Mathematical Finance, Vol. 20, No. 2 (April 2010), 289–312

PRICING AND HEDGING OF DERIVATIVES BASED ON NONTRADABLE UNDERLYINGS

STEFAN ANKIRCHNER

Institute for Applied Mathematics, University of Bonn

PETER IMKELLER

Institut für Mathematik, Humboldt-Universität zu Berlin

GONÇALO DOS REIS

CMAP - Ecole Polytechnique

This paper is concerned with the study of insurance related derivatives on financial markets that are based on nontradable underlyings, but are correlated with tradable assets. We calculate exponential utility-based indifference prices, and corresponding derivative hedges. We use the fact that they can be represented in terms of solutions of forward-backward stochastic differential equations (FBSDE) with quadratic growth generators. We derive the Markov property of such FBSDE and generalize results on the differentiability relative to the initial value of their forward components. In this case the optimal hedge can be represented by the price gradient multiplied with the correlation coefficient. This way we obtain a generalization of the classical "delta hedge" in complete markets.

KEY WORDS: financial derivatives, hedging, utility-based pricing, BSDE, forward-backward stochastic differential equation (FBSDE), quadratic growth, differentiability, stochastic calculus of variations, Malliavin calculus, pricing by marginal utility.

1. INTRODUCTION

In recent years more and more financial instruments have been created which are not derived from exchange traded securities. For instance in 1999 the Chicago Mercantile Exchange introduced weather futures contracts, the payoffs of which are based on average temperatures at specified locations. Another example of derivatives with nontradable underlyings are catastrophe futures based on an insurance loss index regulated by an independent agency or simply derivatives based on equity indices such as S&P or DAX.

Financial or insurance derivatives of this type are impossible to perfectly hedge, since it is impossible to trade the underlying variable that carries independent uncertainty. To circumvent this problem, in practice one looks for a tradable asset that is correlated to the nontradable underlying of the derivative. Even though investing in the correlated asset cannot provide a total hedge of the derivative, and a nonhedgeable *basis risk* remains, it is better than not hedging at all.

In the following we will investigate utility-based pricing principles for derivatives based on nontradable underlyings. Moreover we will show how the derivatives can be partially

Manuscript received July 2007; final revision received May 2008.

Address correspondence to Peter Imkeller, Humboldt-Universität zu Berlin - Mathematics, Unter den Linden 6 Berlin 10099, Germany; e-mail: imkeller@math.hu-berlin.de.

hedged by investing in correlated assets. We present *explicit* hedging strategies that optimize the expected utility of a portfolio of such derivatives. To this end we will establish some structure and smoothness properties of indifference prices such as the Markov property and differentiability with respect to the underlyings. Once these properties are established, we can explicitly describe the optimal hedging strategies in terms of the price gradient and correlation coefficients. This way we obtain a generalization of the classical *delta hedge* of the Black–Scholes model.

The hedging of claims based on nontradable underlyings has already been studied by many authors, see for example, Henderson and Hobson (2002), Henderson (2002), Musiela and Zariphopoulou (2004), Davis (2006), Monoyios (2004), Ankirchner, Imkeller, and Popier (2008). As a common feature of all these papers, optimal hedging strategies are derived with standard stochastic control techniques. The essential components of this analytical approach consist in a formulation of the optimization problem in terms of Hamilton–Jacobi–Bellman (HJB) partial differential equations, and the use of a verification theorem and uniqueness result in order to obtain a representation of the indifference price and the optimal control strategy. We instead employ an approach with a stochastic focus. It starts with the well-known observation that the maximal expected *exponential* utility may be computed by appealing to the martingale optimality principle which leads to a description of price and optimal hedging strategy in terms of a *forward-backward stochastic differential equation* (FBSDE) with a nonlinearity of quadratic type (see Rouge and El Karoui 2000; Hu, Imkeller, and Müller 2005). This immediately implies that the utility indifference price, respectively, hedge is equal to the difference of initial states, respectively, control processes of two FBSDE with a quadratic nonlinearity in the generator. The forward component is given by a Markov process describing the nontradable underlying. The main mathematical contribution of this paper is that it provides simple sufficient conditions for general FBSDE with quadratic nonlinearity to satisfy a Markov property, and—for the BSDE component—to be differentiable with respect to the initial condition of the forward equation. The techniques for proving differentiability of BSDE with quadratic nonlinearity have been developed independently in Briand and Confortola (2008) and Ankirchner, Imkeller, and Dos Reis (2007). Unfortunately, the setup of both papers is not general enough to cover the BSDE needed to calculate exponential indifference prices. Therefore, a slight generalization of these differentiability results is given in the last section of this paper.

As a consequence of the explicit description of indifference prices and hedges in terms of the solution processes of the FBSDE, and in view of the smoothness results mentioned, it is straightforward to describe optimal hedging strategies in terms of the indifference price gradient and the correlation coefficients explicitly. An economics related contribution of the paper is that the framework presented allows to refine the results obtained, for example in Musiela and Zariphopoulou (2004), Davis (2006). First, no longer do we need to impose any restrictions on the coefficients of the diffusion modeling the tradable asset price. More importantly, the BSDE techniques allow to deal with multidimensional underlyings and traded assets. In the approach based on the HJB equation, a solution of the PDE is obtained by using an exponential Hopf-Cole transformation that in general seems to require that there exists only one traded asset. In practice many derivatives are based on more than one underlying, such as spread options or basket options. In order to illustrate how to hedge with more than one asset, we will study in more detail so-called crack spreads, which are written for instance on the difference of crude oil futures and kerosene prices (see Example 2.2 and 5.9).

Finally we address the pricing of derivatives by the *marginal utility approach*. If a company wishes to trade risk not covered by securities on an exchange, they are forced to go outside the exchange to get tailored products to serve their specific needs. These deals that do not go through the exchange trading (although the underlyings may be traded there) and are done directly between buyer and seller are called *over-the-counter* (OTC). For example, airlines regularly make this kind of OTC deals in order to protect themselves against kerosene price fluctuations, which underlines that the amount of money involved in this type of deals is nonnegligible! Investment banks offering OTC deals face the problem of finding a fair price of these agreements. Indifference prices are often a reasonable solution. However, they are not linear! The standard way out, as suggested in the economics literature, is pricing by marginal utility. The marginal utility price is the differential quotient of the indifference price with respect to a marginal amount of the derivative. Here again the first thing to verify is the differentiability of the FBSDE. This in turn allows to derive the dynamics of the marginal utility price as a BSDE with a driver satisfying a random Lipschitz condition.

BSDE with generators of quadratic nonlinearity in the control variable (which will in the sequel sometimes simply be called *quadratic BSDE*) are described by equations of the type

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)\, ds - \int_t^T Z_s\, dW_s, \quad 0 \leq t \leq T,$$

where f is a predictable function satisfying $|f(t, y, z)| \leq C(1 + |y| + |z|^2)$ with some constant C. Our differentiability results are based on the assumption that the derivative to be hedged, denoted by ξ, is essentially bounded. This guarantees that the integral process $\int_0^\cdot Z\, dW$ is a so-called BMO martingale, and hence the density process of a new equivalent probability measure, say Q. By switching to the measure Q one can derive moment estimates needed in order to prove differentiability. The assumption that the derivative has to be bounded seems to be a disadvantage of using BSDE in the stochastic approach instead of working with the HJB partial differential equation in the analytical approach. In practice, this is of no importance.

The paper is organized as follows: in Section 2 we introduce the model, in Section 3 we briefly recall results from Hu et al. (2005) concerning the solution of the problem of exponential expected utility maximization in terms of stochastic control problems and FBSDE with nonlinearities of quadratic type. In Section 4 we show structure properties of indifference prices of derivatives based on a nontradable Markovian index process. In Section 5 we derive explicit formulas for the optimal hedges of such derivatives, and in Section 6 we describe the dynamics of the marginal utility price. All the economics related results are based on mathematical properties of quadratic FBSDE, which will be proved in the last section. For further details, comments, and complete proofs we refer to Dos Reis (2010).

2. THE MODEL

Let $d \in \mathbb{N}$ and let W be a d-dimensional Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$. We denote by $(\mathcal{F}_t)$ the completion of the filtration generated by W. Suppose that a derivative with maturity $T > 0$ is based on a $\mathbb{R}^m$-dimensional nontradable index (think of a stock, temperature, or loss index) with dynamics

$$dR_t = b(t, R_t)\, dt + \rho(t, R_t)\, dW_t, \tag{2.1}$$

where $b : [0, T] \times \mathbb{R}^m \to \mathbb{R}^m$ and $\rho : [0, T] \times \mathbb{R}^m \to \mathbb{R}^{m\times d}$ are measurable deterministic functions. Throughout we assume that there exists a $C \in \mathbb{R}_+$ such that for all $t \in [0, T]$ and $x,\ x' \in \mathbb{R}^m$

$$\text{(R1)} \qquad \begin{aligned} |b(t,x) - b(t,x')| + |\rho(t,x) - \rho(t,x')| &\le C|x - x'|, \\ |b(t,x)| + |\rho(t,x)| &\le C(1 + |x|). \end{aligned}$$

We consider a derivative of the form $F(R_T)$, where $F : \mathbb{R}^m \to \mathbb{R}$ is a bounded and measurable function. Note that at time t, the expected payoff of $F(R_T)$, conditioned on $R_t = r$, is given by $F(R_T^{t,r})$, where $R^{t,r}$ is the solution of the SDE

$$R_s^{t,r} = r + \int_t^s b\big(u, R_u^{t,r}\big)\, du + \int_t^s \rho\big(u, R_u^{t,r}\big)\, dW_u, \quad s \in [t, T]. \tag{2.2}$$

Our correlated financial market consists of k risky assets and one nonrisky asset. We use the nonrisky asset as numeraire and suppose that the prices of the risky assets in units of the numeraire evolve according to the SDE

$$dS_t^i = S_t^i(\alpha_i(t, R_t)\, dt + \beta_i(t, R_t)\, dW_t), \quad i = 1, \ldots, k,$$

where $\alpha_i(t, r)$ is the ith component of a measurable and vector-valued map $\alpha : [0, T] \times \mathbb{R}^m \to \mathbb{R}^k$ and $\beta_i(t, r)$ is the i-th row of a measurable and matrix-valued map $\beta : [0, T] \times \mathbb{R}^m \to \mathbb{R}^{k\times d}$. Notice that W is the same $\mathbb{R}^d$-dimensional Brownian motion as the one driving the index process (2.1), and hence the correlation between the index and the tradable assets is determined by the matrices ρ and β.

In order to exclude arbitrage opportunities in the financial market we assume $d \ge k$. For technical reasons we suppose that

(M1) α is bounded,
(M2) there exist constants $0 < \varepsilon < K$ such that $\varepsilon I_k \le (\beta(t, r)\beta^*(t, r)) \le K I_k$ for all $(t, r) \in [0, T] \times \mathbb{R}^m$,

where $\beta^*(t, r)$ is the transpose of $\beta(t, r)$, and I_k is the k-dimensional unit matrix.

Before we proceed with the model description we will illustrate the range of possible applications by giving some examples of derivatives our model may apply to.

EXAMPLE 2.1. Weather derivatives are a typical example of financial instruments derived from nontradable underlyings. One of the most common types of weather derivatives are based on the so-called *accumulated heating degree days* (cHDD). The heating degree of a day with average temperature τ in Celsius degrees is defined as HDD $= \max\{0, 18 - \tau\}$, that is, HDD describes the (positive) difference between the average daily temperature measured and the temperature above which rooms are usually heated. The cHDDs are defined as a moving average sum of HDDs over a fixed time length, for instance a month. Real data shows the cHDD to be almost lognormally distributed, and therefore they can be modelled as geometric Brownian motions (see Davis 2001). This means that in (2.1) we would have to choose $b(t, R_t) = \alpha_1 R_t$ and $\rho(t, R_t) = \alpha_2 R_t$, with $\alpha_1 \in \mathbb{R}$ and $\alpha_2 \in \mathbb{R} \setminus \{0\}$ depending on the season. Tradable assets that are more or less correlated with average temperatures are for example, electricity futures and natural gas futures.

The derivative explained in the next example is based on more than one underlying.

EXAMPLE 2.2. Spread options in general involve two or more underlying structures (prices, indices, interest rates, and many other possible quantities), and measure the distance between them. We do not go into details since spread options are well known (see Carmona and Durrleman 2003 for an overview). For simplicity we refer to a two-dimensional example of *Crack spreads*.

Crack spreads consist in the simultaneous purchase or sale of crude against the sale or purchase of refined petroleum products. We concentrate on the kerosene crack spread, which pits crude oil price (co) against kerosene price (ke). A company producing kerosene (from crude oil) wishes to cover part of its risk arising from a sudden boost of the crude oil price by buying kerosene crack spreads. It thereby faces the problem that kerosene trading is not done on a sufficiently liquid market to warrant a futures contract or some other type of exchange-traded contract. So derivative contracts of this type must be arranged on OTC basis.

Knowing that the price of heating oil (ho) is highly correlated with the kerosene price—except during the Iraq war—crack spreads themselves can be hedged by using heating oil futures.

We model prices in the following way, where the superscripts represent the underlying products,

$$\begin{aligned} dR_t^{ke} &= R_t^{ke}\left(b_1 dt + \gamma_2 dW_t^1 + \gamma_3 dW_t^2 + \gamma_4 dW_t^3\right)\\ dR_t^{co} &= R_t^{co}\left(b_2 dt + \gamma_1 dW_t^1\right)\\ dS_t^{ho} &= S_t^{ho}\left(b_3 dt + \beta_1 dW_t^1 + \beta_2 dW_t^2\right), \end{aligned}$$

where we assume that $b_1, b_2, b_3 \in \mathbb{R}$, $\gamma_1, \gamma_2, \gamma_3, \gamma_4, \beta_1, \beta_2 \in \mathbb{R}\backslash\{0\}$ and the correlation between heating oil and kerosene is given by $\sigma = (\gamma_2\beta_1 + \gamma_3\beta_2)/\sqrt{(\gamma_2^2+\gamma_3^2+\gamma_4^2)(\beta_1^2+\beta_2^2)}$.

A European call on the spread is of the form $\xi(R_T^{ke}, S_T^{co}) = (R_T^{ke} - S_T^{co} - K)^+$, with K being the strike.

Throughout let U be the exponential utility function with risk aversion coefficient $\eta > 0$, that is

$$U(x) = -e^{-\eta x}.$$

In what follows let $(t, r) \in [0, T] \times \mathbb{R}^m$. By an *investment strategy* we mean any predictable process $\lambda = (\lambda^i)_{1 \le i \le k}$ with values in $\mathbb{R}^k$ such that the integral process $\int_0^t \lambda_r^i \frac{dS_r^i}{S_r^i}$ is defined for all $i \in \{1, \dots, k\}$. We interpret λ^i as the value of the portfolio fraction invested in the i-th asset. Investing according to a strategy λ leads to a total gain due to trading during the time interval $[t, s]$ which amounts to $G_s^{\lambda,t} = \sum_{i=1}^k \int_t^s \lambda_u^i \frac{dS_u^i}{S_u^i}$. We will denote by $G_s^{\lambda,t,r}$ the gain conditional on $R_t = r$.

REMARK 2.3. As one can see the wealth process is given by

$$G_s^{\lambda,t,r} = \sum_{i=1}^k \int_t^s \lambda_u^i\left[\alpha_i\left(u, R_u^{t,r}\right) du + \beta_i\left(u, R_u^{t,r}\right) dW_u\right)\right],$$

and hence does *not* depend on the value of the correlated price process! This is a feature of our model that will later imply the indifference price at time t to depend *only* on the value of the index process at a given time t.

Let $\mathcal{A}^{t,r}$ be the set of all strategies λ such that $E\int_t^T |\lambda_s \beta(s, R_s^{t,r})|^2 ds < \infty$ and the family $\{e^{-\eta G_\tau^{\lambda,t,r}} : \tau$ is a stopping time with values in $[t, T]\}$ is uniformly integrable. If $\lambda \in \mathcal{A}^{t,r}$, then we say that λ is *admissible*. We use the same admissibility criteria as in Section 3 in Hu et al. (2005), so that later we may invoke their results. The *maximal expected utility* at time T, conditioned on the wealth to be v at time t and the index to satisfy $R_t = r$, is defined by

$$V^0(t, v, r) = \sup\{EU(v + G_T^{\lambda,t,r}) : \lambda \in \mathcal{A}^{t,r}\}. \tag{2.3}$$

One can show that there exists a strategy π, called *optimal strategy*, such that $EU(v + G_T^{\pi,t,r}) = V^0(v, t, r)$. The convexity of the utility functions implies that π is a.s. unique on $[t, T]$, and it follows from Theorem 7 in Hu et al. (2005) that $\pi \in \mathcal{A}^{t,r}$.

Suppose an investor is endowed with a derivative $F(R_T)$ and is keeping it in his portfolio until maturity T. Then his maximal expected utility is given by

$$V^F(t, v, r) = \sup\{EU(v + G_T^{\lambda,t,r} + F(R_T^{t,r})) : \lambda \in \mathcal{A}^{t,r}\}. \tag{2.4}$$

Also in this case there exists an optimal strategy, denoted by $\widehat{\pi}$, that satisfies $EU(v + G_T^{\widehat{\pi},t,r} + F(R^{t,r})) = V^F(v, t, r)$.

The presence of the derivative $F(R_T)$ leads to a change in the optimal strategy from π to $\widehat{\pi}$. The difference

$$\Delta = \widehat{\pi} - \pi$$

is needed in order to hedge, at least partially, the risk associated with the derivative in the portfolio. We therefore call Δ *derivative hedge*. In the following sections we shall analyze by how much the optimal strategies change if a derivative is added to the portfolio, and we aim at getting an explicit expression for the derivative hedge Δ.

One can easily show that for all $(t, r) \in [0, T] \times \mathbb{R}^m$ there exists a real number $p(t, r)$ such that for all $v \in \mathbb{R}$

$$V^F(t, v - p(t, r), r) = V^0(t, v, r).$$

If an investor has to pay $p(t, r)$ for the derivative $F(R_T^{t,r})$, then he is indifferent between buying and not buying the derivative. Therefore the number $p(t, r)$ is called *indifference price* at time t and level r.

It turns out that the derivative hedge Δ is closely related to the indifference price of the derivative. The derivative either diversifies or amplifies the risk exposure of the portfolio. The difference between $\widehat{\pi}$ and π measures the diversifying impact of F. The price sensitivity, that is the derivative of p relative to the index evolution, is also a measure of the diversification of F (which will be called *diversification pressure* of the derivative F). We will see that the derivative hedge is indeed equal to the price sensitivity multiplied with some correlation parameters.

The problem of finding the optimal strategies π and $\widehat{\pi}$ is a standard stochastic control problem. One can tackle it by solving the related HJB equation, using a verification theorem and proving a uniqueness result. This approach has been chosen for example in Ankirchner et al. (2007). Here, however, we prefer a stochastic approach, using the fact that the stochastic control problem can be solved by finding the solution of a BSDE. In the following section we briefly recall the definition of a BSDE.

3. SOLVING STOCHASTIC OPTIMAL CONTROL PROBLEMS VIA BSDE

Let $\mathcal{H}^2(\mathbb{R}^d)$ be the set of all $\mathbb{R}^d$-valued predictable processes ζ such that $E\int_0^T |\zeta_t|^2\,dt < \infty$, and let $\mathcal{S}^2(\mathbb{R})$ be the set of all $\mathbb{R}$-valued predictable processes δ satisfying $E(\sup_{s\in[0,T]}|\delta_s|^2) < \infty$. By $\mathcal{S}^\infty(\mathbb{R})$ we denote the set of all essentially bounded $\mathbb{R}$-valued predictable processes. Let ξ be $\mathcal{F}_T$-measurable and f a predictable mapping defined on $\Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^d$ with values in $\mathbb{R}$. A solution of the BSDE with *terminal condition* ξ and *generator* f is defined to be a pair of processes $(Y, Z) \in \mathcal{S}^2(\mathbb{R}) \times \mathcal{H}^2(\mathbb{R}^d)$ satisfying

$$Y_t = \xi - \int_t^T Z_s\,dW_s + \int_t^T f(s, Y_s, Z_s)\,ds,\ t \in [0, T].$$

Let us now come back to our control problem of finding the optimal investment strategy π and $\widehat{\pi}$ respectively. It is known that there exists a quadratic BSDE which solves these control problems (see for example, Hu et al. 2005). We first specify the generator of the suitable BSDE, starting with $\hat{\pi}$.

Fix again $(t, r) \in [0, T] \times \mathbb{R}^m$. Let $\vartheta(t, r) = \beta^*(t, r)(\beta(t, r)\beta^*(t, r))^{-1}\alpha(t, r)$ and $C(t, r) = \{x\beta(t, r) : x \in \mathbb{R}^k\}$. Observe that our assumptions imply that $\vartheta(t, r)$ is bounded. The distance of a vector $z \in \mathbb{R}^d$ to the closed and convex set $C(t, r)$ will be defined as $\text{dist}(z, C(t, r)) = \min\{|z - u| : u \in C(t, r)\}$. Let f be the deterministic function

$$\begin{aligned} f : [0, T] \times \mathbb{R}^m \times \mathbb{R}^d \to \mathbb{R},\ (t, r, z) \mapsto z\vartheta(t, r) &+ \frac{1}{2\eta}|\vartheta(t, r)|^2 \\ &- \frac{\eta}{2}\,\text{dist}^2\left(z + \frac{1}{\eta}\vartheta(t, r), C(t, r)\right). \end{aligned}$$

Since $d \geq k$, we have to find the orthogonal projection of the $d-$dimensional vector z to the linear space $C(t, r)$ of image strategies. In Hu et al. (2005) the set $C(t, r)$ is understood as imposing restrictions on the investor when trading in the market that happens to be convex in the setting given.

Notice that f is differentiable in z and satisfies the growth condition

$$|f(t, r, z)| \leq c(1 + |z|^2) \quad \text{a.s.}$$

with some $c \in \mathbb{R}_+$. The growth condition along with the boundedness of the terminal condition guarantees that there exists a unique solution $(\widehat{Y}^{t,r}, \widehat{Z}^{t,r}) \in \mathcal{S}^\infty(\mathbb{R}) \otimes \mathcal{H}^2(\mathbb{R}^d)$ of the BSDE

$$\widehat{Y}_s^{t,r} = F\big(R_T^{t,r}\big) - \int_s^T \widehat{Z}_u^{t,r}\,dW_u - \int_s^T f\big(u, R_u^{t,r}, \widehat{Z}_u^{t,r}\big)\,du, \quad s \in [t, T], \tag{3.1}$$

(see Theorem 2.3 and 2.6. in Kobylanski 2000). Notice that the terminal condition of the BSDE stems from a standard forward SDE. The system of equations consisting of (2.2) and (3.1) is often called FBSDE.

The conditional maximal expected wealth, or in other words the value function of our stochastic control problem, is equal to the utility of the starting point of the BSDE, that is

$$V^F(t, v, r) = -e^{-\eta(v-\widehat{Y}_t^{t,r})}$$

(see Theorem 7 in Hu et al. 2005). Moreover we can reconstruct the optimal strategy $\widehat{\pi}$ starting from $\widehat{Z}$. To this end denote by $\Pi_{C(t,r)}(z)$ the projection of a vector $z \in \mathbb{R}^d$ onto the linear subspace $C(t, r)$. If $R_t = r$, then the optimal strategy $\widehat{\pi}_t$ on $[t, T]$ satisfies

$$\widehat{\pi}_s\beta\big(s, R_s^{t,r}\big) = \Pi_{C(s, R_s^{t,r})}\left[\widehat{Z}_s^{t,r} + \frac{1}{\eta}\vartheta\big(s, R_s^{t,r}\big)\right], \quad s \in [t, T]. \tag{3.2}$$

The last statement follows equally from Theorem 7 in Hu et al. (2005). Since $\vartheta(s, r) \in C(t, r)$ for all $(t, r) \in [0, T] \times \mathbb{R}^m$, equation (3.2) simplifies to $\widehat{\pi}_s\beta(s, R_s^{t,r}) = \Pi_{C(s, R_s^{t,r})}[\widehat{Z}_s^{t,r}] + \frac{1}{\eta}\vartheta(s, R_s^{t,r})$.

Analogously, let $(Y^{t,r}, Z^{t,r})$ be the solution of

$$Y_s^{t,r} = -\int_s^T Z_u^{t,r}\, dW_u - \int_s^T f\big(u, R_u^{t,r}, Z_u^{t,r}\big)\, du, \quad s \in [t, T], \tag{3.3}$$

which represents a stochastic control problem as above, just without the derivative as terminal condition, that is the derivative is not in the portfolio. In this case the maximal expected utility verifies

$$V^0(t, v, r) = -e^{-\eta(v-Y_t^{t,r})},$$

and the optimal strategy π on $[t, T]$ satisfies

$$\pi_s\beta\big(s, R_s^{t,r}\big) = \Pi_{C\left(s, R_s^{t,r}\right)}\left[Z_s^{t,r} + \frac{1}{\eta}\vartheta\big(s, R_s^{t,r}\big)\right], \quad s \in [t, T]. \tag{3.4}$$

Since $\Pi_{C(s, R_s^{t,r})}$ is a linear operator, the derivative hedge is given by the explicit formula

$$\Delta_s\beta\big(s, R_s^{t,r}\big) = \Pi_{C(s, R_s^{t,r})}\big[\widehat{Z}_s^{t,r} - Z_s^{t,r}\big],$$

which will be further determined in the subsequent sections.

4. THE MARKOV PROPERTY OF THE INDIFFERENCE PRICES

In this section we will establish the Markov property of the indifference prices. This will follow from the fact that the solutions of the BSDEs (3.1) and (3.3) are deterministic functions of time and the underlying. To give the precise statement we need to introduce the following σ-algebras. Fixing $t \in [0, T]$, we denote by $\mathcal{D}^m$ the σ-algebra generated by the functions $r \mapsto E[\int_t^T \phi(s, R_s^{t,r})\, ds]$, where $t \in [0, T]$ and ϕ is a bounded continuous $\mathbb{R}-$valued function.

Moreover we assume that the mapping $(t, r) \mapsto \vartheta(t, r)$ is Lipschitz continuous in r, noting that due to (M1) and (M2) this is guaranteed if β and α are Lipschitz continuous.

LEMMA 4.1. *There exist $\mathcal{B}[0, T] \otimes \mathcal{D}^m$-measurable deterministic functions u and $\widehat{u}$: $[0, T] \times \mathbb{R}^m \to \mathbb{R}$ such that*

$$Y_s^{t,r} = u\big(s, R_s^{t,r}\big) \quad \textit{and} \quad \widehat{Y}_s^{t,r} = \widehat{u}\big(s, R_s^{t,r}\big),$$

for $P \otimes \lambda$-a.a. $(\omega, s) \in \Omega \times [t, T]$.

Proof. The generator function f is a polynomial of the components of z of at most second degree. This implies, together with the assumption that ϑ is Lipschitz continuous in r, that there exist functions $f_n : [0, T] \times \mathbb{R}^m \times \mathbb{R}^d \to \mathbb{R}$, globally Lipschitz continuous in z, such that for all compact sets $K \subset \mathbb{R}^m \times \mathbb{R}^d$ the sequence f_n converges to f uniformly on $[0, T] \times K$. Thus the statement follows from Theorem 7.6. □

Lemma 4.1 immediately implies that there exists a nice version of the indifference price p as a function of (t, r).

THEOREM 4.2. *There exists a* $\mathcal{B}[0, T] \otimes \mathcal{D}^m$*-measurable deterministic function* p : $[0, T] \times \mathbb{R}^m \to \mathbb{R}$ *such that for all* $v \in \mathbb{R}$, $(t, r) \in [0, T] \times \mathbb{R}^m$

$$V^F(t, v - p(t, r), r) = V^0(t, v, r). \tag{4.1}$$

Proof. Let $v \in \mathbb{R}$, $(t, r) \in [0, T] \times \mathbb{R}^m$ be given. Recall that $V^F(v, t, r) = -e^{-\eta(v - \widehat{Y}_t^{t,r})}$ and $V^0(v, t, r) = -e^{-\eta(v - Y_t^{t,r})}$. Then put $p(t, r) = u(t, r) - \widehat{u}(t, r)$, where u and $\widehat{u}$ are given from Lemma 4.1. □

In the remainder the function p is always assumed to be measurable in both t and r. In fact it inherits this property from the functions u and $\hat{u}$.

We now turn to an explicit description of the optimal strategies, and in particular their difference, the derivative hedge. These will be derived from the BSDE solutions of the preceding section. We start by noting that similarly to the indifference price the optimal strategies only depend on the time and the index process R.

THEOREM 4.3. *There exist* $\mathcal{B}[0, T] \otimes \mathcal{D}^m$*-measurable deterministic functions* ν *and* $\widehat{\nu}$*, defined on* $[0, T] \times \mathbb{R}^m$ *and taking values in* $\mathbb{R}^d$ *such that for* $(t, r) \in [0, T] \times \mathbb{R}^m$*, the optimal strategies, conditioned on* $R_t = r$*, are given by* $\pi_s = \nu(s, R_s^{t,r})$ *and* $\widehat{\pi}_s = \widehat{\nu}(s, R_s^{t,r})$ for all $s \in [t, T]$.

Proof. Fix $(t, r) \in [0, T] \times \mathbb{R}^m$. Theorem 7.6 implies that there exist $\mathcal{B}[0, T] \otimes \mathcal{D}^m$-measurable deterministic functions v and $\widehat{v}$ mapping $[0, T] \times \mathbb{R}^m$ to $\mathbb{R}^m$ such that for all $s \in [t, T]$

$$Z_s^{t,r} = v(s, R_s^{t,r})\rho(s, R_s^{t,r}) \quad \text{and} \quad \widehat{Z}_s^{t,r} = \widehat{v}(s, R_s^{t,r})\rho(s, R_s^{t,r}).$$

Now let $\gamma(t, r) = \Pi_{C(t,r)}[v(t, r)\rho(t, r) + \frac{1}{\eta}\vartheta(t, r)]$ and $\widehat{\gamma}(t, r) = \Pi_{C(t,r)}[\widehat{v}(t, r)\rho(t, r) + \frac{1}{\eta}\vartheta(t, r)]$. Then, by (3.2) and (3.4), the optimal strategies conditioned on $R_t = r$ satisfy

$$\widehat{\pi}_s \beta(s, R_s^{t,r}) = \widehat{\gamma}(s, R_s^{t,r}) \quad \text{and} \quad \pi_s \beta(s, R_s^{t,r}) = \gamma(s, R_s^{t,r}),$$

for all $s \in [t, T]$. Since the rank of $\beta(t, r)$ is k, then both $\widehat{\nu}(t, r) = \widehat{\gamma}(t, r)\beta^*(t, r)(\beta(t, r)\beta^*(t, r))^{-1}$ and $\nu(t, r) = \gamma(t, r)\beta^*(t, r)(\beta(t, r)\beta^*(t, r))^{-1}$ are well defined. Then uniqueness of π and $\widehat{\pi}$ yields the result. □

REMARK 4.4. Theorem 4.3 implies that the optimal strategies are the so-called *Markov controls*.

We close this section by noting that Theorem 4.2 implies a dynamic principle for the indifference price. Abbreviate $\mathcal{A} = \mathcal{A}^{0,r}$ for some $r \in \mathbb{R}^m$. For any stopping time

$\tau \le T$ and $\mathcal{F}_\tau$-measurable random variable G_τ let $V^F(\tau, G_\tau) = \text{esssup}\{E[U(G_\tau + G_T^{\lambda,\tau} + F(R_T^{0,r})|\mathcal{F}_\tau] : \lambda \in \mathcal{A}\}$. Similarly we define $V^0(\tau, G_\tau)$.

COROLLARY 4.5. *We have*

$$V^F\big(\tau, G_\tau - p\big(\tau, R_\tau^{0,r}\big)\big) = V^0(\tau, G_\tau).$$

Proof. As is shown in Prop. 9 in Hu et al. (2005), the value function V^F satisfies the dynamic principle

$$V^F\big(\tau, G_\tau - p\big(\tau, R_\tau^{0,r}\big)\big) = U\big(G_\tau - p\big(\tau, R_\tau^{0,r}\big) - \widehat{Y}_\tau^{0,r}\big).$$

Since $p(\tau, R_\tau^{0,r}) = Y_\tau^{\tau, R_\tau^{0,r}} - \widehat{Y}_\tau^{\tau, R_\tau^{0,r}} = Y_\tau^{0,r} - \widehat{Y}_\tau^{0,r}$ we obtain $V^F(\tau, G_\tau - p(\tau, R_\tau^{0,r})) = U(G_\tau - Y_\tau^{0,r}) = V^0(\tau, G_\tau)$. □

5. DIFFERENTIABLE INDIFFERENCE PRICES AND EXPLICIT HEDGING STRATEGIES

If we impose stronger conditions on the coefficients of the index process R and the function F, then we can show that the price function p is differentiable in r, and we can obtain an explicit representation of the derivative hedge in terms of the price gradient. To this end we need to introduce the following class of functions.

DEFINITION 5.1. Let $n, \ p \ge 1$. We denote by $\mathbf{B}^{\mathbf{n}\times\mathbf{p}}$ the set of all functions $h : [0, T] \times \mathbb{R}^m \to \mathbb{R}^{n\times p}$, $(t, x) \mapsto h(t, x)$, differentiable in x, for which there exists a constant $C > 0$ such that $\sup_{(t,x)\in[0,T]\times\mathbb{R}^m} \sum_{i=1}^m |\frac{\partial h(t,x)}{\partial x_i}| \le C$, for all $t \in [0, T]$ we have $\sup_{x\in\mathbb{R}^m} \frac{|h(t,x)|}{1+|x|} \le C$, and $x \mapsto \frac{\partial h(t,x)}{\partial x}$ is Lipschitz continuous with Lipschitz constant C.

We will assume that the coefficients of the index diffusion satisfy in addition to (R1)

(R2) $\rho \in \mathbf{B}^{\mathbf{m}\times\mathbf{d}}$, $\mathbf{b} \in \mathbf{B}^{\mathbf{m}\times\mathbf{1}}$, and
(R3) F is a bounded and twice differentiable function such that

$$\nabla F \cdot \rho \in \mathbf{B}^{1\times d} \quad \text{and} \quad \sum_{i=1}^m b_i(t, r)\frac{\partial}{\partial r_i} F(r) + \frac{1}{2}\sum_{i,j=1}^m [\rho\rho^*]_{ij}(t, r)\frac{\partial^2}{\partial r_i \partial r_j} F(r) \in \mathbf{B}^{1\times 1}.$$

The next result guarantees Lipschitz continuity and differentiability of the functions u and $\widehat{u}$ obtained from Theorem 4.1.

THEOREM 5.2. *Suppose that (R1), (R2), and (R3) are satisfied. Besides, suppose that the volatility matrix β and the drift density α are bounded, Lipschitz continuous in r, differentiable in r and that for all $1 \le i \le k, \ 1 \le j \le d$ the derivatives $\nabla_r \beta_{ij}$ and $\nabla_r \alpha_i$ are also Lipschitz continuous in r. Then the functions u and $\widehat{u}$ are Lipschitz continuous in r, and continuously differentiable in r.*

Proof. The theorem follows from Lemma 7.3 and Theorem 7.7 in Section 7. All we have to show at this stage is that the assumptions of both results are satisfied. We only show Conditions (7.8) and (7.11), since the remaining ones are easily seen to be fulfilled.

Notice that the conditions on α and β imply that ϑ is differentiable in r, and that ϑ and $\nabla_r \vartheta$ are globally Lipschitz continuous in r. Moreover, since ϑ is bounded, ϑ^2 is Lipschitz continuous in r, too. Recalling the definition of the generator f, note further that

$$\begin{aligned}
&\left|\operatorname{dist}^2\left(z+\frac{1}{\eta}\vartheta(t,r), C(t,r)\right)-\operatorname{dist}^2\left(z+\frac{1}{\eta}\vartheta(t,r'), C(t,r)\right)\right| \\
&\quad \le 2\left(\frac{1}{\eta}\|\vartheta\|_\infty+|z|\right)\left|\operatorname{dist}\left(z+\frac{1}{\eta}\vartheta(t,r), C(t,r)\right)-\operatorname{dist}\left(z+\frac{1}{\eta}\vartheta(t,r'), C(t,r)\right)\right| \\
&\quad \le 2\left(\frac{1}{\eta}\|\vartheta\|_\infty+|z|\right)\left|\left|z+\frac{1}{\eta}\vartheta(t,r)-\Pi_{C(t,r)}\left(z+\frac{1}{\eta}\vartheta(t,r)\right)\right|\right. \\
&\qquad \left.-\left|z+\frac{1}{\eta}\vartheta(t,r')-\Pi_{C(t,r)}\left(z+\frac{1}{\eta}\vartheta(t,r')\right)\right|\right| \\
&\quad \le 4\left(\frac{1}{\eta}\|\vartheta\|_\infty+|z|\right)\frac{1}{\eta}|\vartheta(t,r)-\vartheta(t,r')|, \quad t\in[0,T], r, \quad r'\in\mathbb{R}^m, \quad z\in\mathbb{R}^d.
\end{aligned}$$

This shows that there exists a constant $K\in\mathbb{R}_+$ such that $|f(t,r,z)-f(t,r',z)|\le K(1+|z|)|r-r'|$, and hence Assumption (7.8) of Lemma 7.3 is satisfied.

Observe that $\Pi_{C(t,r)}(y)=y\beta^*(\beta\beta^*)^{-1}\beta(t,r)$ for all $y\in\mathbb{R}^k$, and hence the mapping $r\mapsto\Pi_{C(t,r)}(z+\frac{1}{\eta}\vartheta(t,r))$ is differentiable. Consequently, also f is differentiable and for $t\in[0,T]$, $r\in\mathbb{R}^m$ and $z\in\mathbb{R}^d$ we have

$$\begin{aligned}
\nabla_r f(t,r,z) &= z\nabla_r\vartheta(t,r)+\frac{1}{\eta}\vartheta(t,r)\nabla_r\vartheta(t,r) \\
&\quad -\eta\left(z+\frac{1}{\eta}\vartheta(t,r)-\Pi_{C(t,r)}\left(z+\frac{1}{\eta}\vartheta(t,r)\right)\right) \\
&\quad \times\left(\frac{1}{\eta}\nabla_r\vartheta(t,r)-\nabla_r\Pi_{C(t,r)}\left(z+\frac{1}{\eta}\vartheta(t,r)\right)\right).
\end{aligned}$$

By using that ϑ, $\nabla_r\vartheta$, β and $\nabla_r\beta$ are Lipschitz continuous and bounded, it is straightforward to show that for all $t\in[0,T]$, $r,r'\in\mathbb{R}^m$ and $z,z'\in\mathbb{R}^d$ we have

$$|\nabla_r f(t,r,z)-\nabla_r f(t,r',z')|\le K(1+|z|+|z'|)(|r-r'|+|z-z'|),$$

and hence the generator satisfies Assumption (7.11) of Theorem 7.7.

Now Lemma 7.3 yields the Lipschitz continuity in r of the functions u and $\widehat{u}$. Theorem 7.7 implies the differentiability of $\widehat{Y}^{t,r}$ and $Y^{t,r}$ with respect to r, and hence also of u and $\widehat{u}$. □

As an immediate consequence we obtain smoothness of the indifference price function.

COROLLARY 5.3. *Suppose that the assumptions of Theorem 5.2 are satisfied. Then the indifference price function p is continuously differentiable in r.*

Having shown smoothness of the indifference price, we can finally derive an explicit formula for the derivative hedge in terms of the price gradient. To this end we denote the *conditional derivative hedge* by $\Delta(t,r)=\widehat{\nu}(t,r)-\nu(t,r)$, $(t,r)\in[0,T]\times\mathbb{R}^m$.

THEOREM 5.4. *Under the assumptions of Theorem 5.2, and with the notation of Section 3, the derivative hedge satisfies*

$$\Delta(t,r)=-\nabla_r p(t,r)\rho(t,r)\beta^*(t,r)(\beta(t,r)\beta^*(t,r))^{-1}, \quad (t,r)\in[0,T]\times\mathbb{R}^m. \tag{5.1}$$

REMARK 5.5. Note that Theorem 5.4 implies that the derivative hedge at time t depends only on R_t.

Proof of Theorem 5.4. Note that $C(t,r)$ is a linear subspace of $\mathbb{R}^d$ for all $(t,r) \in [0,T] \times \mathbb{R}^m$. Therefore, the projection operator $\Pi_{C(t,r)}$ is linear and hence

$$\begin{aligned}\Delta(t,r) &= (\widehat{\gamma}(t,r) - \gamma(t,r))\beta^*(t,r)(\beta(t,r)\beta^*(t,r))^{-1} \\ &= \left(\Pi_{C(t,r)}\left[\widehat{Z}_t^{t,r} + \frac{1}{\eta}\vartheta(t,r)\right] - \Pi_{C(t,r)}\left[Z_t^{t,r} + \frac{1}{\eta}\vartheta(t,r)\right]\right) \\ &\quad \times \beta^*(t,r)(\beta(t,r)\beta^*(t,r))^{-1} \\ &= \left(\Pi_{C(t,r)}\left[\widehat{Z}_t^{t,r} - Z_t^{t,r}\right]\right)\beta^*(t,r)(\beta(t,r)\beta^*(t,r))^{-1}.\end{aligned}$$

It follows from Theorem 7.7 that $\widehat{Z}_t^{t,r} - Z_t^{t,r} = (\nabla_r \widehat{u}(t,r) - \nabla_r u(t,r))\rho(t,r) = -\nabla_r p(t,r)\rho(t,r)$, and hence we obtain the result. □

If the market consists of only one risky asset, then the optimal strategy simplifies to the following formula.

COROLLARY 5.6. *Let $k = 1$. Then the derivative hedge is given by*

$$\begin{aligned}\Delta(t,r) &= -\frac{\langle \beta(t,r), \nabla_r p(t,r)\rho(t,r)\rangle}{|\beta(t,r)|^2} \\ &= -\frac{\sum_{i=1}^{d} \beta_i(t,r) \sum_{j=1}^{m} \frac{\partial}{\partial r_j} p(t,r)\rho_{ji}(t,r)}{\sum_{i=1}^{d} \beta_i^2(t,r)}, \qquad (t,r) \in [0,T] \times \mathbb{R}^m.\end{aligned}$$

Proof. Fix $(t,r) \in [0,T] \times \mathbb{R}^m$. Note that $C(t,r) = \{x\beta(t,r) : x \in \mathbb{R}\}$ is a one-dimensional subspace of $\mathbb{R}^d$. For all $z = (z_i)_{1 \le i \le d} \in \mathbb{R}^d$ let $g(z) = \frac{\langle \beta(r,t), z\rangle}{|\beta(t,r)|^2} = \frac{\sum_{i=1}^d \beta_i(t,r)\, z_i}{\sum_{i=1}^d \beta_i^2(t,r)}$. Then $g(z)\beta(t,r)$ is the orthogonal projection of z onto $C(t,r)$. Thus Theorem 5.4 yields that $\Delta(t,r) = -g(\nabla_r p(t,r)\rho(t,r))$. □

REMARK 5.7.

(1) Suppose the derivative $F(R_T)$ is traded on an exchange. By pretending the price observed is approximately equal to an indifference price, the hedging formula (5.1) provides a very simple tool for hedging the derivative. Notice that the risk aversion coefficient η does not appear explicitly in (5.1).

(2) If $k = d$ and the matrices $\beta(t,r)$ are all invertible, then our financial market is complete and the derivative $F(R_T)$ can be fully replicated. Moreover the derivative hedge satisfies

$$\Delta(t,r) = -\nabla_r p(t,r)\rho(t,r)\beta^{-1}(t,r).$$

If S is chosen to be the index, that is $R = S$, then we obtain $\Delta = (\frac{\partial p}{\partial s_1}(t,r)S^1, \ldots, \frac{\partial p}{\partial s_k}(t,r)S^k)$. Moreover, the number of shares to invest into asset i is given by $\frac{\Delta^i(t,r)}{S^i(t,r)} = \frac{\partial p}{\partial s_i}$. Thus Δ coincides with the classical "delta hedge."

EXAMPLE 5.8. As in Example 2.1 suppose that R is the moving average cHDD process modelled as a geometric Brownian motion, and assume that there exists one tradable correlated risky asset. More precisely let $d = 2,\ k = m = 1,\ \rho = (\alpha_2\ 0),\ \beta = (\beta_1\ \beta_2)$ with $\alpha_2, \beta_1, \beta_2 \in \mathbb{R} \setminus \{0\}$. Then

$$\Delta(t,r) = -\alpha_2 \frac{\partial p(t,r)}{\partial r} \frac{\beta_1}{\beta_1^2 + \beta_2^2}.$$

EXAMPLE 5.9. Applying our results to Example 2.2, we have to take $m = 2,\ k = 2$ and $d = 3$. Hence

$$\rho = \begin{pmatrix} \gamma_1 & 0 & 0 \\ \gamma_2 & \gamma_3 & \gamma_4 \end{pmatrix}, \qquad \beta = \begin{pmatrix} \gamma_1 & 0 & 0 \\ \beta_1 & \beta_2 & 0 \end{pmatrix}, \qquad \beta^*(\beta\beta^*)^{-1} = \frac{1}{\gamma_1\beta_2} \begin{pmatrix} \beta_2 & 0 \\ -\beta_1 & \gamma_1 \\ 0 & 0 \end{pmatrix}.$$

With a simple computation we have for $(t,r) \in [0, T] \times \mathbb{R}^m$

$$\Pi_{C(t,r)}[\nabla_r p(t,r)\rho(t,r)] = \left(\gamma_1 \frac{\partial}{\partial r_1} p(t,r) + \gamma_2 \frac{\partial}{\partial r_2} p(t,r) \quad \gamma_3 \frac{\partial}{\partial r_2} p(t,r) \quad 0 \right).$$

Equation (5.1) applied to our example produces the following Delta hedge for $(t,r) \in [0, T] \times \mathbb{R}^2$

$$\Delta(t,r) = \left(-\frac{\partial}{\partial r_1} p(t,r) + \left(\frac{\beta_1\gamma_3}{\gamma_1\beta_2} - \frac{\gamma_2}{\gamma_1} \right) \frac{\partial}{\partial r_2} p(t,r) \quad - \frac{\gamma_3}{\beta_2} \frac{\partial}{\partial r_2} p(t,r) \right),$$

where r_1 represents the crude oil and r_2 the kerosene variable. If $\gamma_4 = 0$ then we have a perfect hedge and if $\gamma_3 = 0$, then the price of heating oil does not play a role in the hedge, as one would expect.

6. PRICING BY MARGINAL UTILITY

Suppose there is no exchange and the derivative $F(R_T)$ is sold OTC. What is a reasonable price a seller could ask for the derivative? The indifference price seems to be a natural candidate, though it has the disadvantage that the price of a single derivative depends on the total quantity sold, that is the indifference price is nonlinear. For instance the indifference price of $2 \times F(R_T)$ does not equal twice the indifference price of $F(R_T)$. In order to obtain a linear version one may take the limit of the indifference price as the quantity converges to 0. The object thus derived is the indifference price for a vanishing amount of derivatives, and it is therefore called *marginal utility price* (MUP). Having to pay the MUP for each derivative an investor is indifferent between buying and not buying an infinitesimal amount of the derivative.

We continue requiring (R1)-(R3) to be satisfied. We update the notation and, for $q \in \mathbb{R}$ and $(t,r) \in [0, T] \times \mathbb{R}^m$ define by $p(t,r,q)$ the indifference price of q units of $F(R_T^{t,r})$, that is $p(t,r,q)$ is the unique real satisfying

$$\sup_\lambda \left\{ EU\left(v + G_T^{\lambda,t,r} + qF\left(R_T^{t,r}\right) - p(t,r,q)\right\} = \sup_\lambda \left\{ EU\left(v + G_T^{\lambda,t,r}\right)\right\}.$$

The price of one unit is equal to $\frac{p(t,r,q)}{q}$, ($q \neq 0$), and the MUP is defined by

$$\mathrm{MUP}(t,r) = \frac{\partial}{\partial q} p(t,r,q)\bigg|_{q=0}.$$

Recall that $p(t,r,q) = Y_t^{t,r} - \widehat{Y}_t^{t,r,q}$, where $(Y^{t,r}, Z^{t,r})$ is the solution of BSDE (3.3) and $(\widehat{Y}^{t,r,q}, \widehat{Z}^{t,r,q})$ is the solution of the BSDE

$$\widehat{Y}_s^{t,r,q} = qF(R_T^{t,r}) - \int_s^T \widehat{Z}_u^{t,r,q}\, dW_u - \int_s^T f(u, R_u^{t,r}, \widehat{Z}_u^{t,r,q})\, du, \quad s \in [t,T].$$

Naming $\xi(q) = qF(R_T^{t,r})$, then clearly $\xi(q)$ is a globally bounded differentiable Lipschitz function (with bounded derivatives). The boundedness of ξ is trivial since F is bounded and we are only interested in the differentiability of the process with relation to q in a neighborhood of zero. And so, due to the boundedness of F and the quadratic growth hypothesis for f the conditions of Theorem 7.8 are satisfied. Hence, the process $\widehat{Y}^{t,r,q}$ is continuous in t and continuously differentiable in q.

Writing the BSDE differentiated with respect to q gives

$$\begin{aligned}\frac{\partial}{\partial q}\widehat{Y}_s^{t,r,q} &= F(R_T^{t,r}) - \int_s^T \frac{\partial}{\partial q}\widehat{Z}_u^{t,r,q}\, dW_u \\ &\quad - \int_s^T \nabla_z f(u, R_u^{t,r}, \widehat{Z}_u^{t,r,q}) \frac{\partial}{\partial q}\widehat{Z}_u^{t,r,q}\, du, \quad s \in [t,T].\end{aligned}$$

Setting $q = 0$ and renaming the processes for ease of notation we obtain

$$U_s^{t,r} = F(R_T^{t,r}) - \int_s^T V_s\, dW_s - \int_s^T \nabla_z f(s, R_s^{t,r}, Z_s^{t,r}) \cdot V_s\, ds. \tag{6.1}$$

As an end product of these calculations we obtain the following explicit formula for the (MUP) of our derivative.

THEOREM 6.1. *The explicit formula for the Marginal Utility Price of the derivative $F(R_T)$ is given by*

$$MUP(t,r) = U_t^{t,r},$$

where $U_t^{t,r}$ is the first component of the solution pair of the BSDE

$$U_s^{t,r} = F(R_T^{t,r}) - \int_s^T V_s\, dW_s - \int_s^T \nabla_z f(s, R_s^{t,r}, Z_s^{t,r}) \cdot V_s\, ds. \tag{6.2}$$

REMARK 6.2. Note that by performing a Girsanov change of probability measure to the one making the process $\tilde{W} = W + \int_0^\cdot \nabla_z f(s, R_s^{t,r}, Z_s^{t,r})\, ds$ a Brownian motion, solving (6.2) reduces to taking conditional expectations with respect to the underlying filtration. Hence, denoting by $\mathcal{E}(\cdot)$ the stochastic exponential operator, we can represent the marginal utility price explicitly by the following expression

$$MUP(t,r) = E\left[\mathcal{E}\left(\int_0^\cdot \nabla_z f(s, R_s^{t,r}, Z_s^{t,r})\, dW_s\right)_t^T F(R_T^{t,r})\right].$$

7. SOME MATHEMATICAL TOOLS: SMOOTHNESS OF QUADRATIC FBSDE

7.1. Moment Estimates for BSDE with Random Lipschitz Condition

In the following we provide moment estimates for BSDE with generators that satisfy Lipschitz conditions with random bounds for the slopes. More precisely, we assume that for our generator $f : \Omega \times [0, T] \times \mathbb{R}^d \to \mathbb{R}$ there exists an $\mathbb{R}_+$-valued predictable process H such that for all $(\omega, t, z) \in \Omega \times [0, T] \times \mathbb{R}^d$ we have

$$|f(\omega, t, z) - f(\omega, t, z')| \leq H_t |z - z'|. \tag{7.1}$$

We will assume that H is such that the stochastic integral $\int_0^\cdot H\,dW$ with respect to a Brownian motion W is a so-called BMO martingale. Recall that $\int_0^\cdot H\,dW$ is a BMO martingale (we also say it belongs to BMO) if and only if there exists a constant $C \in \mathbb{R}_+$ independent of ω such that for all stopping times τ with values in $[0, T]$ we have

$$E\left[\int_\tau^T H_s^2\,ds \middle| \mathcal{F}_\tau\right] \leq C, \qquad \text{a.s.} \tag{7.2}$$

We refer to Kazamaki (1994) for basic information about BMO martingales. We will abuse the definition and refer to the smallest $C \in \mathbb{R}_+$ that satisfies inequality (7.2) as the BMO norm of H.

Throughout let W be a d-dimensional Brownian motion. Consider the BSDE

$$Y_t = \xi - \int_t^T Z_s\,dW_s + \int_t^T f(s, Z_s)\,ds, \quad 0 \leq t \leq T, \tag{7.3}$$

where ξ is a bounded $\mathcal{F}_T$-measurable random variable, and f satisfies (7.1) relative to a predictable H with finite BMO norm.

We refer to Briand and Confortola (2008) for sufficient criteria for the existence of solutions of such BSDEs.

The moment estimate we shall give next will be needed later for establishing smoothness of the solution of the quadratic BSDE with respect to the parameters the terminal condition depends on.

LEMMA 7.1. *Suppose that for all $\beta \geq 1$ we have $\int_0^T |f(s, 0)|\,ds \in L^\beta(P)$. Let $p > 1$. Then there exist constants $q > 1$ and $C > 0$, depending only on p, T, and the BMO-norm of H, such that we have*

$$E\left[\sup_{t\in[0,T]} |Y_t|^{2p}\right] + E\left[\left(\int_0^T |Z_s|^2\,ds\right)^p\right] \leq C\left(E\left[|\xi|^{2pq} + \left(\int_0^T |f(s, 0)|\,ds\right)^{2pq}\right]\right)^{\frac{1}{q}}.$$

Proof. This follows from Corollary 3.4 in Briand and Confortola (2008) by a straightforward generalization to the multidimensional case considered here. It can also be shown with the method used in the proof of Theorem 6.1 in Ankirchner et al. (2007). □

7.2. Differentiability of Quadratic FBSDE

Consider now a FBSDE of the form

$$X_s^x = x + \int_0^t b\big(s, X_s^x\big)\, ds + \int_0^t \rho\big(s, X_s^x\big)\, dW_s, \tag{7.4}$$
$$Y_s^x = F\big(X_T^x\big) - \int_t^T Z_s^x\, dW_s + \int_t^T f\big(s, X_s^x, Z_s^x\big)\, ds,$$

where $b : [0, T] \times \mathbb{R}^m \to \mathbb{R}^m$ and $\rho : [0, T] \times \mathbb{R}^m \to \mathbb{R}^{m\times d}$ and W is the $d-$dimensional Brownian motion of the preceding subsection. Note that ρ is a $n \times d$ matrix. We will denote its transpose by ρ^*. The generator of the backward part is assumed to be a $\mathcal{P}(\mathcal{F}_t) \otimes \mathcal{B}(\mathbb{R}^m) \otimes \mathcal{B}(\mathbb{R}^d)$-measurable process $f : \Omega \times [0, T] \times \mathbb{R}^m \times \mathbb{R}^d \to \mathbb{R}$ such that there exists a constant $M \in \mathbb{R}_+$ such that for all $(t, x, z) \in [0, T] \times \mathbb{R}^m \times \mathbb{R}^d$ we have

$$|f(t, x, z)| \le M(1 + |z|^2) \quad \text{a.s.} \tag{7.5}$$

Here $\mathcal{P}(\mathcal{F}_t)$ denotes the σ-field of predictable sets with respect to the filtration $(\mathcal{F}_t)$. Moreover we assume that

(7.6) f is differentiable in x and z and

$$|\nabla_z f(t, x, z)| \le M(1 + |z|) \quad \text{for all} \quad (t, x, z) \in [0, T] \times \mathbb{R}^m \times \mathbb{R}^d \quad \text{a.s.}$$

We will give sufficient conditions for the process Y^x in the solution of the FBSDE (7.4) to be differentiable in x. A further assumption we need is that the coefficients of the forward equation belong to the function space $\mathbf{B}^{\mathbf{m\times d}}$ and $\mathbf{B}^{\mathbf{m\times 1}}$ respectively (see Definition 5.1). To simplify notation, to the pair (b, ρ) of coefficient functions we associate the second order differential operator $\mathcal{L} = \sum_{i=1}^m b_i(\cdot)\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^m [\rho\rho^*]_{ij}(\cdot)\frac{\partial^2}{\partial x_i \partial x_j}$.

We will assume that the coefficients of the forward equation (7.4) satisfy

(D1) $\rho \in \mathbf{B}^{\mathbf{m\times d}}$, $b \in \mathbf{B}^{\mathbf{m\times 1}}$, and that

(D2) $F : \mathbb{R}^m \to \mathbb{R}$ is a twice differentiable function such that $\nabla F \cdot \rho \in \mathbf{B}^{\mathbf{1\times d}}$ and $\mathcal{L}F \in \mathbf{B}^{\mathbf{1\times 1}}$.

It is known that the conditions (D1) and (D2) ensure that X^x is differentiable in x and its difference quotients can be nicely controlled. For the convenience of the reader we quote a standard result which will be needed later. Denote by e_i the unit vector in $\mathbb{R}^m$ in the direction of coordinate i, $1 \le i \le m$.

LEMMA 7.2. *Suppose (D1) and (D2) are satisfied. For all* $x \in \mathbb{R}^m$, $h \neq 0$ *and* $i \in \{1, \dots, m\}$, *let* $\zeta^{x,h,i} = \frac{1}{h}(F(X_T^{x+he_i}) - F(X_T^x))$. *Then for every* $p > 1$ *there exists a* $C > 0$, *dependent only on p and the bounds of* b, ρ, F *and its derivatives, such that for all* x, $x' \in \mathbb{R}^m$ *and* h, $h' \neq 0$,

$$E\big[|\zeta^{x,h,i} - \zeta^{x',h',i}|^{2p}\big] \le C(|x - x'|^2 + |h - h'|^2)^p. \tag{7.7}$$

Proof. Note that by Ito's formula $F(X_t^x) = F(X_0^x) + \int_0^t \nabla F(X_s^x) \cdot \rho(s, X_s^x)\, dW_s + \int_0^t \mathcal{L}F ds$. Thus $F(X_t^x)$ is a diffusion with coefficients $\tilde{\rho}(s, x) = \nabla F(x) \cdot \rho(s, x)$ and $\tilde{b}(s, x) = \sum_{i=1}^m b_i(s, x)\frac{\partial F(x)}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^m \rho_{ij}(s, x)\frac{\partial^2 F(x)}{\partial x_i \partial x_j}$, $(s, x) \in [0, T] \times \mathbb{R}^m$. By (D2) we have $\tilde{\rho} \in \mathbf{B}^{\mathbf{1\times d}}$ and $\tilde{b} \in \mathbf{B}^{\mathbf{1\times 1}}$. Therefore, by using standard results on stochastic flows (see Lemma 4.6.3 in Kunita 1990), we obtain the result. □

Notice that since F is bounded and the growth condition (7.5) holds, there exists a unique solution $(Y^x, Z^x) \in \mathcal{S}^\infty(\mathbb{R}) \otimes \mathcal{H}^2(\mathbb{R}^d)$ of the BSDE in (7.4) for all $x \in \mathbb{R}^d$. One can even show that we may choose the family $(Y^x)_{x\in\mathbb{R}^m}$ such that it is continuous in x.

LEMMA 7.3. *Let (D1), (D2), (7.5) and (7.6) be satisfied, and assume that F is bounded and that there exists a constant $K \in \mathbb{R}_+$ such that for all $t \in [0, T]$, $x, x' \in$ and $z \in \mathbb{R}^d$*

$$|f(t, x, z) - f(t, x', z)| \le K(1 + |z|)|x - x'|. \tag{7.8}$$

Then for all $p > 1$ there exists a constant $C \in \mathbb{R}_+$ such that for all $x, x' \in \mathbb{R}^m$,

$$E \sup_{t\in[0,T]} \left| Y_t^x - Y_t^{x'} \right|^{2p} \le C|x - x'|^{2p}, \tag{7.9}$$

$$E\left[\left(\int_0^T \left|Z_t^x - Z_t^{x'}\right|^2 dt\right)^p\right] \le C|x - x'|^{2p}. \tag{7.10}$$

In particular, Kolmogorov's continuity criterion implies that there exists a measurable process $\widetilde{Y} : \Omega \times [0, T] \times \mathbb{R}^m$ such that $(t, x) \mapsto \widetilde{Y}_t^x$ is continuous for a.a. ω; and for all $(t, x) \in [0, T] \times \mathbb{R}^m$ we have $\widetilde{Y}_t^x = Y_t^x$ a.s.

Proof. For $\alpha \in \mathbb{R}$, let $\chi(y) = e^{\alpha y}$. By applying Ito's formula to $\chi(Y^x)$ and using standard arguments one can show that $\int_0^\cdot Z^x \, dW \in \text{BMO}$ with the BMO norm depending only on the bound of F and the growth constant of f in z.

For all $x, x' \in \mathbb{R}^m$ let $U_t = Y_t^x - Y_t^{x'}$, $V_t = Z_t^x - Z_t^{x'}$ and $\zeta = F(X^x) - F(X^{x'})$. We use a line integral transformation in order to show that U^x can be seen as a BSDE with generator satisfying a Lipschitz condition of the type (7.1). Define $J_t = \int_0^1 \nabla_x f(t, X_t^x - \vartheta(X_t^x - X_t^{x'}), Z_t^x) \, d\vartheta$ and $H_t = \int_0^1 \nabla_z f(t, X_t^{x'}, Z_t^{x'} - \vartheta(Z_t^x - Z_t^{x'})) \, d\vartheta$ and observe that

$$\begin{aligned} U_t &= \zeta - \int_t^T V_s \, dW_s + \int_t^T \big(f\big(s, X_s^x, Z_s^x\big) - f\big(s, X_s^{x'}, Z_s^x\big)\big) \\ &\quad + \big(f\big(s, X_s^{x'}, Z_s^x\big) - f\big(s, X_s^{x'}, Z_s^{x'}\big)\big) \, ds \\ &= \zeta - \int_t^T V_s \, dW_s + \int_t^T \big(J_s\big(X_s^x - X_s^{x'}\big) + H_s V_s\big) \, ds. \end{aligned}$$

The moment estimate of Lemma 7.1 applied to the pair (U, V) leads to

$$\begin{aligned} &E\left[\sup_{t\in[0,T]} |U_t|^{2p}\right] + E\left[\left(\int_0^T |V_s|^2 \, ds\right)^p\right] \\ &\quad \le C \left(E\left[|\zeta|^{2pq} + \left(\int_0^T \left|J_s\big(X_s^x - X_s^{x'}\big)\right| ds\right)^{2pq}\right]\right)^{\frac{1}{q}}, \end{aligned}$$

for some constants $C > 0$ and $q > 1$. By (7.8) we have $\nabla_x f(t, x, z) \leq K(1 + |z|)$, and hence

$$E\left(\int_0^T |J_s(X_s^x - X_s^{x'})|\, ds\right)^{2pq}$$
$$\leq K^{2pq}\left(E\left(\int_0^T (1 + |Z_s^x|)^2 ds\right)^{2pq}\right)^{\frac{1}{2}}\left(E\left(\int_0^T |X_s^x - X_s^{x'}|^2 ds\right)^{2pq}\right)^{\frac{1}{2}}.$$

Lemma 7.1 implies that $E(\int_0^T (1 + |Z^x|)^2\, ds)^{2pq}$ is bounded, and by standard results on moment estimates of SDEs we have $E(\int_0^T |X_s^x - X_s^{x'}|^2 ds)^{2pq} \leq C'|x - x'|^{4pq}$ for some constant $C' \in \mathbb{R}_+$ (see Theorem 3.2 in Kunita 2004). Moreover, the Lipschitz property of F guarantees that there exists a constant C'' such that $E|\zeta|^{2pq} \leq C''|x - x'|^{2pq}$, and hence the desired result follows. □

The following theorem guarantees pathwise continuous differentiability of an appropriate modification of the solution process.

THEOREM 7.4. *Let (D1), (D2), (7.5), and (7.6) be satisfied, and suppose that F is bounded and f satisfies (7.8). Besides suppose that* $\nabla_z f$ *is globally Lipschitz continuous in (x,z) and that* $\nabla_x f$ *satisfies for all* $t \in [0, T]$, $x, x' \in$ *and* $z, z' \in \mathbb{R}^d$

$$|\nabla_x f(t, x, z) - \nabla_x f(t, x', z')| \leq K(1 + |z| + |z'|)(|x - x'| + |z - z'|). \tag{7.11}$$

Then there exists a function $\Omega \times [0, T] \times \mathbb{R}^m \to \mathbb{R}^{m+1+d}$, $(\omega, t, x) \mapsto (X_t^x, Y_t^x, Z_t^x)(\omega)$, *such that for almost all* ω, X_t^x *and* Y_t^x *are continuous in t and continuously differentiable in x, and for all x,* (X_t^x, Y_t^x, Z_t^x) *is a solution of FBSDE (7.4). Moreover, there exists a process* $\nabla_x Z^x \in \mathcal{H}^2$ *such that the pair* $(\nabla_x Y^x, \nabla_x Z^x)$ *solves the BSDE*

$$\nabla_x Y_t^x = \nabla_x F(X_T^x)\nabla_x X_T^x - \int_t^T \nabla_x Z_s^x\, dW_s + \int_t^T [\nabla_x f(s, X_s^x, Z_s^x)\nabla_x X_s^x + \nabla_z f(s, X_s^x, Z_s^x)\nabla_x Z_s^x]\, ds. \tag{7.12}$$

We will use Kolmogorov's continuity criterion in order to prove the theorem. Let $x \in \mathbb{R}^m$. For all $h \neq 0$, let $\Delta_t^{x,h} = \frac{1}{h}(X_t^{x+he_i} - X_t^x)$, $U_t^{x,h} = \frac{1}{h}(Y_t^{x+e_ih} - Y_t^x)$, $V_t^{x,h} = \frac{1}{h}(Z_t^{x+he_i} - Z_t^x)$ and $\zeta^{x,h} = \frac{1}{h}(\xi(x + he_i) - \xi(x))$. We need the following estimates.

LEMMA 7.5. *For all* $p > 1$, $x, x' \in \mathbb{R}^m, h, h' \neq 0$ *we have with some constant C*

$$E\left[\sup_{t\in[0,T]} |U_t^{x,h} - U_t^{x',h'}|^{2p}\right] \leq C(|x - x'|^2 + |h - h'|^2)^p. \tag{7.13}$$

Proof. Let $p > 1$. Note that for all $h \neq 0$

$$U_t^{x,h} = \zeta^{x,h} - \int_t^T V_s^{x,h} dW_s + \int_t^T \frac{1}{h}[f(s, X_s^{x+he_i}, Z_s^{x+he_i}) - f(s, X_s^x, Z_s^x)]\, ds.$$

We use a line integral transformation in order to show that $U^{x,h}$ can be seen as a BSDE with random Lipschitz bound. To this end define two $(\mathcal{F}_t)$-adapted processes by

$$A_t^{x,h} = \int_0^1 \nabla_x f\big(t, X_t^x + \vartheta\big(X_t^{x+he_i} - X_t^x\big), Z_t^x\big)\, d\vartheta,$$
$$I_t^{x,h} = \int_0^1 \nabla_z f\big(t, X_t^{x+he_i}, Z_t^x + \vartheta\big(Z_t^{x+he_i} - Z_t^x\big)\big)\, d\vartheta.$$

Then

$$\frac{1}{h}\big[f\big(t, X_t^{x+he_i}, Z_t^{x+he_i}\big) - f\big(t, X_t^x, Z_t^x\big)\big] = A_t^{x,h}\Delta_t^{x,h} + I_t^{x,h} V_t^{x,h}.$$

The growth condition (7.6) implies that $|I^{x,h}| \le M(1 + |Z^x| + |Z^{x+he_i}|)$, and hence $\int_0^\cdot I^{x,h}\, dW \in \mathrm{BMO}$. Thus we obtain a BSDE with generator satisfying condition (7.1).

Now let $x, x' \in \mathbb{R}^m$ and $h, h' \neq 0$. Then the difference $(U^{x,h} - U^{x',h'}, V^{x,h} - V^{x',h'})$ solves again a BSDE with generator of the type (7.1), namely

$$U_t^{x,h} - U_t^{x',h'} = \zeta^{x,h} - \zeta^{x',h'} - \int_t^T \big(V_s^{x,h} - V_s^{x',h'}\big)\, dW_s$$
$$- \int_t^T \big(I^{x,h}\big(V_s^{x,h} - V_s^{x',h'}\big) + \big(I_s^{x,h} - I_s^{x',h'}\big)V^{x',h'} + A_s^{x,h}\Delta_s^{x,h} - A_s^{x',h'}\Delta_s^{x',h'}\big)\, ds.$$

Therefore Lemma 7.1 yields for $q > 1$

$$E\Big[\sup_{t\in[0,T]} \big|U_t^{x,h} - U_t^{x',h'}\big|^{2p}\Big]$$
$$\le C\Bigg\{E\bigg[|\zeta^{x,h} - \zeta^{x',h'}|^{2pq} + E\bigg[\bigg(\int_0^T \Big(\big|A_s^{x,h}\Delta_s^{x,h} - A_s^{x',h'}\Delta_s^{x',h'}\big|$$
$$+ \big|I_s^{x,h} - I_s^{x',h'}\big|\big|V^{x',h'}\big|\Big)\, ds\bigg)^{2pq}\bigg]^{\frac{1}{q}}\Bigg\}.$$

To treat the first term on the right hand side, use Lemma 7.2 to see that $E[|\zeta^{x,h} - \zeta^{x',h'}|^{2pq}]^{\frac{1}{q}} \le C(|x - x'|^2 + |h - h'|^2)^p$.

For the second term, recall that $\nabla_z f$ is Lipschitz continuous, say with Lipschitz constant $L \in \mathbb{R}_+$. We therefore have for any $t \in [0, T]$

$$\big|I_t^{x,h} - I_t^{x',h'}\big| \le L\big(\big|X_t^{x,h} - X_t^{x',h'}\big| + \big|Z_t^x - Z_t^{x'}\big| + \big|Z_t^{x+he_i} - Z_t^{x'+h'e_i}\big|\big).$$

Now Cauchy-Schwarz' inequality leads to

$$E\bigg[\bigg(\int_0^T \big|I_s^{x,h} - I_s^{x',h'}\big|\big|V^{x',h'}\big|\, ds\bigg)^{2pq}\bigg]^{\frac{1}{q}}$$
$$\le \bigg(E\bigg[\bigg(\int_0^T \big|I_s^{x,h} - I_s^{x',h'}\big|^2\, ds\bigg)^{2pq}\bigg] E\bigg[\bigg(\int_0^T |V^{x,h}|^2\, ds\bigg)^{2pq}\bigg]\bigg)^{\frac{1}{2q}}.$$

So Lemma 7.3 and Lemma 4.5.6 in Kunita (1990) imply with some constant C

$$E\left[\left(\int_0^T \left|I_s^{x,h} - I_s^{x',h'}\right|^2 ds\right)^{2pq}\right]^{\frac{1}{2q}} \le C(|x-x'|^2 + |h-h'|^2)^p.$$

The term $E[(\int_0^T |V^{x,h}|^2\, ds)^{2pq}]$ is seen to be bounded by an appeal to Lemma 7.3.

It remains to show that $E[(\int_0^T |A_s^{x,h}\Delta_s^{x,h} - A_s^{x',h'}\Delta_s^{x',h'}|\, ds)^{2pq}]^{\frac{1}{q}} \le C(|x-x'|^2 + |h-h'|^2)^p$. For this we separately estimate the two summands on the right hand side of the following inequality

$$\left|A_s^{x,h}\Delta_s^{x,h} - A_s^{x',h'}\Delta_s^{x',h'}\right| \le \left|A_s^{x,h}\right|\left|\Delta_s^{x,h} - \Delta_s^{x',h'}\right| + \left|\Delta_s^{x',h'}\right|\left|A_s^{x,h} - A_s^{x',h'}\right|.$$

First note that due to (7.8) we have for some constants $C_1, C_2 \dots$

$$\int_0^T \left|A_s^{x,h}\right|\left|\Delta_s^{x,h} - \Delta_s^{x',h'}\right| ds \le C_1 \left(\int_0^T \left(1+\left|Z_s^x\right|\right)^2 ds\right)^{\frac{1}{2}} \left(\int_0^T \left|\Delta_s^{x,h} - \Delta_s^{x',h'}\right|^2 ds\right)^{\frac{1}{2}},$$

which implies, together with Lemma 7.1 and standard estimates of differences of the $\Delta^{x,h}$ (see Theorem 3.3 in Kunita 2004),

$$\begin{aligned} E\left(\int_0^T \left|A_s^{x,h}\right|\left|\Delta_s^{x,h} - \Delta_s^{x',h'}\right| ds\right)^{2pq} &\le C_2 \left(E\left(\int_0^T \left|\Delta_s^{x,h} - \Delta_s^{x',h'}\right|^2 ds\right)^{2pq}\right)^{\frac{1}{2}} \\ &\le C_3(|x-x'|^2 + |h-h'|^2)^{pq}. \end{aligned}$$

Second, from (7.11) we obtain

$$\begin{aligned} &\int_0^T \left|\Delta_s^{x',h'}\right|\left|A_s^{x,h} - A_s^{x',h'}\right| ds \\ &\quad \le C_4 \int_0^T \left(1+\left|Z_s^x\right|+\left|Z_s^{x'}\right|\right)\left(\left|X_s^x - X_s^{x'}\right| + \left|X_s^{x+he_i} - X_s^{x'+h'e_i}\right| + \left|Z_s^x - Z_s^{x'}\right|\right) ds \end{aligned}$$

and hence, with Lemma 7.3 and the moment estimates for X^x,

$$\begin{aligned} &E\left(\int_0^T \left|\Delta_s^{x',h'}\right|\left|A_s^{x,h} - A_s^{x',h'}\right| ds\right)^{2pq} \\ &\quad \le C_5 \left(E\left(\int_0^T \left(\left|X_s^x - X_s^{x'}\right| + \left|X_s^{x+he_i} - X_s^{x'+h'e_i}\right| + \left|Z_s^x - Z_s^{x'}\right|\right)^2 ds\right)^{2pq}\right)^{\frac{1}{2}} \\ &\quad \le C_6(|x-x'|^2 + |h-h'|^2)^{pq}. \end{aligned}$$

Combining the estimates just derived, we conclude

$$E\left[\sup_{t\in[0,T]} \left|U_t^{x,h} - U_t^{x',h'}\right|^{2p}\right] \le C(|x-x'|^2 + |h-h'|^2)^p.$$

This completes the proof of the lemma. □

Proof of Theorem 7.4. Note that by Lemma 7.3 we may assume that $(t,x) \mapsto Y_t^x$ is continuous for all ω. Then $U^{x,h}$ has continuous paths for all $x \in \mathbb{R}^m$ and $h \neq 0$.

Let $\mathcal{Q}$ be the collection of all pairs (x, h) where x is a vector of dyadic rationals in $\mathbb{R}^m$ and $h \neq 0$ a dyadic rational in $\mathbb{R}$. Since inequality (7.13) is valid, Kolmogorov's continuity criterion implies that there exists a null set N such that for all $\omega \in N^c$ the function $\mathcal{Q} \ni (x, h) \mapsto U^{x,h}$ can be uniquely extended to a continuous function from $\mathbb{R}^{m+1}$ into the space of continuous functions endowed with the sup norm (see Thm 74 or 75, Ch. IV, Protter 2004). Such a null set N can be chosen for any direction i in which we differentiate, and hence there exists a modification of Y^x such that for all t the mapping $x \mapsto Y_t^x$ possesses continuous partial derivatives.

Finally it is straightforward to show that the derivative $\nabla_x Y^x$ together with a process $\nabla_x Z^x$, defined as an $\mathcal{H}^2$ limit of the processes $V^{x,h}$ as $h \to 0$, solve the BSDE (7.12). □

7.3. The Markov Property of FBSDE

The forward part of our FBSDE (7.4) is solved by a time inhomogeneous Markov process. We will now investigate the consequences of this fact in more detail. Let us fix an initial time $t \in [0, T)$, as well as an initial state x to be taken by our forward process at this time. Then, conditioned on taking the value x at time t, the forward process satisfies the SDE

$$X_s^{t,x} = x + \int_t^s b\big(r, X_r^{t,x}\big)\, dr + \int_t^s \rho\big(r, X_r^{t,x}\big)\, dW_r, \tag{7.14}$$

where $x \in \mathbb{R}^m$ and $s \in [t, T]$. We will assume that the coefficients satisfy a growth and a Lipschitz condition. More precisely, assume that there exists a constant $C \in \mathbb{R}_+$ such that for all $x,\ x' \in \mathbb{R}^m$ and $t \in [0, T]$

$$\begin{aligned} |b(t, x) - b(t, x')| + |\rho(t, x) - \rho(t, x')| &\leq C(|x - x'|), \\ |b(t, x)| + |\rho(t, x)| &\leq C(1 + |x|). \end{aligned} \tag{7.15}$$

Condition (7.15) guarantees that there exists a unique solution of (7.14). It moreover implies that $X_r^{t,x}$ is Malliavin differentiable and that its Malliavin gradient has a representation involving, for (t, x) fixed, the global flow on the space of nonsingular linear operators $\Phi^{t,x}$ on $\mathbb{R}^m$ defined by the equation

$$\Phi_s^{t,x} = 1_{\mathbb{R}^m} + \int_t^s \nabla_x b\big(u, X_u^{t,x}\big)\Phi_u^{t,x}\, du + \int_t^s \nabla_x \rho\big(u, X_u^{t,x}\big)\Phi_u^{t,x}\, dW_u, \quad s \geq t.$$

Here $\nabla_x b$ and $\nabla_x \rho$ describe the gradients of b, respectively, ρ existing in the weak sense under (7.15), $1_{\mathbb{R}^m}$ the $m \times m$ unit matrix. The Malliavin gradient is then given by the formula (see Nualart 1995, p. 126)

$$D_\vartheta X_s^{t,x} = \Phi_s^{t,x}\big(\Phi_\vartheta^{t,x}\big)^{-1} \rho\big(\vartheta, X_\vartheta^{t,x}\big), \quad t \leq \vartheta \leq s. \tag{7.16}$$

With the Markov process $X^{t,x}$ starting at time t in x in mind, we now consider BSDE of the form

$$Y_s^{t,x} = F\big(X_T^{t,x}\big) - \int_s^T Z_r^{t,x}\, dW_r + \int_s^T f\big(r, X_r^{t,x}, Z_r^{t,x}\big)\, dr. \tag{7.17}$$

In accordance with Section 3, we now assume that the generator is a *deterministic* Borel measurable function $f : [0, T] \times \mathbb{R}^m \times \mathbb{R}^d \to \mathbb{R}$. Again we assume that f is differentiable in (x, z) and that there exists a constant $M \in \mathbb{R}_+$ such that for all $(t, x, z) \in [0, T] \times \mathbb{R}^m \times \mathbb{R}^d$ we have

$$(7.18) \qquad |f(t, x, z)| \leq M(1 + |z|^2) \text{ a.s.} \qquad \text{and} \qquad |\nabla_z f(t, x, z)| \leq M(1 + |z|) \text{ a.s.}$$

for all $(t, z) \in [0, T] \times \mathbb{R}^m$. If F is bounded, then it follows from Theorem 2.3 and 2.6 in Kobylanski (2000) that there exists a unique solution $(Y^{t,x}, Z^{t,x}) \in \mathcal{S}^\infty(\mathbb{R}) \otimes \mathcal{H}^2(\mathbb{R}^d)$ of the BSDE (7.17). The next result states that the solution of the BSDE is already determined by the forward process $X^{t,x}$. In order to formulate it, for all $m \in \mathbb{N}$ we denote by $\mathcal{D}^m$ the σ-algebra on $\mathbb{R}^m$ generated by the family of functions $\mathbb{R}^m \ni x \mapsto E \int_t^T \varphi(s, X_s^{t,x})\, ds$, where $t \in [0, T]$ and $\varphi : [0, T] \times \mathbb{R}^m \to \mathbb{R}$ is bounded and continuous.

THEOREM 7.6. *Let $F : \mathbb{R}^m \to \mathbb{R}$ be a bounded Borel function, suppose that f satisfies (7.18) and the coefficients of the forward diffusion (7.15). Suppose that there exist functions $f_n : [0, T] \times \mathbb{R}^m \times \mathbb{R}^d \to \mathbb{R}$, globally Lipschitz continuous in (x, z), such that for almost all ω and for all compact sets $K \subset \mathbb{R}^m \times \mathbb{R}^d$ the sequence f_n converges to f uniformly on $[0, T] \times K$. Then there exist two $\mathcal{B}[0, T] \otimes \mathcal{D}^m$- and $\mathcal{B}[0, T] \otimes \mathcal{D}^m$-measurable deterministic functions u and v on $[0, T] \times \mathbb{R}^m$ such that*

$$(7.19) \qquad Y_s^{t,x} = u(s, X_s^{t,x}) \quad \textit{and} \quad Z_s^{t,x} = v(s, X_s^{t,x})\rho(s, X_s^{t,x}),$$

for $P \otimes \lambda$-a.a. $(\omega, s) \in \Omega \times [t, T]$.

Proof. Let f^n be Lipschitz continuous in (x, z) such that f^n converges locally uniformly on $\mathbb{R}_+ \times \mathbb{R}^m \times \mathbb{R}^d$. Let $(t, x) \in [0, T] \times \mathbb{R}^m$ and denote by $(Y^n, Z^n) = ((Y^n)^{t,x}, (Z^n)^{t,x})$ the solution of the BSDE with generator f^n and terminal condition $\xi = F(X_T^{t,x})$. It follows from Theorem 2.8 in Kobylanski (2000) that Y^n converges to $Y^{t,x}$ in $\mathcal{H}^\infty(\mathbb{R})$, and Z^n converges to $Z^{t,x}$ in $\mathcal{H}^2(\mathbb{R}^d)$. By taking a subsequence if necessary, we may assume that Z^n converges to $Z^{t,x}$ a.s. on $\Omega \times [0, T]$.

According to Theorem 4.1 in El Karoui, Peng, and Quenez (1997), there exist $\mathcal{B}[0, T] \otimes \mathcal{D}^m$- and $\mathcal{B}[0, T] \otimes \mathcal{D}^m$-measurable deterministic functions $u_n(t, x)$ and $v_n(t, x)$ that satisfy the representations $Y_s^n = u_n(s, X_s^{t,x})$ and $Z_s^n = v_n(s, X_s^{t,x})\rho(s, X_s^{t,x})$ for all $s \in [t, T]$ a.s. Now define

$$u(t, x) = \liminf_n u_n(t, x) \qquad \text{and} \qquad v(t, x) = \liminf_n v_n(t, x).$$

Then $Y_s^{t,x} = u(s, X_s^{t,x})$ and $Z_s^{t,x} = v(s, X_s^{t,x})\rho(s, X_s^{t,x})$, a.s. □

By combining Theorem 7.6 with Theorem 7.4 we obtain a representation of the control process $Z^{t,x}$ in terms of the derivative of $Y^{t,x}$ with respect to x.

THEOREM 7.7. *Suppose that the assumptions of Theorem 7.6 are satisfied. Besides assume that $\nabla_z f$ is globally Lipschitz continuous, that (7.8) and (7.11) are satisfied, and further that the forward coefficients satisfy the stronger conditions (D1) and (D2). Then $u(t, x)$ is differentiable in x for a.a. $t \in [0, T]$. Moreover,*

$$(7.20) \qquad Z_s^{t,x} = \nabla_x u(t, X_s^{t,x})\rho(s, X_s^{t,x}),$$

for $P \otimes \lambda$-a.a. $(\omega, s) \in \Omega \times [t, T]$.

Proof. Recall that $X_s^{t,x}$ is Malliavin differentiable and that the assumptions of Lemma 7.3 are satisfied. Equation (7.9) implies that $x \mapsto u(t,x) = Y_t^{t,x}$ is Lipschitz continuous. Therefore $Y_s^{t,x} = u(s, X_s^{t,x})$ is Malliavin differentiable (see Proposition 1.2.2, Nualart 1995). By Theorem 7.4, $u(t,x)$ is differentiable in x, and by the chain rule we have $D_\vartheta Y_s^{t,x} = \nabla_x u(s, X_s^{t,x})\, D_\vartheta X_s^{t,x}$. Since due to (7.16) $D_s X_s^{t,x} = \rho(s, X_s^{t,x})$ and $Z_s^{t,x} = D_s Y_s^{t,x}$ (the later following f.ex. from Lemma 5.1 in El Karoui et al. 1997), Theorem 7.6 implies (7.20). □

7.4. Differentiability of Quadratic BSDE with Parameterized Terminal Condition

For this subsection we pass to a more abstract parameter dependence of the solution of a BSDE than studied above in a pair of forward and backward SDE. We consider the BSDE

$$Y_t^x = \xi(x) - \int_t^T Z_s^x \, dW_s + \int_t^T f\big(s, Z_s^x\big)\, ds, \quad t \in [0,T], \quad x \in \mathbb{R}^m. \tag{7.21}$$

Throughout we assume that

(E1) $\mathbb{R}^m \ni x \mapsto \xi(x) \in \mathbb{R}$ is a bounded random field which as a function of x is differentiable with bounded partial derivatives; $\nabla\xi(x)$ is also Lipschitz in x; also $f(t,0)$ is $(\mathcal{F}_t)$−adapted and satisfies $f(t,0) \in L^p$ for all $p \geq 1$.

(E2) there exists $M \in \mathbb{R}_+$ such that $|f(t,z)| \leq M(1+|z|^2)$ a.s.; f is differentiable in z such that $|\nabla_z f(t,z)| \leq M(1+|z|)$ for all $(t,z) \in [0,T] \times \mathbb{R}^d$ a.s.

(E3) for all $x \in \mathbb{R}^m$, $h \neq 0$ and $i \in \{1,\dots,m\}$, let $\zeta^{x,h,i} = \frac{1}{h}(\xi(x+he_i) - \xi(x))$. Then for every $p > 1$ there exists a $C > 0$, dependent only on p, such that for all $x, x' \in \mathbb{R}^m$ and $h, h' \neq 0$,

$$E[|\zeta^{x,h,i} - \zeta^{x',h',i}|^{2p}] \leq C(|x-x'|^2 + |h-h'|^2)^p. \tag{7.22}$$

Although the terminal condition does not depend on a forward diffusion (see Lemma 7.2 for a derivation of (7.22) in a FBSDE setting), Hypothesis (E1)-(E3) allow to apply the methods we used in Subsection 7.2 and obtain the following theorem.

THEOREM 7.8. *Let (E1), (E2), and (E3) be satisfied. Then there exists a function* $\Omega \times [0,T] \times \mathbb{R}^m \to \mathbb{R}^{1+d}$, $(\omega, t, x) \mapsto (Y_t^x, Z_t^x)(\omega)$, *such that for almost all* ω, *the process* Y_t^x *is continuous in t and continuously differentiable in x, and for all x,* (Y_t^x, Z_t^x) *is a solution of BSDE (7.21). Moreover, there exists a process* $\nabla_x Z^x \in \mathcal{H}^2(\mathbb{R}^{m\times d})$ *such that the pair* $(\nabla_x Y^x, \nabla_x Z^x)$ *solves the BSDE*

$$\nabla_x Y_t^x = \nabla_x \xi(x) - \int_t^T \nabla_x Z_s^x \, dW_s + \int_t^T \big[\nabla_z f\big(s, Z_s^x\big) \nabla_x Z_s^x\big]\, ds.$$

Proof. Conditions (E1) and (E3) guarantee that the solutions of the BSDE (7.21) exist and $(Y^x, Z^x) \in \mathcal{S}^\infty(\mathbb{R}) \otimes \mathcal{H}^2(\mathbb{R}^d)$.

Condition (E1), (E2), (E3), and the BMO property of the martingale $\int_0^\cdot Z^x dW$ allow us to prove moment estimates that correspond to Lemma 7.1, Lemma 7.3, and Lemma 7.5. Hence a simple adaptation of the proof of Theorem 7.4 provides the proof of Theorem 7.8. □

REFERENCES

ANKIRCHNER, S., P. IMKELLER, and G. DOS REIS (2007): Classical and Variational Differentiability of BSDEs with Quadratic Growth, *Electron. J. Probab.* 12(53), 1418–1453 (electronic).

ANKIRCHNER, S., P. IMKELLER, and A. POPIER (2008): Optimal Cross Hedging of Insurance Derivatives, *Stochastic Anal. Appl.* 26(4), 679–709.

BRIAND, P., and F. CONFORTOLA (2008): BSDEs with Stochastic Lipschitz Condition and Quadratic PDEs in Hilbert Spaces, *Stochastic Process. Appl*. 118(5), 818–838.

CARMONA, R., and V. DURRLEMAN (2003): Pricing and Hedging Spread Options, *SIAM Rev.* 45(4), 627–685 (electronic).

DAVIS, M. H. A. (2001): Pricing Weather Derivatives by Marginal Value, *Quant. Finance* 1(3), 305–308.

DAVIS, M. H. A. (2006): Optimal Hedging with Basis Risk, in *From Stochastic Calculus to Mathematical Finance*, pp. 169–187. Berlin: Springer.

DOS REIS, G. (2010): On Some Properties of Solutions of Quadratic Growth BSDE and Applications in Finance and Insurance. PhD Thesis, Humboldt University in Berlin.

EL KAROUI, N., S. PENG, and M. C. QUENEZ (1997): Backward Stochastic Differential Equations in Finance, *Math. Finance* 7(1), 1–71.

HENDERSON, V. (2002): Valuation of Claims on Nontraded Assets Using Utility Maximization, *Math. Finance* 12(4), 351–373.

HENDERSON, V., and D. HOBSON (2002): Real Options with Constant Relative Risk Aversion, *J. Econom. Dynam. Control* 27(2), 329–355.

HU, Y., P. IMKELLER, and M. MÜLLER (2005): Utility Maximization in Incomplete Markets, *Ann. Appl. Probab.* 15(3), 1691–1712.

KAZAMAKI, N. (1994): Continuous Exponential Martingales and BMO, *Lecture Notes in Mathematics* , Vol. 1579, Berlin: Springer-Verlag.

KOBYLANSKI, M. (2000): Backward Stochastic Differential Equations and Partial Differential Equations with Quadratic Growth, *Ann. Probab.* 28(2), 558–602.

KUNITA, H. (1990): Stochastic Flows and Stochastic Differential Equations, *Cambridge Studies in Advanced Mathematics*, Vol. 24, Cambridge: Cambridge University Press.

KUNITA, H. (2004): Stochastic Differential Equations Based on Lévy Processes and Stochastic Flows of Diffeomorphisms, in *Real and Stochastic Analysis*, M. M. Rao, ed., pp. x+405., Boston, MA: Birkhäuser Boston Inc.

MONOYIOS, M. (2004): Performance of Utility-Based Strategies for Hedging Basis Risk, *Quant. Finance* 4(3), 245–255.

MUSIELA, M., and T. ZARIPHOPOULOU (2004): An Example of Indifference Prices Under Exponential Preferences, *Finance Stoch.* 8(2), 229–239.

NUALART, D. (1995): The Malliavin Calculus and Related Topics, *Probability and Its Applications*. New York: Springer-Verlag.

PROTTER, P. (2004): Stochastic Integration and Differential Equations, 2nd ed. *Applications of Mathematics*, Vol. 21 of *Stochastic Modelling and Applied Probability*, Berlin: Springer-Verlag.

ROUGE, R., and N. EL KAROUI (2000): Pricing via Utility Maximization and Entropy, *Math. Finance* 10(2), 259–276. INFORMS Applied Probability Conference (Ulm, 1999).