*

$$\mathbb{P}\{g_-(X_T^2) \le Y_T \le g_+(X_T^2)\} = 1\,,$$

where g_- *(resp.* g_+*) is the left continuous (resp. right continuous) version of* g*.*

Remark 2.6 One way to prove this theorem is to first "regularise" the solution of the above system by adding some noise in the component X^2 and by smoothing the terminal condition. In this context, the existence and uniqueness will follow from the results of the previous section. Then, delicate arguments [7] have to be used to retrieve a solution of (2.46) by letting the noise go to zero and by removing the smoothing on the terminal condition.

References

1. Antonelli, Fabio. 1993. Backward-forward stochastic differential equations. *The Annals of Applied Probability*, 777–793.
2. Barles, Guy, Rainer Buckdahn, and Etienne Pardoux. 1997. Backward stochastic differential equations and integral-partial differential equations. *Stochastics: An International Journal of Probability and Stochastic Processes* 60 (1–2): 57–83.
3. Bismut, Jean-Michel. 1973. Conjugate convex functions in optimal stochastic control. *Journal of Mathematical Analysis and Applications* 44 (2): 384–404.
4. Bismut, Jean-Michel. 1978. Contrôle des systèmes linéaires quadratiques: applications de l'intégrale stochastique. In *Séminaire de Probabilités XII*, 180–264. Springer.
5. Cardaliaguet, Pierre, Delarue, François, Lasry, Jean-Michel, Lions, Pierre-Louis. 2015. The master equation and the convergence problem in mean field games. arXiv:1509.02505.
6. Carmona, René, and François Delarue. 2013. Probabilistic analysis of mean-field games. *SIAM Journal on Control and Optimization* 51 (4): 2705–2734.
7. Carmona, René, and François Delarue. 2013. Singular FBSDEs and scalar conservation laws driven by diffusion processes. *Probability Theory and Related Fields* 157 (1–2): 333–388.
8. Carmona, René, Delarue, Francois. 2014. The master equation for large population equilibriums. In *Stochastic Analysis and Applications 2014*, 77–128. Springer.
9. Carmona, René, François Delarue, Gilles-Edouard Espinosa, and Nizar Touzi. 2013. Singular forward-backward stochastic differential equations and emissions derivatives. *The Annals of Applied Probability* 23 (3): 1086–1128.
10. Carmona, René, and François Delarue. 2013. Mean field forward-backward stochastic differential equations. *Electronic Communications in Probability* 18 (68): 15.
11. Carmona, René, and François Delarue. 2015. Forward-backward stochastic differential equations and controlled McKean-Vlasov dynamics. *The Annals of Probability* 43 (5): 2647–2700.
12. Carmona, René, François Delarue, and Aimé Lachapelle. 2013. Control of McKean-Vlasov dynamics versus mean field games. *Mathematics and Financial Economics* 7 (2): 131–166.
13. Chassagneux, Jean-François, Crisan, Dan, Delarue, François. 2014. A probabilistic approach to classical solutions of the master equation for large population equilibria. arXiv:1411.3009.
14. Chassagneux, Jean-François, Romuald Elie, and Idris Kharroubi. 2011. A note on existence and uniqueness for solutions of multidimensional reflected BSDEs. *Electronic Communications in Probability* 16: 120–128.
15. Crandall, Michael G., Hitoshi Ishii, and Pierre-Louis Lions. 1992. User's guide to viscosity solutions of second order partial differential equations. *Bulletin of the American Mathematical Society* 27 (1): 1–67.
16. Crépey, Stéphane. 2015. Bilateral counterparty risk under funding constraints-part I: Pricing. *Mathematical Finance* 25 (1): 1–22.
17. Crépey, Stéphane. 2015. Bilateral counterparty risk under funding constraints-part II: CVA. *Mathematical Finance* 25 (1): 23–50.
18. Cvitanic, Jaksa, Karatzas, Ioannis. 1996. Backward stochastic differential equations with reflection and Dynkin games. *The Annals of Probability*, 2024–2056.
19. Cvitanic, Jaksa, Karatzas, Ioannis, Mete Soner, H. 1998. Backward stochastic differential equations with constraints on the gains-process. *Annals of Probability*, 1522–1551.
20. Darling, Richward W.R., and Etienne Pardoux. 1997. Backwards SDE with random terminal time and applications to semilinear elliptic PDE. *The Annals of Probability* 25 (3): 1135–1159.
21. Delarue, François. 2002. On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case. *Stochastic Processes and their Applications* 99 (2): 209–286.
22. Delarue, François. 2003. Estimates of the solutions of a system of quasi-linear PDEs. A probabilistic scheme. In *Séminaire de Probabilités XXXVII*, 290–332. Springer.
23. Elie, Romuald, Possamaï, Dylan. 2016. Contracting theory with competitive interacting agents. arXiv:1605.08099.

24. El Karoui, Nicole, Kapoudjian, Christophe. 1997. Etienne Pardoux, Shige Peng, and Marie-Claire Quenez. Reflected solutions of backward SDE's, and related obstacle problems for PDE's. *The Annals of Probability*, 702–737.
25. El Karoui, Nicole, Shige Peng, and Marie-Claire Quenez. 1997. Backward stochastic differential equations in finance. *Mathematical Finance* 7 (1): 1–71.
26. Frei, Christoph, and Gonçalo Dos Reis. 2011. A financial market with interacting investors: does an equilibrium exist? *Mathematics and Financial Economics* 4 (3): 161–182.
27. Gegout Petit, Anne, and E. Pardoux. 1996. Equations différentielles stochastiques rétrogrades réfléchies dans un convexe. *Stochastics: An International Journal of Probability and Stochastic Processes* 57 (1–2): 111–128.
28. Hamadene, Said, Lepeltier, Jean-Pierre, Peng, Shige. 1997. BSDEs with continuous coefficients and stochastic differential games. *Pitman Research Notes in Mathematics Series*, 115–128.
29. Harter, Jonathan, Richou, Adrien. 2016. A stability approach for solving multidimensional quadratic BSDEs. arXiv:1606.08627.
30. Hu, Ying, Peter Imkeller, and Matthias Müller. 2005. Utility maximization in incomplete markets. *The Annals of Applied Probability* 15 (3): 1691–1712.
31. Huang, Minyi, Roland P. Malhamé, and Peter E. Caines. 2006. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Communications in Information & Systems* 6 (3): 221–252.
32. Kobylanski, Magdalena. 2000. Backward stochastic differential equations and partial differential equations with quadratic growth. *Annals of Probability*, 558–602.
33. Kunita, Hiroshi. 1997. *Stochastic flows and stochastic differential equations*, vol. 24. Cambridge University Press.
34. Ladyzhenskaya, O.A., Solonnikov, V.A., Ural'Ceva, N.N. 1968. *Linear and quasilinear equations of parabolic type, translations of mathematical monographs*, vol. 23. Providence RI: American Mathematical Society.
35. Lasry, Jean-Michel, and Pierre-Louis Lions. 2007. Mean field games. *Japanese Journal of Mathematics* 2 (1): 229–260.
36. Ma, Jin, Philip Protter, and Jiongmin Yong. 1994. Solving forward-backward stochastic differential equations explicitly - a four step scheme. *Probability Theory and Related Fields* 98 (3): 339–359.
37. Ma, Jin, Wu Zhen, Detao Zhang, and Jianfeng Zhang. 2015. On well-posedness of forward-backward SDEs-a unified approach. *The Annals of Applied Probability* 25 (4): 2168–2214.
38. Ma, Jin, Yong, Jiongmin. 1999. *Forward-backward stochastic differential equations and their applications*. Number 1702. Springer Science & Business Media.
39. Pardoux, Étienne. 1998. Backward stochastic differential equations and viscosity solutions of systems of semilinear parabolic and elliptic PDEs of second order. In *Stochastic Analysis and Related Topics VI*, 79–127. Springer.
40. Pardoux, Étienne. 1999. Homogenization of linear and semilinear second order parabolic PDEs with periodic coefficients: a probabilistic approach. *Journal of Functional Analysis* 167 (2): 498–520.
41. Pardoux, Etienne, and Shige Peng. 1990. Adapted solution of a backward stochastic differential equation. *Systems & Control Letters* 14 (1): 55–61.
42. Pardoux, Etienne, Peng, Shige. 1992. Backward stochastic differential equations and quasilinear parabolic partial differential equations. In *Stochastic partial differential equations and their applications*, 200–217. Springer.
43. Pardoux, Étienne, and Aurel Răşcanu. 2014. *Stochastic differential equations, Backward SDEs, Partial differential equations*, vol. 69. Springer.
44. Pardoux, Etienne, and Shanjian Tang. 1999. Forward-backward stochastic differential equations and quasilinear parabolic PDEs. *Probability Theory and Related Fields* 114 (2): 123–150.
45. Pardoux, Étienne, and Shuguang Zhang. 1998. Generalized BSDEs and nonlinear Neumann boundary value problems. *Probability Theory and Related Fields* 110 (4): 535–558.
46. Peng, Shige, and Wu Zhen. 1999. Fully coupled forward-backward stochastic differential equations and applications to optimal control. *SIAM Journal on Control and Optimization* 37 (3): 825–843.

47. Rouge, Richard, and Nicole El Karoui. 2000. Pricing via utility maximization and entropy. *Mathematical Finance* 10 (2): 259–276.
48. Soner, H.Mete, Nizar Touzi, and Jianfeng Zhang. 2012. Wellposedness of second order backward SDEs. *Probability Theory and Related Fields* 153 (1–2): 149–190.
49. Tevzadze, Revaz. 2008. Solvability of backward stochastic differential equations with quadratic growth. *Stochastic Processes and their Applications* 118 (3): 503–515.
50. Xing, Hao, Žitković, Gordan. 2016. A class of globally solvable Markovian quadratic BSDE systems and applications. arXiv:1603.00217.
51. Yong, Jiongmin. 1997. Finding adapted solutions of forward-backward stochastic differential equations: method of continuation. *Probability Theory and Related Fields* 107 (4): 537–572.

http://www.springer.com/978-3-319-63114-1

A Forward-Backward SDEs Approach to Pricing in Carbon Markets
Chassagneux, J.-F.; Chotai, H.; Muûls, M.
2017, VI, 104 p. 35 illus., 29 illus. in color., Softcover
ISBN: 978-3-319-63114-1