DOI: 10.1214/EJP.v12-462

Electronic Journal of Probability

Vol. 12 (2007), Paper no. , pages 1418–1453.

Journal URL
http://www.math.washington.edu/~ejpecp/

Classical and Variational Differentiability of BSDEs with Quadratic Growth*

Stefan Ankirchner and Peter Imkeller and Gonçalo Dos Reis
Institut für Mathematik
Humboldt-Universität zu Berlin
Unter den Linden 6, 10099 Berlin, Germany
ankirchn@mathematik.hu-berlin.de imkeller@mathematik.hu-berlin.de gnreis@mathematik.hu-berlin.de

Abstract

We consider Backward Stochastic Differential Equations (BSDEs) with generators that grow quadratically in the control variable. In a more abstract setting, we first allow both the terminal condition and the generator to depend on a vector parameter x. We give sufficient conditions for the solution pair of the BSDE to be differentiable in x. These results can be applied to systems of forward-backward SDE. If the terminal condition of the BSDE is given by a sufficiently smooth function of the terminal value of a forward SDE, then its solution pair is differentiable with respect to the initial vector of the forward equation. Finally we prove sufficient conditions for solutions of quadratic BSDEs to be differentiable in the variational sense (Malliavin differentiable).

Key words: BSDE, forward-backward SDE, quadratic growth, differentiability, stochastic calculus of variations, Malliavin calculus, Feynman-Kac formula, BMO martingale, reverse Hölder inequality.

AMS 2000 Subject Classification: Primary 60H10; Secondary: 60H07, 65C30.

Submitted to EJP on January 31, 2007, final version accepted November 5, 2007.

*This research was supported by the DFG Research Center MATHEON "Mathematics for Key Technologies" (FZT86) in Berlin.

Introduction

Problems of stochastic control treated by the crucial tool of *backward stochastic differential equations (BSDEs)* have been encountered in many areas of application of mathematics in recent years. A particularly important area is focused around optimal hedging problems for contingent claims in models of financial markets. Recently, a special class of hedging problems in incomplete financial markets has been considered in the area where finance and insurance concepts meet. At this interface problems of *securitization* arise, i.e. insurance risk is transferred to capital markets. One particularly interesting risk source is given by climate or environmental hazards affecting insurance companies or big branches of the economy that depend on weather such as agriculture and fishing, transportation and tourism. The public awareness of climate hazards such as floods or hurricanes is continually increasing with the intensity of the discussion about irreversible changes due to human impact.

BSDEs typically appear in the following setting. On a financial market some small investors are subject to an external risk source described for instance by weather or climate influences. There may also be big investors such as re-insurance companies that depend in a possibly different way on the same risk source. In this situation market incompleteness stems from the external risk not hedgeable by the market assets. One may complete the market either by making the external risk tradable through the introduction of an insurance asset traded among small agents, or by introducing a risk bond issued by a big agent. In this setting, treating the utility maximization problem for the agents under an equilibrium condition describing basically market clearing for the additional assets, leads to the determination of the market price of external risk through a BSDE which in case of exponential utility turns out to be quadratic in the control variable (see [7], [3] and [4]). Alternatively, instead of maximizing utility with respect to exponential utility functions we might minimize risk measured by the *entropic risk measure*. In this setting we again encounter a BSDE with quadratic nonlinearity, of the type

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

where W is a finite-dimensional Wiener process of the same dimension as the control process Z, with a generator f that depends at most quadratically on Z, and a bounded terminal condition ξ. In the meantime, the big number of papers published on general BSDEs is rivalled by the number of papers on BSDEs of this type of nonlinearity. For a more complete list of references see [5] or [9]. In particular, there are papers in which the boundedness condition on ξ is relaxed to an exponential integrability assumption, or where the stochastic integral process of Z is supposed to be a BMO martingale.

In a particularly interesting case the terminal variable ξ is given by a function $g(X_T^x)$ at terminal time T of the solution process X of a forward SDE

$$X_t^x = x + \int_0^t b(s, X_s^x)ds + \int_0^t \sigma(s, X_s^x)dW_s, \quad 0 \le t \le T,$$

with initial vector $x \in \mathbb{R}$. Similarly, the driver f may depend on the diffusion dynamics of X^x. Via the famous link given by the generalized Feynman-Kac formula, systems as the above of forward-backward stochastic differential equations are seen to yield a stochastic access to solve nonlinear PDE in the viscosity sense, see [9].

In this context, questions related to the regularity of the solutions (X^x, Y^x, Z^x) of the stochastic forward-backward system in the classical sense with respect to the initial vector x or in the sense of the stochastic calculus of variations (Malliavin calculus) are frequently encountered. Equally, from a more analytic point of view also questions of smoothness of the viscosity solutions of the PDE associated via the Feynman-Kac link are seen to be very relevant.

For instance, Horst and Müller (see [7]) ask for existence, uniqueness and regularity of a global classical solution of our PDE from the *analytic point of view*. Not attempting a systematic approach of the problem, they use the natural access of the problem by asking for smoothness of the solutions of the stochastic system in terms of the stochastic calculus of variations. But subsequently they work under the restrictive condition that the solutions of the BSDE have bounded variational derivatives, which is guaranteed only under very restrictive assumptions on the coefficients.

The question of smoothness of the *stochastic solutions* in the parameter x arises for instance in an approach of cross hedging of environmental risks in [1]. Here the setting is roughly the one of an incomplete market generated by a number of big and small agents subject to an external (e.g. climate related) risk source, and able to invest in a given capital market. The risk exposure of different types of agents may be negatively correlated, so that typically one type profits from the risky event, while at the same time the other type suffers. Therefore the concept of hedging one type's risk by transferring it to the agents of the other type in a *cross hedging* context makes sense. Mathematically, in the same way as described above, it leads to a BSDE of the quadratic type, the solution (Y^x, Z^x) of which depends on the initial vector x of a forward equation with solution X^x. Under certain assumptions, the cross-hedging strategy can be explicitly given in a formula depending crucially on x, and in which the sensitivity with respect to x describes interesting quality properties of the strategy.

In this paper, we tackle regularity properties of the solutions (Y^x, Z^x) of BSDEs of the quadratic type such as the two previously sketched in a systematic and thorough way. Firstly, the particular dependence on the starting vector x of the forward component of a forward-backward system will be generalized to the setting of a terminal condition $\xi(x)$ depending in a smooth way to be specified on some vector x in a certain Euclidean state space. We both consider the smoothness with respect to x in the classical sense, as well as the smoothness in the sense of Malliavin's calculus.

The common pattern of reasoning in order to tackle smoothness properties of any kind starts with a priori estimates for difference and differential quotients, or for infinite dimensional gradients in the sense of variational calculus. In the estimates, these quantities are related to corresponding difference and differential quotients or Malliavin gradients of the terminal variable and the driver. To obtain the a priori estimates, we make use to changes of probability of the Girsanov type, by which essentially nonlinear parts of the driver are eliminated. Since terminal conditions in our treatment are usually bounded, the exponential densities in these measure changes are related to BMO martingales. Known results about the inverse Hölder inequality allow to show that as a consequence the exponential densities are r-integrable for some $r > 1$ related to the BMO norm. This way we are able to reduce integrability properties for the quantities to be estimated to a natural level. In a second step, the a priori inequalities are used to derive the desired smoothness properties from corresponding properties of driver and terminal condition. To the best of our knowledge, only Malliavin differentiability results of this type have been obtained so far, with strong conditions on the coefficients restricting generality considerably (see [7]).

After finishing this work we found out that there exists a paper by Briand and Confortola (see [2]) with related results based on similar techniques. The two studies were carried out simultaneously and in a completely independent way.

The paper is organized as follows. In section 1 we fix the notation and recall some process properties needed in the proofs of the main body of the paper. Section 2 contains the main results on classical differentiability. In sections 3, 4 and 5 we give a priori bounds for classes of non-linear BSDEs. Section 6 contains the proofs of the theorems stated in Section 2. Section 7 is devoted to the application of the proven results to the forward-backward SDE setting. In Section 8 we state and prove the Malliavin differentiability results.

1 Preliminaries

Throughout this paper let $(\Omega, \mathcal{F}, P)$ be a complete probability space and $W = (W_t)_{t\geq 0}$ a $d-$dimensional Brownian motion. Let $\{\mathcal{F}_t\}_{t\geq 0}$ denote the natural filtration generated by W, augmented by the $P-$null sets of $\mathcal{F}$.

Let $T > 0$, ξ be an $\mathcal{F}_T$-measurable random variable and $f : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$. We will consider Backward Stochastic Differential Equations (BSDEs) of the form

$$Y_t = \xi + \int_t^T f(t, Y_t, Z_t)\mathrm{d}t - \int_t^T Z_t \mathrm{d}W_t. \tag{1}$$

As usual we will call ξ the *terminal condition* and the function f the *generator* of the BSDE (1). A *solution* consists of a pair (Y, Z) of adapted processes such that (1) is satisfied. To be correct we should write $\int_t^T \langle Z_t, \mathrm{d}W_t\rangle$ or $\sum_{i=1}^d \int_t^T Z_s^i dW_s^i$ instead of $\int_t^T Z_t \mathrm{d}W_t$, since W and Z are $d-$dimensional vectors; but for simplicity we use this notation as it is without ambiguity. It is important to know which process spaces the solution of a BSDE belongs to. We therefore introduce the following notation for the spaces we will frequently use. Let $p \in [1, \infty]$. Then, for $m \in \mathbb{N}^*$

- $\mathbb{L}^p(\mathbb{R}^m)$ is the space of all progressively measurable processes $(X_t)_{t\in[0,T]}$ with values in $\mathbb{R}^m$ such that $\|X_t\|^p_{\mathbb{L}^p} = \mathbb{E}[\Big(\int_0^T |X_s|^2\mathrm{d}s\Big)^{p/2}] < \infty$.

- $\mathcal{R}^p(\mathbb{R}^m)$ is the space of all measurable processes $(X_t)_{t\in[0,T]}$ with values in $\mathbb{R}^m$ such that $\|X\|^p_{\mathcal{R}^p} = \mathbb{E}[\Big(\sup_{t\in[0,T]} |X_t|\Big)^p] < \infty$. Note that $\mathcal{R}^\infty(\mathbb{R}^m)$ is the space of bounded measurable processes.

- $H^p(\mathbb{R}^m)$ is the class of all local martingales X such that $\|X\|^p_{H^p} = \mathbb{E}^P[\langle X\rangle_T^{\frac{p}{2}}] < \infty$.

- $L^p(\mathbb{R}^m; P)$ is the space of $\mathcal{F}_T$-measurable random variables $X : \Omega \mapsto \mathbb{R}^m$ such that $\|X\|^p_{L^p} = \mathbb{E}^P[|X|^p] < \infty$. We will omit reference to the space or the measure when there is no ambiguity.

Furthermore, we use the notation $\partial_t = \frac{\partial}{\partial t}$, $\nabla = (\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_d})$ for $(t,x) \in [0,T] \times \mathbb{R}^d$.
Suppose that the generator satisfies, for $a \geq 0$ and $b, c > 0$

$$|f(t,x,y,z)| \leq a(1+b|y|) + \frac{c}{2}|z|^2. \tag{2}$$

Kobylanski has shown in [9] that if ξ is bounded and the generator f satisfies (2), then there exists a solution $(Y,Z) \in \mathcal{R}^\infty \times \mathbb{L}^2$. Moreover, it follows from the results in [11], that in this case the process Z is such that the stochastic integral process relative to the Brownian motion $\int_0^\cdot Z dW$ is a so-called Bounded Mean Oscillation (BMO) martingale.
Since the BMO property is crucial for the proofs we present in this paper we recall its definition and some of its basic properties. For an overview on BMO martingales see [8].

Definition 1.1 (BMO)**.** Let M be a uniformly integrable $(\mathcal{F}_t)$-martingale satisfying $M_0 = 0$. For $1 \leq p < \infty$ set

$$\|M\|_{BMO_p} = \sup_{\tau \text{ stopping time}} \Big(\mathbb{E}\Big[|M_\infty - M_\tau|^p | \mathcal{F}_\tau\Big]\Big)^{1/p}.$$

The normed linear space $\{M : \|M\|_{BMO_p} < \infty\}$ with norm $\|M\|_{BMO_p}$ is denoted by BMO_p. If we want to stress the measure P we are referring to we will write BMO(P).
It can be shown that for any p, $q \in [1,\infty]$ we have $\mathrm{BMO}_p = \mathrm{BMO}_q$ (see [8]). Therefore we will often omit the index and simply write BMO for the set of BMO martingales.

In the following Lemma we state the properties of BMO martingales we will frequently use.

Lemma 1.2 (Properties of BMO martingales)**.**

1) *Given a BMO martingale M with quadratic variation $\langle M\rangle$, its stochastic exponential*
$$\mathcal{E}(M)_T = \exp\{M_T - \frac{1}{2}\langle M\rangle_T\}$$
has integral 1*, and thus the measure defined by $dQ = \mathcal{E}(M)_T dP$ is a probability measure.*

2) *Let M be a BMO martingale relative to the measure P. Then the process $\hat{M} = M - \langle M\rangle$ is a BMO martingale relative to the measure Q (see Theorem 3.3 in* [8]*).*

3) *For any BMO Martingale, it is always possible to find a $p > 1$ such that $\mathcal{E}(M) \in L^p$, i.e. if $\|M\|_{BMO_2} < \Psi(p)$, then $\mathcal{E}(M) \in L^p$ (see for example Theorem 3.1* [8]*). Where $\Psi(x) = \left\{1 + \frac{1}{x^2}\log\frac{2x-1}{2(x-1)}\right\}^{\frac{1}{2}} - 1$ for all $1 < x < \infty$ and verifies $\lim_{x\to 1^+}\Psi(x) = \infty$ and $\lim_{x\to\infty}\Psi(x) = 0$.*

2 Differentiability of quadratic BSDEs in the classical sense

Suppose that the terminal condition and the generator of a quadratic BSDE depend on the Euclidean parameter set $\mathbb{R}^n$ for some $n \in \mathbb{N}^*$. We will show that the smoothness of the terminal condition and the generator is transferred to the solution of the BSDE

$$Y_t^x = \xi(x) - \int_t^T Z_s^x \mathrm{d}W_s + \int_t^T f(s,x,Y_s^x,Z_s^x)\mathrm{d}s, \quad x \in \mathbb{R}^n, \tag{3}$$

where terminal condition and generator are subject to the following conditions

(C1) $f:\Omega\times[0,T]\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^d\to\mathbb{R}$ is an adapted measurable function such that $f(\omega,t,x,y,z)=l(\omega,t,x,y,z)+\alpha|z|^2$, where $l(\omega,t,x,y,z)$ is globally Lipschitz in (y,z) and continuously differentiable in (x,y,z); for all $r\geq 1$ and (t,y,z) the mapping $\mathbb{R}^d\to L^r$, $x\mapsto l(\omega,t,x,y,z)$ is differentiable and for all $x\in\mathbb{R}^n$

$$\lim_{x'\to x}\mathbb{E}^P\Big[\Big(\int_0^T|l(s,x',Y_s^x,Z_s^x)-l(s,x,Y_s^x,Z_s^x)|\mathrm{d}s\Big)^r\Big]=0 \qquad \text{and}$$
$$\lim_{x'\to x}\mathbb{E}^P\Big[\Big(\int_0^T|\frac{\partial}{\partial x}l(s,x',Y_s^{x'},Z_s^{x'})-\frac{\partial}{\partial x}l(s,x,Y_s^x,Z_s^x)|\mathrm{d}s\Big)^r\Big]=0,$$

(C2) the random variables $\xi(x)$ are $\mathcal{F}_T-$adapted and for every compact set $K\subset\mathbb{R}^n$ there exists a constant $c\in\mathbb{R}$ such that $\sup_{x\in K}\|\xi(x)\|_\infty\leq c$; for all $p\geq 1$ the mapping $\mathbb{R}^n\to L^p$, $x\mapsto\xi(x)$ is differentiable with derivative $\nabla\xi$.

If (C1) and (C2) are satisfied, then there exists a unique solution (Y^x,Z^x) of Equation (3). This follows from Theorems 2.3 and 2.6 in [9]. We will establish two differentiability results for the pair (Y^x,Z^x) in the variable x. We first consider differentiability of the vector valued map

$$x\mapsto(Y^x,Z^x)$$

with respect to the Banach space topology defined on $\mathcal{R}^p(\mathbb{R}^1)\times\mathbb{L}^p(\mathbb{R}^d)$. This will be stated in Theorem 2.1. A slightly more stringent result will be obtained in the subsequent Theorem 2.2. Here, we consider pathwise differentiability of the maps

$$x\mapsto(Y_t^x(\omega),Z_t^x(\omega))$$

in the usual sense, for almost all pairs (ω,t). In both cases, the derivatives will be identified with $(\nabla Y^x,\nabla Z^x)$ solving the BSDE

$$\begin{aligned}\nabla Y_t^x &= \nabla\xi(x)-\textstyle\int_t^T\nabla Z_s^x\mathrm{d}W_s\\ &\quad+\textstyle\int_t^T\left[\partial_x l(s,x,Y_s^x,Z_s^x)+\partial_y l(s,x,Y_s^x,Z_s^x)\nabla Y_s^x+\partial_z l(s,x,Y_s^x,Z_s^x)\nabla Z_s^x+2\alpha Z_s^x\nabla Z_s^x\right]\mathrm{d}s.\end{aligned}\tag{4}$$

We emphasize at this place that it is not immediate that this BSDE possesses a solution. In fact, without considering it as a component of a system of BSDEs also containing the original quadratic one, it can only be seen as a linear BSDE with global, but random (and not bounded) Lipschitz constants.

Theorem 2.1. *Assume (C1) and (C2). Then for all $p\geq 1$, the function $\mathbb{R}^n\to\mathcal{R}^p(\mathbb{R}^1)\times\mathbb{L}^p(\mathbb{R}^d)$, $x\mapsto(Y^x,Z^x)$, is differentiable, and the derivative is a solution of the BSDE (4).*

Under slightly stronger conditions one can show the existence of a modification of Y^x which is P-a.s. differentiable as a mapping from $\mathbb{R}^n$ to $\mathbb{R}$. Let $e_i=(0,\dots,1,\dots,0)$ be the unit vector in $\mathbb{R}^n$ where the ith component is 1 and all the other components 0. For $x\in\mathbb{R}^n$ and $h\neq 0$ let $\zeta(x,h,e_i)=\frac{1}{h}[\xi(x+he_i)-\xi(x)]$. For the existence of differentiable modifications we will assume that

(C3) for all $p\geq 1$ there exists a constant $C>0$ such that for all $i\in\{1,\dots,n\}$, x, $x'\in\mathbb{R}^n$ and $h,h'\in\mathbb{R}\setminus\{0\}$

$$\mathbb{E}\Big[|\xi(x+he_i)-\xi(x'+h'e_i)|^{2p}+|\zeta(x,h,e_i)-\zeta(x',h',e_i)|^{2p}\Big]\leq C(|x-x'|^2+|h-h'|^2)^p.$$

Theorem 2.2. *Suppose, in addition to the assumptions of Theorem 2.1, that (C3) is satisfied and that $l(t,x,y,z)$ and its derivatives are globally Lipschitz continuous in (x,y,z). Then there exists a function $\Omega \times [0,T] \times \mathbb{R}^n \to \mathbb{R}^{1+d}$, $(\omega,t,x) \mapsto (Y_t^x, Z_t^x)(\omega)$, such that for almost all ω, 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 (3).*

3 Moment estimates for linear BSDEs with stochastic Lipschitz generators

By formally deriving a quadratic BSDE with generator satisfying (C1) and (C2) we obtain a linear BSDE with a stochastic Lipschitz continuous generator. The Lipschitz constant depends on the second component of the solution of the original BSDE. In order to show differentiability, we start deriving a priori estimates for this type of linear BSDE with stochastic Lipschitz continuous generator. For this purpose, we first need to show that the moments of the solution can be effectively controlled. Therefore this section is devoted to moment estimates of solutions of BSDEs of the form

$$U_t = \zeta - \int_t^T V_s dW_s + \int_t^T [l(s,U_s,V_s) + H_s V_s + A_s]\, \mathrm{d}s. \tag{5}$$

We will make the following assumptions concerning the drivers:

(A1) For all $p \geq 1$, ζ is $\mathcal{F}_T$−adapted and we have $\zeta \in L^p(\mathbb{R}^1)$,

(A2) H is a predictable $\mathbb{R}^d$−valued process, integrable with respect to W, such that $\int H dW$ is a BMO-martingale,

(A3) $l : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ is such that for all (u,v), the process $l(\omega,t,u,v)$ is $(\mathcal{F}_t)$-predictable and there exists a constant $M > 0$ such that for all (ω,t,u,v),

$$|l(\omega,t,u,v)| \leq M(|u| + |v|),$$

(A4) A is a measurable adapted process such that for all $p \geq 1$ we have $\mathbb{E}[\big(\int_0^T |A_s|\mathrm{d}s\big)^p] < \infty$.

Moreover, we assume that (U,V) is a solution of (5) satisfying

(A5) $[\int_0^T U_s^2 |V_s|^2 \mathrm{d}s]^{\frac{1}{2}}$ and $\int_0^T |U_s A_s|\mathrm{d}s$ are p-integrable for all $p \geq 1$.

Under the assumptions (A1), (A2), (A3), (A4) and (A5) one obtains the following estimates.

Theorem 3.1 (Moment estimates)**.** *Assume that (A1)-(A5) are satisfied. Let $p > 1$ and $r > 1$ such that $\mathcal{E}(\int H dW)_T \in L^r(P)$. Then there exists a constant $C > 0$, depending only on p, T, M and the BMO-norm of $\int H dW$), such that with the conjugate exponent q of r we have*

$$\mathbb{E}^P\Big[\sup_{t\in[0,T]} |U_t|^{2p}\Big] + \mathbb{E}^P\Big[\Big(\int_0^T |V_s|^2 \mathrm{d}s\Big)^p\Big] \;\; \leq \;\; C\mathbb{E}^P\Big[|\zeta|^{2pq^2} + \Big(\int_0^T |A_s|\mathrm{d}s\Big)^{2pq^2}\Big]^{\frac{1}{q^2}}. \tag{6}$$

Moreover we have

$$\mathbb{E}^P[\int_0^T |U_s|^2 \mathrm{d}s] + \mathbb{E}^P[\int_0^T |V_s|^2 \mathrm{d}s] \;\; \leq \;\; C\mathbb{E}^P\Big[|\zeta|^{2q^2} + \Big(\int_0^T |A_s|^2 \mathrm{d}s\Big)^{q^2}\Big]^{\frac{1}{q^2}}. \tag{7}$$

The proof is divided into several steps. First let $\beta > 0$ and observe that by applying Itô's formula to $e^{\beta t}U_t^2$ we obtain

$$
\begin{aligned}
e^{\beta t}U_t^2 &= e^{\beta T}U_T^2 - 2\int_t^T e^{\beta s}U_sV_s\mathrm{d}W_s \\
&\quad + \int_t^T e^{\beta s}\big[-\beta U_s^2 + 2U_s\big(l(s,U_s,V_s) + H_sV_s + A_s\big) - |V_s|^2\big]\mathrm{d}s.
\end{aligned}
$$

By (A2), the auxiliary measure defined by $Q = \mathcal{E}(H\cdot W)_T \cdot P$ is in fact a probability measure. Then $\hat{W}_t = W_t - \int_0^t H_s\mathrm{d}s$ is a Q-Brownian motion, and

$$
\begin{aligned}
e^{\beta t}U_t^2 &\leq e^{\beta T}U_T^2 - 2\int_t^T e^{\beta s}U_sV_s\mathrm{d}\hat{W}_s \\
&\quad + \int_t^T e^{\beta s}\big[(-\beta + 2M)U_s^2 + 2M|U_s||V_s| - |V_s|^2 + |U_sA_s|\big]\mathrm{d}s
\end{aligned}
$$

By choosing $\beta = M^2 + 2M$, we obtain

$$
e^{\beta t}U_t^2 + \int_t^T e^{\beta s}(M|U_s| - |V_s|)^2\mathrm{d}s \leq e^{\beta T}U_T^2 - 2\int_t^T e^{\beta s}U_sV_s\mathrm{d}\hat{W}_s + \int_t^T e^{\beta s}|U_sA_s|\mathrm{d}s. \quad (8)
$$

We therefore first prove moment estimates under the measure Q.

Lemma 3.2. *For all $p > 1$ there exists a constant C, depending only on p, T and M, such that*

$$
\mathbb{E}^Q\Big[\sup_{t\in[0,T]}|U_t|^{2p}\Big] + \mathbb{E}^Q\left[\left(\int_0^T |V_s|^2\mathrm{d}s\right)^p\right] \leq C\mathbb{E}^Q\Big[|\zeta|^{2p} + \Big(\int_0^T |A_s|\mathrm{d}s\Big)^{2p}\Big]. \quad (9)
$$

Moreover we have

$$
\mathbb{E}^Q\Big[\int_0^T |U_s|^2\mathrm{d}s\Big] + \mathbb{E}^Q\left[\int_0^T |V_s|^2\mathrm{d}s\right] \leq C\mathbb{E}^Q\Big[|\zeta|^2 + \int_0^T |A_s|^2\mathrm{d}s\Big]. \quad (10)
$$

Proof. Throughout this proof let $C_1, C_2, \ldots,$ be constants depending only on p, T and M. Inequality (8) implies

$$
e^{\beta t}U_t^2 \leq e^{\beta T}U_T^2 - 2\int_t^T e^{\beta s}U_sV_s\mathrm{d}\hat{W}_s + \int_t^T e^{\beta s}|U_sA_s|\mathrm{d}s, \quad (11)
$$

and (A5) together with the existence of the rth moment for $\mathcal{E}(\int H dW)_T$ yield $\int_0^T U_s^2|V_s|^2\mathrm{d}s \in L^1(Q)$. Hence, since $e^{\beta t}U_t^2$ is $(\mathcal{F}_t)$-adapted,

$$
e^{\beta t}U_t^2 \leq e^{\beta T}\mathbb{E}^Q\Big[|\zeta|^2 + \int_t^T e^{\beta s}|U_sA_s|\mathrm{d}s|\mathcal{F}_t\Big]. \quad (12)
$$

Integrating both sides and using Young's inequality, we obtain

$$
\begin{aligned}
\mathbb{E}^Q[\int_0^T U_s^2\mathrm{d}s] &\leq C_1\mathbb{E}^Q[|\zeta|^2 + \int_0^T |U_sA_s|\mathrm{d}s] \\
&\leq C_1\mathbb{E}^Q[\zeta^2 + 2C_1\int_0^T |A_s|^2\mathrm{d}s] + \frac{1}{2}\mathbb{E}^Q[\int_0^T U_s^2\mathrm{d}s],
\end{aligned}
$$

and hence

$$\mathbb{E}^Q[\int_0^T U_s^2 \mathrm{d}s] \le C_2\mathbb{E}^Q[|\zeta|^2 + \int_0^T |A_s|^2\mathrm{d}s]. \tag{13}$$

Inequality (12), (A5) and Doob's L^p inequality imply for $p > 1$

$$\begin{aligned}\mathbb{E}^Q[\sup_{t\in[0,T]}|U_t|^{2p}] &\le C_3\mathbb{E}^Q\Big[\Big(|\zeta|^2 + \int_0^T |U_sA_s|\mathrm{d}s\Big)^p\Big]\\ &\le C_4\mathbb{E}^Q\Big[|\zeta|^{2p} + \Big(\sup_{t\in[0,T]}|U_t|\int_0^T |A_s|\mathrm{d}s\Big)^p\Big].\end{aligned}$$

By Young's inequality, $(\sup_{t\in[0,T]}|U_t|^p)(\int_0^T|A_s|\mathrm{d}s)^p \le \frac{1}{2C_4}\sup_{t\in[0,T]}|U_t|^{2p} + 2C_4(\int_0^T|A_s|\mathrm{d}s)^{2p}$, and hence

$$\mathbb{E}^Q[\sup_{t\in[0,T]}|U_t|^{2p}] \le C_5\mathbb{E}^Q\Big[|\zeta|^{2p} + \Big(\int_0^T |A_s|\mathrm{d}s\Big)^{2p}\Big]. \tag{14}$$

In order to complete the proof, note that (8) implies

$$\begin{aligned}&\int_t^T e^{\beta s}|V_s|^2\mathrm{d}s\\ &\quad\le e^{\beta T}U_T^2 - 2\int_t^T e^{\beta s}U_sV_s\mathrm{d}\hat{W}_s + 2\int_t^T e^{\beta s}M|U_s||V_s|\mathrm{d}s + \int_t^T e^{\beta s}|U_s||A_s|\mathrm{d}s.\end{aligned} \tag{15}$$

By Young's inequality, $2\int_t^T e^{\beta s}M|U_s||V_s|\mathrm{d}s \le \frac{1}{2}\int_t^T e^{\beta s}|V_s|^2\mathrm{d}s + 8M^2\int_t^T e^{\beta s}U_s^2\mathrm{d}s$, and hence

$$\begin{aligned}\frac{1}{2}\mathbb{E}^Q\Big[\int_0^T e^{\beta s}|V_s|^2\mathrm{d}s\Big] &\le \mathbb{E}^Q[e^{\beta T}U_T^2 + 8M^2\int_0^T e^{\beta s}U_s^2\mathrm{d}s + \int_0^T e^{\beta s}U_s^2 + e^{\beta s}|A_s|^2\mathrm{d}s\Big]\\ &\le C_6\mathbb{E}^Q[\zeta^2 + \int_0^T |A_s|^2\mathrm{d}s]\end{aligned}$$

which, combined with (13) leads to the desired Inequality (10).

Equation (15), Young's inequality, Doob's L^p-inequality and the Burkholder-Davis-Gundy inequality imply

$$\begin{aligned}&\mathbb{E}^Q\left[\left(\int_0^T e^{\beta s}|V_s|^2\mathrm{d}s\right)^p\right]\\ \le\;& C_7\mathbb{E}^Q\Big[|\zeta|^{2p} + \Big(T\sup_{t\in[0,T]}e^{\beta t}U_t^2\Big)^p + \Big(\int_0^T e^{\beta s}U_sV_s\mathrm{d}\hat{W}_s\Big)^p + \Big(\int_0^T e^{\beta s}|U_s||A_s|\mathrm{d}s\Big)^p\Big]\\ \le\;& C_8\mathbb{E}^Q\Big[|\zeta|^{2p} + \sup_{t\in[0,T]}e^{\beta tp}|U_t|^{2p} + \Big(\int_0^T e^{2\beta s}U_s^2|V_s|^2\mathrm{d}s\Big)^{\frac{p}{2}} + \sup_{t\in[0,T]}|U_t|^p\Big(\int_0^T e^{\beta s}|A_s|\mathrm{d}s\Big)^p\Big]\\ \le\;& C_8\mathbb{E}^Q\Big[|\zeta|^{2p} + \sup_{t\in[0,T]}e^{\beta tp}|U_t|^{2p} + \Big(\sup_{t\in[0,T]}e^{\beta t}U_t^2\Big)^{\frac{p}{2}}\Big(\int_0^T e^{\beta s}|V_s|^2\mathrm{d}s\Big)^{\frac{p}{2}}\\ &\quad + \sup_{t\in[0,T]}|U_t|^{2p} + \Big(\int_0^T e^{\beta s}|A_s|\mathrm{d}s\Big)^{2p}\Big]\end{aligned}$$

By Young's inequality,

$$\Big(\sup_{t\in[0,T]} e^{\beta t}U_t^2\Big)^{\frac{p}{2}}\Big(\int_0^T e^{\beta s}|V_s|^2\mathrm{d}s\Big)^{\frac{p}{2}} \le 2C_8\Big(\sup_{t\in[0,T]} e^{\beta t}U_t^2\Big)^p + \frac{1}{2C_8}\Big(\int_0^T e^{\beta s}|V_s|^2\mathrm{d}s\Big)^p,$$

which implies

$$\begin{aligned}\mathbb{E}^Q\left[\Big(\int_0^T e^{\beta s}|V_s|^2\mathrm{d}s\Big)^p\right] &\le C_9\mathbb{E}^Q\Big[|\zeta|^{2p} + \sup_{t\in[0,T]}|U_t|^{2p} + \Big(\int_0^T |A_s|\mathrm{d}s\Big)^{2p}\Big] \\ &\le C_{10}\mathbb{E}^Q\Big[|\zeta|^{2p} + \Big(\int_0^T |A_s|\mathrm{d}s\Big)^{2p}\Big].\end{aligned}$$

Thus, with Inequality (14), the proof is complete. □

Proof of Theorem 3.1. Notice that by the second statement of Lemma 1.2, the process $\int Hd\hat{W} = \int HdW - \int_0^\cdot H_s^2\mathrm{d}s$ belongs to BMO(Q), and hence $-\int Hd\hat{W}$ also. Moreover, $\mathcal{E}(\int HdW)^{-1} = \mathcal{E}(-\int Hd\hat{W})$. Consequently, by the third statement of Lemma 1.2, there exists an $r > 1$ such that $\mathcal{E}(H\cdot W)_T \in L^r(P)$ and $\mathcal{E}(H\cdot W)_T^{-1} \in L^r(Q)$. Throughout let $D = \max\{\|\mathcal{E}(H\cdot W)_T\|_{L^r(P)}, \|\mathcal{E}(H\cdot W)_T^{-1}\|_{L^r(Q)}\}$. Hölder's inequality and Lemma 3.2 imply that for the conjugate exponent q of r we have

$$\begin{aligned}\mathbb{E}^P[\sup_{s\in[0,T]}|U_s|^{2p}] &= \mathbb{E}^Q[\mathcal{E}(H\cdot W)_T^{-1}\sup_{s\in[0,T]}|U_s|^{2p}] \le D\mathbb{E}^Q[\sup_{s\in[0,T]}|U_s|^{2pq}]^{\frac{1}{q}} \\ &\le C_1\,D\,\mathbb{E}^Q\Big[|\zeta|^{2pq} + \Big(\int_0^T |A_s|\mathrm{d}s\Big)^{2pq}\Big]^{\frac{1}{q}} \\ &= C_1\,D\,\mathbb{E}^P[\mathcal{E}(H\cdot W)_T\Big(|\zeta|^{2pq} + \Big(\int_0^T |A_s|\mathrm{d}s\Big)^{2pq}\Big)]^{\frac{1}{q}} \\ &\le C_2\,D^{\frac{1+q}{q}}\mathbb{E}^P[|\zeta|^{2pq^2} + \Big(\int_0^T |A_s|\mathrm{d}s\Big)^{2pq^2}]^{\frac{1}{q^2}},\end{aligned}$$

where C_1, C_2 represent constants depending on p, M, T and the BMO norm of $\int HdW$. Similarly, with another constant C_3, $\mathbb{E}^P[\int_0^T |V_s|^{2p}\mathrm{d}s] \le C_3\,D^{\frac{1+q}{q}}\mathbb{E}^P[|\zeta|^{2pq^2} + \Big(\int_0^T |A_s|\mathrm{d}s\Big)^{2pq^2}\Big)]^{\frac{1}{q^2}}$, and hence (6). By applying the same arguments to (10) we finally get (7). □

4 A priori estimates for linear BSDEs with stochastic Lipschitz constants

In this section we shall derive a priori estimates for the variation of the linear BSDEs that play the role of good candidates for the derivatives of our original BSDE. These will be used to prove continuous differentiability of the smoothly parametrized solution in subsequent sections. Let (ζ, H, l_1, A) and (ζ', H', l_2, A') be parameters satisfying the properties (A1), (A2), (A3) and (A4) of Section 3 and suppose that l_1 and l_2 are globally Lipschitz continuous and differentiable in (u, v). Let (U, V) resp. (U', V') be solutions of the linear BSDE

$$U_t = \zeta - \int_t^T V_s dW_s + \int_t^T [l_1(s, U_s, V_s) + H_sV_s + A_s]\mathrm{d}s \tag{16}$$

resp.

$$U'_t = \zeta' - \int_t^T V'_s dW_s + \int_t^T [l_2(s, U'_s, V'_s) + H'_s V'_s + A'_s] \mathrm{d}s$$

both satisfying property (A5). Throughout let $\delta U_t = U_t - U'_t$, $\delta V_t = V_t - V'_t$, $\delta\zeta = \zeta - \zeta'$, $\delta A_t = A_t - A'_t$ and $\delta l(t, u, v) = l_1(t, u, v) - l_2(t, u, v)$.

Theorem 4.1 (A priori estimates). *Suppose we have for all $\beta \geq 1$, $\int_0^T \delta U_s^2 |\delta V_s|^2 \mathrm{d}s \in L^\beta(P)$ and $\int_0^T |\delta U_s \delta A_s| \mathrm{d}s \in L^\beta(P)$. Let $p \geq 1$ and $r > 1$ such that $\mathcal{E}(\int H' dW)_T \in L^r(P)$. Then there exists a constant $C > 0$, depending only on p, T, M and the BMO-norm of $\int H' dW$, such that with the conjugate exponent q of r we have*

$$\begin{aligned}
&\mathbb{E}^P\Big[\sup_{t\in[0,T]} |\delta U_t|^{2p}\Big] + \mathbb{E}^P\Big[\Big(\int_0^T |\delta V_s|^2 \mathrm{d}s\Big)^p\Big] \\
&\quad \leq C\Big\{\mathbb{E}^P\Big[|\delta\zeta|^{2pq^2} + \Big(\int_0^T |\delta l(s, U'_s, V'_s) + \delta A_s| \mathrm{d}s\Big)^{2pq^2}\Big]^{\frac{1}{q^2}} \\
&\qquad + (\mathbb{E}^P[|\zeta|^{2pq^2} + \big(\int_0^T |A_s| \mathrm{d}s\big)^{2pq^2}])^{\frac{1}{2q^2}} \mathbb{E}^P\Big[\Big(\int_0^T |H_s - H'_s|^2 \mathrm{d}s\Big)^{2pq^2}\Big]^{\frac{1}{2q^2}}\Big\}
\end{aligned}$$

We proceed in the same spirit as in the preceding section. Before proving Theorem 4.1 we will show a priori estimates with respect to the auxiliary probability measure Q defined by $Q = \mathcal{E}(\int H' dW)_T \cdot P$. Note that $\hat{W}_t = W_t - \int_0^t H'_s \mathrm{d}s$ is a Q-Brownian motion.

Lemma 4.2. *Let $p > 1$. There exists a constant $C > 0$, depending only on p, T and M, such that*

$$\begin{aligned}
\mathbb{E}^Q\Big[\sup_{t\in[0,T]} |\delta U_t|^{2p}\Big] \quad &\leq \quad C\Big\{\mathbb{E}^Q\Big[|\delta\zeta|^{2p} + \Big(\int_0^T |\delta l(s, U'_s, V'_s) + \delta A_s| \mathrm{d}s\Big)^{2p}\Big] \qquad (17)\\
&\quad + \Big(\mathbb{E}^Q\Big[|\zeta|^{2p} + \big(\int_0^T |A_s| \mathrm{d}s\big)^{2p}\Big]\Big)^{\frac{1}{2}} \mathbb{E}^Q\Big[\Big(\int_0^T |H_s - H'_s|^2 \mathrm{d}s\Big)^{2p}\Big]^{\frac{1}{2}}\Big\},
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}^Q[\Big(\int_0^T |\delta V_s|^2 \mathrm{d}s\Big)^p] \quad &\leq \quad C\Big\{\mathbb{E}^Q\Big[|\delta\zeta|^{2p} + \Big(\int_0^T |\delta l(s, U'_s, V'_s) + \delta A_s| \mathrm{d}s\Big)^{2p}\Big] \qquad (18)\\
&\quad + \Big(\mathbb{E}^Q\Big[|\zeta|^{2p} + \big(\int_0^T |A_s| \mathrm{d}s\big)^{2p}\Big]\Big)^{\frac{1}{2}} \mathbb{E}^Q\Big[\Big(\int_0^T |H_s - H'_s|^2 \mathrm{d}s\Big)^{2p}\Big]^{\frac{1}{2}}\Big\}.
\end{aligned}$$

Proof. The difference δU satisfies

$$\begin{aligned}
\delta U_t \quad &= \quad \delta\zeta - \int_t^T \delta V_s \mathrm{d}W_s + \int_t^T [(H_s V_s - H'_s V'_s) + l_1(s, U_s, V_s) - l_2(s, U'_s, V'_s) + \delta A_s] \mathrm{d}s \\
&= \quad \delta\zeta - \int_t^T \delta V_s \mathrm{d}W_s + \int_t^T [l_1(s, U'_s, V'_s) - l_2(s, U'_s, V'_s) + H'_s \delta V_s + \delta A_s] \mathrm{d}s \\
&\quad + \int_t^T [(H_s - H'_s) V_s + l_1(s, U_s, V_s) - l_1(s, U'_s, V'_s)] \mathrm{d}s.
\end{aligned}$$

Let $\beta > 0$. Applying Itô's formula to $e^{\beta t}\delta U_t^2, t \geq 0$, yields the equation

$$\begin{aligned} e^{\beta t}\delta U_t^2 &= e^{\beta T}\delta U_T^2 - 2\int_t^T e^{\beta s}\delta U_s \delta V_s \mathrm{d}W_s + 2\int_t^T e^{\beta s}\delta U_s\, H_s'\,\delta V_s \mathrm{d}s \\ &\quad + \int_t^T e^{\beta s}\Big[-\beta\delta U_s^2 - |\delta V_s|^2 + 2\big(l_1(s,U_s,V_s) - l_1(s,U_s',V_s')\big)\delta U_s\Big]\mathrm{d}s \\ &\quad +2\int_t^T e^{\beta s}\delta U_s(H_s - H_s')V_s \mathrm{d}s + 2\int_t^T e^{\beta s}\delta U_s(\delta l_s + \delta A_s)\mathrm{d}s, \end{aligned} \tag{19}$$

where $\delta l_s = l_1(s,U_s',V_s') - l_2(s,U_s',V_s')$. Using the Lipschitz property of l_1 we obtain

$$\begin{aligned} e^{\beta t}\delta U_t^2 &\leq e^{\beta T}\delta U_T^2 + \int_t^T e^{\beta s}\Big[(-\beta + 2M)\delta U_s^2 - |\delta V_s|^2 + 2M|\delta U_s|\,|\delta V_s|\Big]\mathrm{d}s \\ &\quad +2\int_t^T e^{\beta s}\delta U_s[(H_s - H_s')V_s + \delta l_s + \delta A_s]\mathrm{d}s - 2\int_t^T e^{\beta s}\delta U_s\delta V_s \mathrm{d}\hat{W}_s. \end{aligned}$$

If $\beta = (M^2 + 2M)$, then

$$\begin{aligned} e^{\beta t}\delta U_t^2 + \int_t^T e^{\beta s}(M|\delta U_s| - |\delta V_s|)^2 \mathrm{d}s &\leq e^{\beta T}\delta U_T^2 + 2\int_t^T e^{\beta s}\delta U_s[(H_s - H_s')V_s + \delta l_s + \delta A_s]\mathrm{d}s \\ &\quad -2\int_t^T e^{\beta s}\delta U_s\delta V_s\mathrm{d}\hat{W}_s. \end{aligned} \tag{20}$$

We will now derive the desired estimates from Equation (20). First observe that by taking conditional expectations, we get

$$e^{\beta t}\delta U_t^2 \leq e^{\beta T}\mathbb{E}^Q\left[\delta U_T^2 + 2\int_t^T e^{\beta s}\delta U_s[(H_s - H_s')V_s + \delta l_s + \delta A_s]\mathrm{d}s\Big|\mathcal{F}_t\right].$$

Let $p > 1$. Then for some constants $C_1, C_2, \ldots$, depending on p, T and M, we obtain

$$\sup_{t\in[0,T]}|\delta U_t|^{2p} \leq C_1 \sup_{t\in[0,T]}\left\{\Big(\mathbb{E}^Q\big[|\delta U_T|^2|\mathcal{F}_t\big] + \mathbb{E}\big[\int_0^T|\delta U_s[(H_s - H_s')V_s + \delta l_s + \delta A_s]|\mathrm{d}s\big|\mathcal{F}_t\big]\Big)^p\right\}$$

and by Doob's L^p inequality we get

$$\mathbb{E}^Q[\sup_{t\in[0,T]}|\delta U_t|^{2p}] \leq C_2\left\{\mathbb{E}^Q\Big[|\delta U_T|^{2p}\Big] + \mathbb{E}[\Big(\int_0^T|\delta U_s[(H_s - H_s')V_s + \delta l_s + \delta A_s]|\mathrm{d}s\Big)^p]\right\}.$$

By using Young's and Hölder's inequalities we have

$$
\begin{aligned}
&\mathbb{E}^Q\Big[\Big(\int_0^T |\delta U_s[(H_s-H'_s)V_s+\delta l_s+\delta A_s]|\mathrm{d}s\Big)^p\Big]\\
&\leq C_3\mathbb{E}^Q\left\{\sup_{t\in[0,T]}|\delta U_t|^p\left[\Big(\int_0^T|H_s-H'_s|^2\mathrm{d}s\Big)^{\frac{p}{2}}\Big(\int_0^T|V_s|^2\mathrm{d}s\Big)^{\frac{p}{2}}+\Big(\int_0^T|\delta l_s+\delta A_s|\mathrm{d}s\Big)^p\right]\right\}\\
&\leq \frac{1}{2C_4}\mathbb{E}^Q\Big[\sup_{t\in[0,T]}|\delta U_t|^{2p}\Big]\\
&\quad +4C_4\mathbb{E}^Q\Big[\Big(\int_0^T|H_s-H'_s|^2\mathrm{d}s\Big)^p\Big(\int_0^T|V_s|^2\mathrm{d}s)\Big)^p+\Big(\int_0^T|\delta l_s+\delta A_s|\mathrm{d}s\Big)^{2p}\Big]\\
&\leq \frac{1}{2C_4}\mathbb{E}^Q\Big[\sup_{t\in[0,T]}|\delta U_t|^{2p}\Big]+C_5\Big\{\mathbb{E}^Q\Big(\int_0^T|\delta l_s+\delta A_s|\mathrm{d}s\Big)^{2p}\\
&\qquad +\mathbb{E}^Q\Big[\Big(\int_0^T(H_s-H'_s)^2\mathrm{d}s\Big)^{2p}\Big]^{\frac{1}{2}}\mathbb{E}^Q\Big[\Big(\int_0^T|V_s|^2\mathrm{d}s\Big)^{2p}\Big]^{\frac{1}{2}}\Big\}.
\end{aligned}
\tag{21}
$$

Therefore, we may further estimate

$$
\begin{aligned}
\mathbb{E}^Q[\sup_{t\in[0,T]}|\delta U_t|^{2p}] \quad\leq\quad & C_6\Big\{\mathbb{E}^Q[\,|\delta\zeta|^{2p}]+\mathbb{E}^Q[(\int_0^T|\delta l_s+\delta A_s|\mathrm{d}s)^{2p}]\\
&+\mathbb{E}^Q\Big[\Big(\int_0^T|H_s-H'_s|^2\mathrm{d}s\Big)^{2p}\Big]^{\frac{1}{2}}\mathbb{E}^Q\Big[\Big(\int_0^T|V_s|^2\mathrm{d}s\Big)^{2p}\Big]^{\frac{1}{2}}\Big\}.
\end{aligned}
$$

Due to Lemma 3.2, $\mathbb{E}^Q\Big[\Big(\int_0^T|V_s|^2\mathrm{d}s\Big)^{2p}\Big]^{\frac{1}{2}}\leq C_7\mathbb{E}^Q\Big[|\zeta|^{2p}+(\int_0^T|A_s|\mathrm{d}s)^{2p}\Big]^{\frac{1}{2}}<\infty$, which implies the δU_s part of Inequality (17).

In order to prove the second inequality, note that (20) also implies

$$
\begin{aligned}
\int_t^T e^{\beta s}|\delta V_s|^2\mathrm{d}s \quad\leq\quad & e^{\beta T}\delta U_T^2+2\int_t^T e^{\beta s}\delta U_s[(H_s-H'_s)V_s+\delta l_s+\delta A_s]\mathrm{d}s\\
&+2M\int_t^T e^{\beta s}|\delta U_s|\,|\delta V_s|\mathrm{d}s-2\int_t^T e^{\beta s}\delta U_s\delta V_s\mathrm{d}\hat{W}_s.
\end{aligned}
\tag{22}
$$

Equation (22), Doob's L^p-inequality and the Burkholder-Davis-Gundy inequality imply

$$
\begin{aligned}
\mathbb{E}^Q[\Big(\int_0^T|\delta V_s|^2\mathrm{d}s\Big)^p] \quad\leq\quad & C_8\Big\{(\mathbb{E}^Q\Big[|\delta\zeta|^{2p}+\int_0^T|\delta U_s|^{2p}\mathrm{d}s\Big]+\mathbb{E}^Q[\Big(\int_0^T\delta U_s^2\delta|V_s|^2\mathrm{d}s\Big)^{\frac{p}{2}}]\\
&+\mathbb{E}^Q[(\int_0^T|\delta U_s[(H_s-H'_s)V_s+\delta l_s+\delta A_s]|\mathrm{d}s)^p]\Big\}.
\end{aligned}
$$

Consequently, Young's inequality allows to deduce

$$
\begin{aligned}
\mathbb{E}^Q[\Big(\int_0^T|\delta V_s|^2\mathrm{d}s\Big)^p] \quad\leq\quad & C_9\Big\{\mathbb{E}^Q\Big[|\delta\zeta|^{2p}+\int_0^T|\delta U_s|^{2p}\mathrm{d}s\Big]+\mathbb{E}^Q[\sup_{t\in[0,T]}|\delta U_t|^{2p}]\\
&+\mathbb{E}^Q[(\int_0^T|\delta U_s[(H_s-H'_s)V_s+\delta l_s+\delta A_s]|\mathrm{d}s)^p]\Big\}.
\end{aligned}
$$

Finally, (17) and (21) imply

$$\begin{aligned}\mathbb{E}^Q[\Big(\int_0^T |\delta V_s|^2 \mathrm{d}s\Big)^p] &\leq C_{10}\mathbb{E}^Q\Big[|\delta\zeta|^{2p} + \Big(\int_0^T |\delta l_s + \delta A_s|\mathrm{d}s\Big)^{2p}\Big] \\ &\quad + C_{10}\mathbb{E}^Q[|\zeta|^{2p} + (\int_0^T |A_s|\mathrm{d}s)^{2p}]^{\frac{1}{2}}\mathbb{E}^Q\Big[\Big(\int_0^T |H_s - H'_s|^2\mathrm{d}s\Big)^{2p}\Big]^{\frac{1}{2}}\end{aligned}$$

and hence the proof is complete. □

Proof of Theorem 4.1. This can be deduced from Lemma 4.2 with arguments similar to those of Theorem 3.1. We just have to invoke Lemma 1.2. □

5 A priori estimates for quadratic BSDEs

Consider the two quadratic BSDEs

$$Y_t = \xi - \int_t^T Z_s dW_s + \int_t^T [l_1(s, Y_s, Z_s) + \alpha Z_s^2]\mathrm{d}s \tag{23}$$

and

$$Y'_t = \xi' - \int_t^T Z'_s dW_s + \int_t^T [l_2(s, Y'_s, Z'_s) + \alpha (Z'_s)^2]\mathrm{d}s, \tag{24}$$

where ξ and ξ' are two bounded $\mathcal{F}_T$-measurable random variables, and l_1 and l_2 are globally Lipschitz and differentiable in (y, z). Put now $\delta Y_t = Y_t - Y'_t$, $\delta Z_t = Z_t - Z'_t$, $\delta\xi = \xi - \xi'$ and $\delta l = l_1 - l_2$. The a priori estimates we shall prove next will serve for establishing (moment) smoothness of the solution of the quadratic BSDE with respect to a parameter on which the terminal variable depends smoothly. Note first that by boundedness of ξ and ξ' we have that both $\int Z dW$ and $\int Z' dW$ are BMO martingales, so that we may again invoke the key Lemma 1.2.

Theorem 5.1. *Suppose that for all $\beta \geq 1$ we have $\int_0^T |\delta l(s, Y_s, Z_s)|\mathrm{d}s \in L^\beta(P)$. Let $p > 1$ and choose $r > 1$ such that $\mathcal{E}(\alpha(Z_s + Z'_s)\cdot W)_T \in L^r(P)$. Then there exists a constant $C > 0$, depending only on p, T, M and the BMO-norm of $(\alpha\int(Z_s + Z'_s)dW)$, such that with the conjugate exponent q of r we have*

$$\begin{aligned}&\mathbb{E}^P\Big[\sup_{t\in[0,T]} |\delta Y_t|^{2p}\Big] + \mathbb{E}^P\left[\left(\int_0^T |\delta Z_s|^2\mathrm{d}s\right)^p\right] \\ &\qquad \leq C\left(\mathbb{E}^P\Big[|\delta\xi|^{2pq^2} + (\int_0^T |\delta l(s, Y_s, Z_s)|\mathrm{d}s)^{2pq^2}\Big]\right)^{\frac{1}{q^2}}.\end{aligned}$$

Moreover we have

$$\mathbb{E}^P[\int_0^T |\delta Y_s|^2\mathrm{d}s] + \mathbb{E}^P[\int_0^T |\delta Z_s|^2\mathrm{d}s] \leq C\left(\mathbb{E}^P\left[|\delta\xi|^{2q^2} + \Big(\int_0^T |\delta l(s, Y_s, Z_s)|\mathrm{d}s\Big)^{2q^2}\right]\right)^{\frac{1}{q^2}}.$$

We give only a sketch of the proof since the arguments are very similar to the ones used in the proofs in Sections 3 and 4.

First observe that

$$\delta Y_t = \delta\xi - \int_t^T \delta Z_s \mathrm{d}W_s + \int_t^T [l_1(s,Y_s,Z_s) - l_1(s,Y'_s,Z'_s) + \delta l(s,Y'_s,Z'_s) + \alpha(Z_s+Z'_s)\delta Z_s]\mathrm{d}s.$$

By applying Itô's formula to $e^{\beta t}|\delta Y_t|^2$ we obtain

$$\begin{aligned}
& e^{\beta t}|\delta Y_t|^2 - e^{\beta T}|\delta Y_T|^2 \\
&= 2\int_t^T e^{\beta s}\delta Y_s\Big(l_1(s,Y_s,Z_s) - l_1(s,Y'_s,Z'_s) + \delta l(s,Y'_s,Z'_s)\Big)\mathrm{d}s - 2\int_t^T e^{\beta s}(\beta|\delta Y_s|^2 + |\delta Z_s|^2)\mathrm{d}s \\
&\quad +2\int_t^T e^{\beta s}\delta Y_s \alpha(Z_s+Z'_s)\delta Z_s \mathrm{d}s - 2\int_t^T e^{\beta s}\delta Y_s \delta Z_s \mathrm{d}W_s. \qquad (25)
\end{aligned}$$

We start with a priori estimates under the auxiliary probability measure Q defined by $Q = \mathcal{E}(\alpha\int(Z_s+Z'_s)dW)\cdot P$. Note that $\tilde{W}_t = W_t - \int_0^t \alpha(Z_s+Z'_s)ds$ is a Q-Brownian motion.

Let $\beta > 0$. Equality (25) and the Lipschitz property of l_1 yield

$$\begin{aligned}
e^{\beta t}|\delta Y_t|^2 \leq\; & e^{\beta T}|\delta\xi|^2 - 2\int_t^T e^{\beta s}\,\delta Y_s\,\delta Z_s \mathrm{d}\tilde{W}_s + 2\int_t^T e^{\beta s}\delta Y_s \delta l(s,Y'_s,Z'_s)\mathrm{d}s \\
&+ \int_t^T e^{\beta s}\Big((-\beta+2M)|\delta Y_s|^2 - |\delta Z_s|^2 + 2M|\delta Y_s||\delta Z_s|\Big)\mathrm{d}s.
\end{aligned}$$

By choosing $\beta = M^2 + 2M$ we obtain the general inequality

$$\begin{aligned}
& e^{\beta t}|\delta Y_t|^2 + \int_t^T e^{\beta s}(|\delta Z_s| - M|\delta Y_s|)^2\mathrm{d}s \\
&\leq e^{\beta T}|\delta\xi|^2 + \int_t^T e^{\beta s}\delta Y_s \delta l(s,Y'_s,Z'_s)\mathrm{d}s - 2\int_t^T e^{\beta s}\delta Y_s\,\delta Z_s \mathrm{d}\tilde{W}_s. \qquad (26)
\end{aligned}$$

Note that the process $\int_0^t e^{\beta s}\,\delta Y_s\,\delta Z_s \mathrm{d}\tilde{W}_s$ is a strict martingale because δY_s is bounded and $(\delta Z\cdot\tilde{W})$ is BMO relative to Q.

Notice that Equation (26) is of similar but simpler form than Equation (20). This is because the $(H_s - H'_s)$ term in (20) has been completely absorbed by the Girsanov measure change. As a consequence, following the proof of Lemma 4.2, we obtain the following estimates:

Lemma 5.2. *For all $p > 1$ there exists a constant $C > 0$, depending only on p, M and T, such that*

$$\mathbb{E}^Q\Big[\sup_{t\in[0,T]}|\delta Y_t|^{2p}\Big] + \mathbb{E}^Q\left[\left(\int_0^T |\delta Z_s|^2\mathrm{d}s\right)^p\right] \leq C\,\mathbb{E}^Q\Big[|\delta\xi|^{2p} + (\int_0^T |\delta l(s,Y_s,Z_s)|\mathrm{d}s)^{2p}\Big].$$

Moreover we have

$$\mathbb{E}^Q[\int_0^T |\delta Y_t|^2\mathrm{d}s] + \mathbb{E}^Q\big[\int_0^T |\delta Z_s|^2\mathrm{d}s\big] \leq C\mathbb{E}^Q\left[|\delta\xi|^2 + \Big(\int_0^T |\delta l(s,Y_s,Z_s)|\mathrm{d}s\Big)^2\right]. \qquad (27)$$

Proof of Theorem 5.1. The arguments are similar to those of the proof of Theorem 3.1. Just make use of Lemma 1.2. □

6 Proof of the differentiability

We now approach the problem of differentiability of the solutions of a quadratic BSDEs with respect to a vector parameter on which the terminal condition depends differentiably. We start with the proof of the weaker property of Theorem 2.1. Our line of reasoning will be somewhat different from the one used for instance by Kunita [10] in the proof of the diffeomorphism property of smooth flows of solutions of stochastic differential equations. He starts with formally differentiating the stochastic differential equation, and showing that the resulting equation possesses a solution. The latter is then used explicitly in moment estimates for *its deviation from difference quotients* of the original equation. The estimates are then used to prove pathwise convergence of the difference quotients to the solution of the differentiated SDE. We emphasize that in our proofs, we will have to derive moment estimates for *differences of difference quotients* instead. They will allow us to show the existence of a derivative process in a Cauchy sequence type argument using the completeness of underlying vector spaces, which of course will be the solution process of the formally differentiated BSDE. So our procedure contains the statement of the existence of a solution of the latter as a by-product of the proof of the Theorem 2.1. It is not already available as a good candidate for the derivative process, since, as we stated earlier, the formally differentiated BSDE is a globally Lipschitz one with random Lipschitz constants for which the classical existence theorems do not immediately apply. Throughout assume that $f(t,x,y,z) = l(t,x,y,z) + \alpha|z|^2$ and $\xi(x)$ satisfy (C1) and (C2) respectively.

For all $x \in \mathbb{R}^n$ let (Y_t^x, Z_t^x) be a solution of the BSDE (3). It is known that the solution is unique and that $(Y^x, Z^x) \in \mathcal{R}^\infty(\mathbb{R}^1) \times \mathbb{L}^2(\mathbb{R}^d)$ (see [9]).

It follows from Lemma 1 in [11] that there exists a constant $D > 0$ such that for all $x \in \mathbb{R}^n$ we have $\|(Z^x \cdot W)_T\|_{BMO_2} \leq D$. Now let $r > 1$ be such that $\Psi(r) > 2\alpha D$ (see property 3) of Lemma 1.2), and denote as before by q the conjugate exponent of r.

Proof of Theorem 2.1. To simplify notation we assume that $M > 0$ is a constant such that $\xi(x)$, $x \in \mathbb{R}^n$, and the derivatives of l in (y,z) are all bounded by M. We first show that all the partial derivatives of Y and Z exist. Let $x \in \mathbb{R}^n$ and $e_i = (0,\dots,1,\dots 0)$ be the unit vector in $\mathbb{R}^n$ the ith component of which is 1 and all the others 0. For all $h \neq 0$, let $U_t^h = \frac{1}{h}(Y_t^{x+e_ih} - Y_t^x)$, $V_t^h = \frac{1}{h}(Z_t^{x+he_i} - Z_t^x)$ and $\zeta^h = \frac{1}{h}(\xi(x+he_i) - \xi(x))$.

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

$$U_t^h = \zeta^h - \int_t^T V_s^h \mathrm{d}W_s + \int_t^T \frac{1}{h}[f(s, x+he_i, Y_s^{x+he_i}, Z_s^{x+he_i}) - f(s,x,Y_s^x,Z_s^x)]\mathrm{d}s.$$

To simplify the last term we use a line integral transformation. For all $(\omega,t) \in \Omega \times \mathbb{R}_+$ let $s_{x,h} = s_{x,h}(\omega,t) : [0,1] \to \mathbb{R}^{n+1+d}$ be defined by $s_{x,h}(\theta) = (x+\theta he_i, Y_t^x + \theta(Y_t^{x+he_i} - Y_t^x), Z_t^x + \theta(Z_t^{x+he_i} - Z_t^x))$. Though $s_{x,h}$ depends on i we omit to indicate this dependence for ease of notation. Note that $\frac{1}{h}s'_{x,h}(\theta) = (e_i, U_t^h, V_t^h)$. Moreover, $A_t^{x,h} = \int_0^1 \frac{\partial l}{\partial x_i}(s_{x,h}(\theta))\mathrm{d}\theta$, $G_t^{x,h} = \int_0^1 \frac{\partial l}{\partial y}(s_{x,h}(\theta))\mathrm{d}\theta$ and $I_t^{x,h} = \int_0^1 \frac{\partial l}{\partial z}(s_{x,h}(\theta))\mathrm{d}\theta$ are $(\mathcal{F}_t)$-adapted processes satisfying

$$\begin{aligned}\frac{1}{h}[l(t,x+he_i,Y_t^{x+he_i},Z_t^{x+he_i}) - l(t,x,Y_t^x,Z_t^x)] &= \int_0^1 \langle \nabla l(s_{x,h}(\theta)), s'_{x,h}(\theta)\rangle \mathrm{d}\theta \\ &= A_t^{x,h} + G_t^{x,h}U_t^h + I_t^{x,h}V_t^h.\end{aligned}$$

Since the derivatives of l are bounded by M, $G_t^{x,h}$ and $I_t^{x,h}$ are bounded by M as well. However, we stress that $A_t^{x,h}$ is not necessarily bounded. We define two random functions $m_s^{x,h}(u,v)$ and $m_s(u,v)$ from $\mathbb{R}^{1+d}$ to $\mathbb{R}$ such that $(u,v) \mapsto m_s^{x,h}(u,v) = (G_s^{x,h}u + I_s^{x,h}v)$ and $m_s(u,v) = [\partial_y l(s, Y_s^x, Z_s^x)u + \partial_z l(s, Y_s^x, Z_s^x)v]$. Observe that these functions satisfy (A3) and that they are Lipschitz continuous and differentiable in (u,v). In these terms,

$$U_t^h = \zeta^h - \int_t^T V_s^h \mathrm{d}W_s + \int_t^T [m_s^{x,h}(U_s^h, V_s^h) + A_s^{x,h} + \alpha(Z_s^{x+he_i} + Z_s^x)V_s^h]\mathrm{d}s,$$

and thus we obtain an equation as modelled by (5). Notice that for all h, $h' \neq 0$ the pairs (U^h, V^h) and $(U^h - U^{h'}, V^h - V^{h'})$ satisfy assumptions (A4) and (A5). Therefore Theorem 4.1 implies with $\delta A_t = A_t^{x,h} - A_t^{x,h'}$

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{t\in[0,T]} |U_t^h - U_t^{h'}|^{2p}\Big] \\
&\leq C\Big\{\mathbb{E}\Big[|\zeta^h - \zeta^{h'}|^{2pq^2} + \left(\int_0^T |m_s^{x,h}(U_s^h, V_s^h) - m_s^{x',h'}(U_s^h, V_s^h) + \delta A_s| ds\right)^{2pq^2}\Big]^{\frac{1}{q^2}} \qquad (28) \\
&\quad + \mathbb{E}\Big[|\zeta^{h'}|^{2pq^2} + \Big(\int_0^T |A_s^{x,h'}|\mathrm{d}s\Big)^{2pq^2}\Big]^{\frac{1}{2q^2}} \mathbb{E}\Big[\left(\int_0^T \alpha^2 |Z_s^{x+he_i} - Z_s^{x+h'e_i}|^2 ds\right)^{2pq^2}\Big]^{\frac{1}{2q^2}}\Big\}.
\end{aligned}$$

Condition (C2) implies that $\mathbb{E}[|\zeta^h - \zeta^{h'}|^{2pq^2}]$ converges to zero as h, $h' \to 0$. Moreover, for some open set $\mathcal{O}$ containing 0 we have $\sup_{h'\in\mathcal{O}\setminus\{0\}}(\mathbb{E}|\zeta^{h'}|^{2pq^2}) < \infty$. Due to Condition (C1), we may also assume that $\sup_{h'\in\mathcal{O}\setminus\{0\}} \mathbb{E}[(\int_0^T |A_s^{x,h'}|\mathrm{d}s)^{2pq^2}] < \infty$. Moreover,

$$\lim_{h\to 0} \mathbb{E}(\int_0^T |l(s, x + he_i, Y_s^x, Z_s^x) - l(s, x, Y_s^x, Z_s^x)|\mathrm{d}s)^\beta = 0$$

for all $\beta \geq 1$, and therefore, with Theorem 5.1, the third summand on the right hand side of (28) converges to zero as h, $h' \to 0$.

In order to prove convergence of the second summand let $P \otimes \lambda$ be the product measure of P and the Lebesgue measure λ on $[0,T]$. It follows from Theorem 5.1 that Z^{x+he_i} converges to Z^x in measure relative to $P \otimes \lambda$. Moreover, for all $t \in [0,T]$, $Y_t^{x+he_i}$ converges to Y_t^x in probability. Since the partial derivatives l_y and l_z are continuous and bounded, dominated convergence implies $\lim_{h\to 0} \mathbb{E}^P \left(\int_0^T |m_s^{x,h}(U_s^h, V_s^h) - m_s(U_s^h, V_s^h)| ds\right)^{2pq^2} = 0$. Condition (C1) guarantees $\lim_{h\to 0} \mathbb{E}^P \left(\int_0^T |\delta A_s| ds\right)^{2pq^2} = 0$, and hence, $\lim_{h,h'\to 0} \mathbb{E}\Big[\sup_{t\in[0,T]} |U_t^h - U_t^{h'}|^{2p}\Big] = 0$.

Finally, Theorem 4.1 and an estimation similar to (28) yield

$$\lim_{h,h'\to 0} \mathbb{E}\left(\int_0^T |V_s^h - V_s^{h'}|^2 \mathrm{d}s\right)^p = 0.$$

Now let (h_n) be a sequence in $\mathbb{R} \setminus \{0\}$ converging to zero. Then, since $\mathcal{R}^{2p}(\mathbb{R}^1)$ and $\mathbb{L}^{2p}(\mathbb{R}^d)$ are Banach spaces, the sequence U^{h_n} converges to a process $\frac{\partial}{\partial x_i} Y_t^x$, and V^{h_n} to a process $\frac{\partial}{\partial x_i} Z_t^x$ with respect to the corresponding norms. By convergence term by term for the difference quotient version of the quadratic BSDE and its formal derivative, which follows from our a priori estimates, we see that the pair $(\frac{\partial}{\partial x_i}Y_t^x, \frac{\partial}{\partial x_i}Z_t^x)$ is a solution of the BSDE

$$\begin{aligned}\frac{\partial Y_t^x}{\partial x_i} &= \frac{\partial}{\partial x_i}\xi(x) - \int_t^T \frac{\partial Z_s^x}{\partial x_i}\mathrm{d}W_s + \int_t^T \Big[\partial_{x_i} l(s,x,Y_s^x,Z_s^x) \\ &\qquad + \partial_y l(s,x,Y_s^x,Z_s^x)\frac{\partial Y_s^x}{\partial x_i} + \partial_z l(s,x,Y_s^x,Z_s^x)\frac{\partial Z_s^x}{\partial x_i} + 2\alpha Z_s^x \frac{\partial Z_s^x}{\partial x_i}\Big]\mathrm{d}s.\end{aligned}$$

Similarly to the first part one can show that $\lim_{h\to 0}\mathbb{E}\Big[\sup_{t\in[0,T]}|U_t^h - \frac{\partial}{\partial x_i}Y_t^x|^{2p}\Big] = 0$ and $\lim_{h\to 0}\mathbb{E}\left(\int_0^T |V_s^h - \frac{\partial}{\partial x_i}Z_s^x|^2\mathrm{d}s\right)^p = 0$, and thus $\mathbb{R}^n \to \mathcal{R}^{2p}(\mathbb{R}^1)\times\mathbb{L}^{2p}(\mathbb{R}^d)$, $x\mapsto (Y_t^x, Z_t^x)$ is partially differentiable. The a priori estimates of Theorem 4.1 imply that the mapping $x \mapsto (\nabla Y_t^x, \nabla Z_t^x)$ is continuous and hence, (Y_t^x, Z_t^x) is totally differentiable. Since differentiability with respect to $2p$th moments implies differentiability with respect to all inferior moments above 1, we have established the result. □

As a byproduct of the previous proof we obtain that for every $x\in\mathbb{R}^n$ there exists a solution $(\nabla Y_t^x, \nabla Z_t^x)$ of the BSDE (4).

We now proceed with the proof of Theorem 2.2, in which we claim pathwise continuous differentiability. To be consistent with the previous proof, we will again compare difference quotients varying in h. To this end we need the following estimates.

Lemma 6.1. *Suppose (C3) is satisfied and that l and the derivatives of l are all Lipschitz continuous in (x,y,z). Then for all $p>1$ there exists a constant $C>0$, dependent only on p, T, M and D, such that for all $x,x'\in\mathbb{R}^n$, $h,h'\in\mathbb{R}$ and $i\in\{1,\dots,n\}$,*

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|Y_t^{x+he_i} - Y_t^{x'+h'e_i}\right|^{2p} + \left(\int_0^T |Z_s^{x+he_i} - Z_s^{x'+h'e_i}|^2 ds\right)^p\right] \le C\big(|x-x'|^2 + |h-h'|^2\big)^p.$$

Proof. This follows from Theorem 5.1, where we put $l_1(s,y,z) = l(s,x+he_i,y,z)$, $l_2(s,y,z) = l(s,x'+h'e_i,y,z)$. □

The preceding Lemma immediately implies a first pathwise smoothness result in x for the process Y^x. In fact, Kolmogorov's continuity criterion applies and yields a modification of Y^x which is continuous in x. More precisely:

Corollary 6.2. *There exists a process $\hat{Y}^x$ such that for all $(t,\omega)\in[0,T]\times\Omega$, the function $x\mapsto \hat{Y}_t^x(\omega)$ is continuous, and for all (t,x) we have $\hat{Y}_t^x = Y_t^x$ almost surely.*

Let e_i be a unit vector in $\mathbb{R}^n$. For all $x\in\mathbb{R}^n$ and $h\neq 0$, let $U_t^{x,h} = \frac{1}{h}(Y_t^{x+he_i} - Y_t^x)$, $V_t^{x,h} = \frac{1}{h}(Z_t^{x+he_i} - Z_t^x)$ and $\zeta^{x,h} = \frac{1}{h}\big(\xi(x+he_i) - \xi(x)\big)$. If $h=0$, then define $U_t^{x,0} = \frac{\partial}{\partial x_i}Y^x$, $V_t^{x,0} = \frac{\partial}{\partial x_i}Z^x$ and $\zeta^{x,0} = \frac{\partial}{\partial x_i}\xi(x)$. The proof of Theorem 2.2 will be based on the following result on the usual difference of difference quotients. Knowing a "good candidate" for the derivative from Theorem 2.1 we allow $h=0$ this time, by replacing the difference quotient with this candidate.

Lemma 6.3. *Let $p > 1$ and $\mathcal{O} \subset \mathbb{R}^{n+1}$ be an open set contained in a ball of radius κ. Suppose that Condition (C3) holds and that l and the derivatives of l in (x, y, z) are Lipschitz continuous in (x, y, z) with Lipschitz constant $L > 0$. Then there exists a constant C, depending on κ, L, p, T, M, D, such that for all (x, h) and $(x', h') \in \mathcal{O}$,*

$$\mathbb{E}\Big[\sup_{t\in[0,T]} |U_t^{x,h} - U_t^{x',h'}|^{2p}\Big] \leq C(|x - x'|^2 + |h - h'|^2)^p. \tag{29}$$

Proof. Throughout the proof, $C_1, C_2, \ldots$ are constants depending on κ, L, p, T, M, D.

Since $\mathcal{O}$ is bounded, (C3) implies that for every $r > 1$ there exists a constant C_1 such that for all $(x, h) \in \mathcal{O}$ we have $\mathbb{E}(\sup_{t\in[0,T]} |\zeta_t^{x,h}|^{2r}) < C_1$. Now let $s_{x,h}$, $m^{x,h}$, $A^{x,h}$, $G^{x,h}$, $I^{x,h}$ and $U^{x,h}$ be defined as in the proof of Theorem 2.2, and denote $A^{x,0} = \frac{\partial l}{\partial x}(x, Y^x, Z^x)$, $G^{x,0} = \frac{\partial l}{\partial y}(x, Y^x, Z^x)$, etc. Then the estimate (29) will be deduced from the inequality

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{t\in[0,T]} |U_t^{x,h} - U_t^{x',h'}|^{2p}\Big] \qquad (30)\\
\leq\;& C_2\Big\{\mathbb{E}\Big[|\zeta^{x,h} - \zeta^{x',h'}|^{2pq^2}\Big]^{\frac{1}{q^2}}\\
+\;& \mathbb{E}\Big[\Big(\int_0^T |m_s^{x',h'}(U_s^{x,h}, V_s^{x,h}) - m_s^{x,h}(U_s^{x,h}, V_s^{x,h})| + |A_s^{x,h} - A_s^{x',h'}|ds\Big)^{2pq^2}\Big]^{\frac{1}{q^2}}\\
+\;& \mathbb{E}\Big[|\zeta^{x',h'}|^{2pq^2} + \Big(\int_0^T |A_s^{x',h'}|\mathrm{d}s\Big)^{2pq^2}\Big]^{\frac{1}{2q^2}} \mathbb{E}\Big[\left(\int_0^T \alpha^2 |Z_s^{x'+h'e_i} - Z_s^{x+he_i}|^2 ds\right)^{2pq^2}\Big]^{\frac{1}{2q^2}}\Big\}
\end{aligned}$$

which follows from Theorem 4.1. We first analyze the order of the convergence of

$$B_1(x, x', h, h') = \mathbb{E}\Big[\left(\int_0^T |m_s^{x',h'}(U_s^{x,h}, V_s^{x,h}) - m_s^{x,h}(U_s^{x,h}, V_s^{x,h})|\mathrm{d}s\right)^{2pq^2}\Big]^{\frac{1}{q^2}} \text{ as } h, h' \to 0.$$

To this end notice that

$$\begin{aligned}
B_1(x, x', h, h') \;\leq\; C_3\Big\{ &\left(\mathbb{E}\left(\int_0^T |G_t^{x',h'} - G_t^{x,h}||U_t^{x,h}|\mathrm{d}t\right)^{2pq^2}\right)^{\frac{1}{q^2}}\\
&+ \left(\mathbb{E}\left(\int_0^T |I_t^{x',h'} - I_t^{x,h}||V_t^{x,h}|\mathrm{d}t\right)^{2pq^2}\right)^{\frac{1}{q^2}}\Big\}.
\end{aligned}$$

Then

$$
\begin{aligned}
&\int_0^T |G_t^{x',h'} - G_t^{x,h}||U_t^{x,h}|\mathrm{d}t \\
&\quad \le \sup_{t\in[0,T]} |U_t^{x,h}| \int_0^T |G_t^{x',h'} - G_t^{x,h}|\mathrm{d}t \\
&\quad \le \sup_{t\in[0,T]} |U_t^{x,h}| \int_0^T \left(\int_0^1 |\partial_y l(s_{x',h'}(\theta)) - \partial_y l(s_{x,h}(\theta))|\mathrm{d}\theta \right) \mathrm{d}t \\
&\quad \le \sup_{t\in[0,T]} |U_t^{x,h}| \int_0^T \int_0^1 L|s_{x',h'}(\theta) - s_{x,h}(\theta)|\mathrm{d}\theta\mathrm{d}t \\
&\quad \le C_4 \sup_{t\in[0,T]} |U_t^{x,h}| \Big(|x'-x| + |h'-h| + \sup_{t\in[0,T]} |Y_t^{x'} - Y_t^x| + \sup_{t\in[0,T]} |Y_t^{x'+h'e_i} - Y_t^{x+he_i}| \\
&\qquad + \int_0^T (|Z_t^{x'} - Z_t^x| + |Z_t^{x'+h'e_i} - Z_t^{x+he_i}|)\mathrm{d}t \Big),
\end{aligned}
$$

and, by applying Hölder's inequality we obtain with Lemma 6.1

$$
\left(\mathbb{E}\left(\int_0^T |G_t^{x',h'} - G_t^{x,h}||U_t^{x,h}|\mathrm{d}t \right)^{2pq^2} \right)^{\frac{1}{q^2}} \quad \le \quad C_5 \left(|h-h'|^2 + |x-x'|^2 \right)^p .
$$

Similarly, $\left(\mathbb{E}\left[\int_0^T |I_t^{x',h'} - I_t^{x,h}||V_t^{x,h}|\mathrm{d}t \right]^{2pq^2} \right)^{\frac{1}{q^2}} \le C_6 \left(|h-h'|^2 + |x-x'|^2 \right)^p$, and so we conclude $B_1(x,x',h,h') \le C_7 \left(|h-h'|^2 + |x-x'|^2 \right)^p$.
By using similar arguments we get

$$
\begin{aligned}
\mathbb{E}\left[\Big(\int_0^T |A_t^{x,h} - A_t^{x',h'}|\mathrm{d}t \Big)^{2pq^2} \right]^{\frac{1}{q^2}} \quad &\le \quad \mathbb{E}\left[\left(\int_0^T \int_0^1 |\partial_x l(s_{x',h'}(\theta)) - \partial_x l(s_{x,h}(\theta))|\mathrm{d}\theta\mathrm{d}t \right)^{2pq^2} \right]^{\frac{1}{q^2}} \\
&\le \quad C_8 \mathbb{E}\left[\left(\int_0^T \int_0^1 |s_{x',h'}(\theta) - s_{x,h}(\theta)|\mathrm{d}\theta\mathrm{d}t \right)^{2pq^2} \right]^{\frac{1}{q^2}} \\
&\le \quad C_9 \left(|h-h'|^2 + |x-x'|^2 \right)^p .
\end{aligned}
$$

Theorem 5.1 and the Lipschitz continuity of l imply

$$
\begin{aligned}
&\mathbb{E}\Big[\left(\int_t^T |Z_s^{x+he_i} - Z_s^{x'+h'e_i}|^2 ds \right)^{2pq^2} \Big]^{\frac{1}{2q^2}} \\
\le \quad & C_{10} \mathbb{E}\Big[|(\xi(x+he_i) - \xi(x'+h'e_i)|^{4pq^4} \\
&\qquad + \Big(\int_0^T |l(s, x+he_i, Y^{x+he_i}, Z^{x+he_i}) - l(s, x'+h'e_i, Y^{x+he_i}, Z^{x+he_i})|\mathrm{d}s \Big)^{4pq^4} \Big]^{\frac{1}{2q^4}} \\
\le \quad & C_{11}(|x-x'|^2 + |h-h'|^2)^p.
\end{aligned}
$$

Finally, (C3) yields $(\mathbb{E}|\zeta^{x,h} - \zeta^{x',h'}|^{2pq^2})^{\frac{1}{q^2}} \le C_{12}(|x-x'|^2 + |h-h'|^2)^p$, and hence

$$
\mathbb{E}\Big[\sup_{t\in[0,T]} |U_t^{x,h} - U_t^{x',h'}|^{2p} \Big] \le C_{13}(|x-x'|^2 + |h-h'|^2)^p.
$$

$\square$

Proof of Theorem 2.2. To simplify notation we may assume that (29) is satisfied for $\mathcal{O} = \mathbb{R}^{n+1}$. Assume that Y_t^x is continuous in x (see Corollary 6.2). Lemma 6.3 and Kolmogorov's continuity criterion imply that $U_t^{x,h}$ has a modification $\hat{U}_t^{x,h}$ continuous in (x, h). Define $\frac{\partial}{\partial x_i} Y_t^x = \hat{U}_t^{x,0}$ and note that we obtain thus a continuous version of the solution of the BSDE (4). For all $(x, h) \in \mathbb{Q}^{n+1}$ let $N(x, h)$ be a null set such that for all $\omega \notin N(x, h)$ we have $\hat{U}_t^{x,h}(\omega) = U_t^{x,h}(\omega)$. Then, $N = \bigcup_{(x,h)\in\mathbb{Q}^{n+1}} N(x, h)$ is a null set such that for all $\omega \notin N$ the following implication holds: If $q_k \in \mathbb{Q}^n$ and $r_k \in \mathbb{Q} \setminus \{0\}$ are sequences with $\lim_{k\to\infty} q_k = x \in \mathbb{R}^n$ and $\lim_{k\to\infty} r_k = 0$, then

$$\lim_{k\to\infty} \frac{1}{r_k}(Y_t^{q_k + r_k e_i} - Y_t^{q_k}) = \frac{\partial}{\partial x_i} Y_t^x.$$

As a consequence of this and the subsequent Lemma 6.4, $Y_t^x(\omega)$ is continuously partially differentiable relative to x_i if $\omega \notin N$. Since we can choose such a null set for any $i \in \{1, \ldots, n\}$, total differentiability follows and the proof is complete. $\square$

Lemma 6.4. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous function and $g : \mathbb{R}^n \to \mathbb{R}^n$ a continuous vector field. Suppose that for all sequences $q_k \in \mathbb{Q}^n$ with $q_k \to x \in \mathbb{R}^n$ and $r_k \in \mathbb{Q} \setminus \{0\}$ with $r_k \to 0$ we have*

$$\lim_{k\to\infty} \frac{1}{r_k}(f(q_k + r_k e_i) - f(q_k)) = g_i(x),$$

where $1 \le i \le n$. Then f is differentiable and $\nabla f = g$.

Proof. To simplify notation assume that $n = 1$. Let $x_k \in \mathbb{R}$ with $x_k \to x \in \mathbb{R}$ and $h_k \in \mathbb{R} \setminus \{0\}$ with $h_k \to 0$. Since f is continuous we may choose $q_k \in \mathbb{Q}$ and $r_k \in \mathbb{Q} \setminus \{0\}$ such that $|f(q_k) - f(x_k)| \le \frac{|h_k|}{2^k}$, $|f(q_k + r_k) - f(x_k + h_k)| \le \frac{|h_k|}{2^k}$ and $|\frac{1}{r_k} - \frac{1}{h_k}| \le \frac{1}{2^k}$. Then

$$\begin{aligned}
&|\frac{1}{h_k}(f(x_k + h_k) - f(x_k)) - g(x)| \\
&\le \left|\frac{1}{h_k}[\big(f(x_k + h_k) - f(x_k)\big) - \big(f(q_k + r_k) - f(q_k)\big)]\right| \\
&\qquad + |(\frac{1}{h_k} - \frac{1}{r_k})(f(q_k + r_k) - f(q_k))| + |\frac{1}{r_k}\big(f(q_k + r_k) - f(q_k)\big) - g(x)| \\
&\le 2\frac{1}{2^k} + \frac{1}{2^k}|f(q_k + r_k) - f(q_k)| + |\frac{1}{r_k}(f(q_k + r_k) - f(q_k)) - g(x)| \\
&\to 0, \qquad (k \to \infty),
\end{aligned}$$

and hence f is partially differentiable. Since the partial derivatives g_i are continuous, f is also totally differentiable. $\square$

7 Differentiability of quadratic Forward-Backward SDEs

In this section we will specify the results obtained in the preceding sections to BSDEs where the terminal conditions are determined by a forward SDE driven by the same Brownian motion as the BSDE. When considering BSDEs with terminal condition determined by a forward SDE we will need regularity of the forward equation. This will be guaranteed if the coefficients are functions belonging to the following space.

Throughout this section let again n be a positive integer and W a d-dimensional Brownian motion.

Definition 7.1. Let k, $m \geq 1$. We denote by $\mathbf{B^{k\times m}}$ the set of all functions $h : [0,T] \times \mathbb{R}^n \to \mathbb{R}^{k\times m}$, $(t,x) \mapsto h(t,x)$, differentiable in x, for which there exists a constant $C > 0$ such that $\sup_{(t,x)\in[0,T]\times\mathbb{R}^n} \sum_{i=1}^n \left|\frac{\partial h(t,x)}{\partial x_i}\right| \leq C$; and for all $t \in [0,T]$ we have $\sup_{x\in\mathbb{R}^n} \frac{|h(t,x)|}{1+|x|} \leq C$ and $x \mapsto \frac{\partial h(t,x)}{\partial x}$ is Lipschitz continuous with Lipschitz constant C.

With any pair $h \in \mathbf{B^{n\times 1}}$ and $\sigma \in \mathbf{B^{n\times d}}$ we associate the second order differential operator $\mathcal{L} = \sum_{i=1}^n h_i(\cdot)\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^n [\sigma\sigma^T]_{ij}(\cdot)\frac{\partial^2}{\partial x_i \partial x_j}$.

We will consider Forward-Backward SDEs (FBSDEs) of the form

$$\begin{cases} X_t^x &= x + \int_0^t b(s,X_s^x)\mathrm{d}s + \int_0^t \sigma(s,X_s^x)\mathrm{d}W_s, \qquad x \in \mathbb{R}^n, \\ Y_t^x &= g(X_T^x) + \int_t^T f(s,X_s^x,Y_s^x,Z_s^x)\mathrm{d}s - \int_t^T Z_s^x \mathrm{d}W_s, \end{cases} \tag{31}$$

where the coefficients satisfy the following assumptions:

(D1) $\sigma \in \mathbf{B^{n\times d}}$, $b \in \mathbf{B^{n\times 1}}$,

(D2) $f : \Omega \times [0,T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ are measurable functions such that $f(\omega,t,x,y,z) = l(\omega,t,x,y,z) + \alpha|z|^2$, where $l(\omega,t,x,y,z)$ is globally Lipschitz and continuously differentiable in (x,y,z),

(D3) $g : \mathbb{R}^n \to \mathbb{R}$ is a twice differentiable function such that $\nabla g \cdot \sigma \in \mathbf{B^{1\times d}}$ and $\mathcal{L}g \in \mathbf{B^{1\times 1}}$.

It follows from standard results on SDEs and from Theorem 2.3 in [9] that there exists a solution (X^x, Y^x, Z^x) of Equation (31). As we will show, the results of Section 2 imply that (X^x, Y^x, Z^x) is differentiable in x and that the derivatives $(\nabla X^x, \nabla Y^x, \nabla Z^x)$ solve the FBSDE

$$\begin{aligned} \nabla X_t^x &= 1 + \int_0^t \partial_x b(X_s^x)\nabla X_s^x \mathrm{d}s + \int_0^t \partial_x \sigma(X_s^x)\nabla X_s^x \mathrm{d}W_s, \\ \nabla Y_t^x &= \partial_x g(X_T^x)\nabla X_T^x - \int_t^T \nabla Z_s^x \mathrm{d}W_s + \int_t^T \Big[\partial_x l(s,X_s^x,Y_s^x,Z_s^x)\nabla X_s^x + \partial_y l(s,X_s^x,Y_s^x,Z_s^x)\nabla Y_s^x \\ &\quad + \partial_z l(s,X_s^x,Y_s^x,Z_s^x)\nabla Z_s^x + 2\alpha Z_s^x \nabla Z_s^x\Big]\mathrm{d}s. \end{aligned} \tag{32}$$

Our first result parallels Theorem 2.1 in which differentiability with respect to vector space topologies is treated.

Theorem 7.2. *Let (D1) and (D2) be satisfied and assume that $g : \mathbb{R}^n \to \mathbb{R}$ is bounded and differentiable. Moreover, suppose that $\frac{\partial l}{\partial x}(t,x,y,z)$ is Lipschitz continuous in x. Then for all $p \geq 2$, the function $\mathbb{R}^n \to \mathcal{R}^p(\mathbb{R}^n) \times \mathcal{R}^p(\mathbb{R}^1) \times \mathbb{L}^p(\mathbb{R}^d)$, $x \mapsto (X^x, Y^x, Z^x)$, is differentiable, and the derivative is a solution of the BSDE (32).*

Proof. By standard results, the mapping $\mathbb{R}^n \to \mathcal{R}^p(\mathbb{R}^1)$, $x \mapsto X^x$ has a continuous version (which we assume being identical to the given one), and for all $p > 1$ there exists a constant $C \in \mathbb{R}_+$ such that for x, $x' \in \mathbb{R}^n$ we have $\mathbb{E}(|X^x - X_t^{x'}|^{2p}) \leq C|x - x'|^{2p}$. See for example Lemma 4.5.4

and Lemma 4.5.6 in [10]. In order to be able to apply Theorem 2.1, we need to verify Condition (C1). For this purpose, note that

$$\begin{aligned}\mathbb{E}\left[\left(\int_0^T |l(t,X_t^x,Y_t^x,Z_t^x)-l(t,X_t^{x'},Y_t^x,Z_t^x)|\mathrm{d}t\right)^{2p}\right]&\\ \leq \mathbb{E}\left[\left(\int_0^T L|X_t^x-X_t^{x'}|\mathrm{d}t\right)^{2p}\right]&\leq \tilde{C}(L,T,C)|x-x'|^{2p},\end{aligned}$$

where L is a Lipschitz constant of l. Similarly,

$$\mathbb{E}\Big[\left(\int_0^T |\frac{\partial}{\partial x}l(t,X^x,Y^x,Z^x)-\frac{\partial}{\partial x}l(t,X^{x'},Y^x,Z^x)\mathrm{d}t|\right)^{2p}\Big]\leq C'|x-x'|^{2p}.$$

This proves (C1). Moreover, notice that $\xi(x)=g(X_T^x)$ satisfies Condition (C2). Thus the statement follows from Theorem 2.1. □

If in addition Condition (D3) is satisfied, we again obtain a sharper result stating pathwise continuous differentiability of an appropriate modification of the solution process.

Theorem 7.3. *Assume that (D1), (D2) and (D3), and suppose that the partial derivatives of l in the variables (x,y,z) are Lipschitz continuous. Then there exists a function $\Omega\times[0,T]\times\mathbb{R}^n\to\mathbb{R}^{n+1+d}$, $(\omega,t,x)\mapsto(X_t^x,Y_t^x,Z_t^x)(\omega)$, 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 (31).*

Let $M>0$ be a constant such that g, the derivatives of g, b and σ, and the partial derivatives of l in (x,y,z) are all bounded by M. For all $x\in\mathbb{R}$ let (X_t^x,Y_t^x,Z_t^x) be the solution of the FBSDE (31). To correspond formally to Theorem 2.2, in the setting of our FBSDE we have to work with

$$\tilde{l}(\omega,t,x,y,z)=l(\omega,t,X_t^x(\omega),y,z).$$

But this functional fails to be globally Lipschitz in x. This is why we have to modify slightly the proof of Theorem 2.2, and cannot just quote it. We start by showing that $\xi(x)=g(X_T^x)$ satisfies Condition (C3).

Lemma 7.4. *For all $x\in\mathbb{R}^n$, $h\neq 0$ and $i\in\{1,\dots,n\}$, let $\zeta^{x,h,i}=\frac{1}{h}(g(X_T^{x+he_i})-g(X_T^x))$. Then for every $p>1$ there exists a $C>0$, dependent only on p and M, such that for all x, $x'\in\mathbb{R}^n$ and h, $h'\neq 0$,*

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

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

$$\mathbb{E}\Big[|\zeta^{x,h,i}-g'(X_T^x)\frac{\partial}{\partial x_i}X_T^x|^{2p}\Big]\leq C(|x-x'|^2)^p.$$

Proof. Note that by Ito's formula $g(X_t^x)=g(X_0^x)+\int_0^t\nabla g(X_s^x)\cdot\sigma(s,X_s^x)dW_s+\int_0^t\mathcal{L}gds$. Thus $g(X_t^x)$ is a diffusion with coefficients $\tilde{\sigma}(s,x)=\nabla g(x)\cdot\sigma(s,x)$ and $\tilde{b}(s,x)=\sum_{i=1}^n b_i(s,x)\frac{\partial g(x)}{\partial x_i}+\frac{1}{2}\sum_{i,j=1}^n\sigma_{ij}(s,x)\frac{\partial^2 g(x)}{\partial x_i\partial x_j}$. By (D3) we have $\tilde{\sigma}\in\mathbf{B^{1\times d}}$ and $\tilde{b}\in\mathbf{B^{1\times 1}}$. Therefore, by using standard results on stochastic flows (see Lemma 4.6.3 in [10]), we obtain the result. □

Proof of Theorem 7.3. First note that it is well-known that X^x may be chosen to be continuous in t and continuously differentiable in x (see for example Theorem 39, Ch. V, [13]). In order to prove that Y^x has such a modification as well, note that Lemma 7.4 implies that $\xi(x) = g(X_T^x)$ satisfies Condition (C3). Now let again $U_t^{x,h} = \frac{1}{h}(Y_t^{x+he_i} - Y_t^x)$, $V_t^{x,h} = \frac{1}{h}(Z_t^{x+he_i} - Z_t^x)$ for all $x \in \mathbb{R}^n$ and $h \neq 0$. If $h = 0$, then define $U_t^{x,0} = \frac{\partial}{\partial x_i} Y^x$, $V_t^{x,0} = \frac{\partial}{\partial x_i} Z^x$ and $\zeta^{x,0} = \frac{\partial}{\partial x_i} g(X_T^x)$. It is enough to show that for all open bounded sets $\mathcal{O} \subset \mathbb{R}^{n+1}$ there exists a constant C such that for all $(x,h) \in \mathcal{O}$

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

As in Lemma 6.3 we will derive this estimate from Inequality (30). Notice that the assumptions of Theorem 7.3 guarantee that all the terms appearing in (30), satisfy the same properties and thus provide the same estimates. There is one essential difference which is due to the appearance of X^x instead of x in the first component of the line described by the integral $\gamma_{x,h}(\theta) = (X^x + \theta(X^{x+he_i} - X^x), Y_t^x + \theta(Y_t^{x+he_i} - Y_t^x), Z_t^x + \theta(Z_t^{x+he_i} - Z_t^x))$. We therefore have to replace the prior $A_t^{x,h}$ by $A_t^{x,h} = B_t^{x,h}\frac{1}{h}(X^{x+he_i} - X^x)$ with $B_t^{x,h} = \int_0^1 \frac{\partial l}{\partial x}(\gamma_{x,h}(\theta))\mathrm{d}\theta$. Therefore we only need to show for $(x,h) \in \mathcal{O}$

$$\mathbb{E}\big[\big(\int_0^T |A_s^{x,h} - A_s^{x',h'}|\mathrm{d}s\big)^{2pq^2}\big]^{\frac{1}{q^2}} \le c(|x-x'|^2 + |h-h'|^2)^p. \tag{33}$$

In fact, with $\Delta^{x,h} = \frac{1}{h}(X^{x+he_i} - X^x)$ if $h \neq 0$ and $\Delta^{x,0} = \nabla X^x$, we have

$$\begin{aligned}
&\mathbb{E}\big(\int_0^T |A_s^{x,h} - A_s^{x',h'}|\mathrm{d}s\big)^{2pq^2}\\
\le\ & \int_0^T |\Delta_s^{x,h}||B_s^{x,h} - B_s^{x',h'}|\mathrm{d}s + \int_0^T |\Delta_s^{x,h} - \Delta_s^{x',h'}||B_s^{x',h'}|\mathrm{d}s.
\end{aligned}$$

The first summand satisfies

$$\begin{aligned}
&\mathbb{E}\big(\int_0^T |\Delta_s^{x,h}||B_s^{x,h} - B_s^{x',h'}|\mathrm{d}s\big)^{2pq^2}\\
&\le \mathbb{E}\big(\int_0^T |\Delta_s^{x,h}| \int_0^1 |\frac{\partial l}{\partial x}(\gamma_{x,h}(\theta)) - \frac{\partial l}{\partial x}(\gamma_{x',h'}(\theta))|\mathrm{d}\theta\mathrm{d}s\big)^{2pq^2}\\
&\le \Big(\mathbb{E}\big(\int_0^T |\Delta_s^{x,h}|\mathrm{d}s\big)^{4pq^2}\Big)^{\frac{1}{2}} \Big(\mathbb{E}\big(\int_0^T\int_0^1 |\gamma_{x,h}(\theta)) - \gamma_{x',h'}(\theta)|\mathrm{d}\theta\mathrm{d}s\big)^{4pq^2}\Big)^{\frac{1}{2}}.
\end{aligned}$$

Lemma 4.6.3 in [10] implies $\sup_{(x,h)\in\mathcal{O}} \mathbb{E}\big(\int_0^T |\Delta_s^{x,h}|\mathrm{d}s\big)^{4pq^2} < \infty$. Besides,

$$\begin{aligned}
&\mathbb{E}\big(\int_0^T\int_0^1 |\gamma_{x,h}(\theta)) - \gamma_{x',h'}(\theta)|\mathrm{d}\theta\mathrm{d}s\big)^{4pq^2}\\
&\le C\,\mathbb{E}\Big(\sup_{t\in[0,T]} |X_t^{x'} - X_t^x| + \sup_{t\in[0,T]} |X_t^{x'+h'e_i} - X_t^{x+he_i}| + \sup_{t\in[0,T]} |Y_t^{x'} - Y_t^x|\\
&\quad + \sup_{t\in[0,T]} |Y_t^{x'+h'e_i} - Y_t^{x+he_i}| + \int_0^T (|Z_t^{x'} - Z_t^x| + |Z_t^{x'+h'e_i} - Z_t^{x+he_i}|)\mathrm{d}t\Big)^{4pq^2}.
\end{aligned}$$

From this we can easily deduce $\mathbb{E}\left[\left(\int_0^T |\Delta_s^{x,h}||B_s^{x,h}-B_s^{x',h'}|\mathrm{d}s\right)^{2pq^2}\right]^{\frac{1}{q^2}} \le C(|x-x'|^2+|h-h'|^2)^p$. Similarly, $\mathbb{E}\left[\left(\int_0^T |\Delta_s^{x,h}-\Delta_s^{x',h'}||B_s^{x',h'}|\mathrm{d}s\right)^{2pq^2}\right]^{\frac{1}{q^2}} \le C(|x-x'|^2+|h-h'|^2)^p$, hence (33) follows and the proof is finished. □

8 Malliavin differentiability of quadratic BSDEs

In this section we shall ask for a different type of smoothness for solutions of quadratic BSDEs, namely differentiability in the variational sense or in the sense of Malliavin's calculus. Of course, this will imply smoothness of the terminal condition in the same sense. If the terminal condition is given by a smooth function of the terminal value of a forward equation, it will also involve variational smoothness of the forward equation.

Let us first review some basic facts about Malliavin calculus. We refer the reader to [12] for a thorough treatment of the theory and to [6] for results related to BSDEs. To begin with, let $C_b^\infty(\mathbb{R}^{n\times d})$ denote the set of functions with partial derivatives of all orders defined on $\mathbb{R}^{n\times d}$ whose partial derivatives are bounded.

Let $\mathcal{S}$ denote the space of random variables ξ of the form

$$\xi = F\Big(\big(\int_0^T h_s^{1,i} dW_s^1\big)_{1\le i\le n}, \cdots, \big(\int_0^T h_s^{d,i} dW_s^d\big)_{1\le i\le n}\Big),$$

where $F \in C_b^\infty(\mathbb{R}^{n\times d})$, $h^1, \cdots, h^n \in L^2([0,T];\mathbb{R}^d)$. To simplify the notation assume that all h^j are written as row vectors.

If $\xi \in \mathcal{S}$ of the above form, we define the d-dimensional operator $D = (D^1, \cdots, D^d) : \mathcal{S} \to L^2(\Omega\times[0,T])^d$ by

$$D_\theta^i \xi = \sum_{j=1}^n \frac{\partial F}{\partial x_{i,j}}\Big(\int_0^T h_t^1 dW_t, \ldots, \int_0^T h_t^n dW_t\Big) h_\theta^{i,j}, \quad 0 \le \theta \le T, \quad 1 \le i \le d.$$

For $\xi \in \mathcal{S}$ and $p > 1$, we define the norm

$$\|\xi\|_{1,p} = \Big(\mathbb{E}\Big[\,|\xi|^p + \Big(\int_0^T |D_\theta \xi|^2 \mathrm{d}\theta\Big)^{\frac{p}{2}}\Big]\Big)^{\frac{1}{p}}.$$

It can be shown (see for example [12]) that the operator D has a closed extension to the space $\mathbb{D}^{1,p}$, the closure of $\mathcal{S}$ with respect to the norm $\|\cdot\|_{1,p}$. Observe that if ξ is $\mathcal{F}_t$−measurable then $D_\theta \xi = 0$ for $\theta \in (t,T]$.

We shall also consider n-dimensional processes depending on a time variable. We define the space $\mathbb{L}_{1,p}^a(\mathbb{R}^n)$ to be the set of $\mathbb{R}^n$−valued progressively measurable processes $u(t,\omega)_{t\in[0,T],\omega\in\Omega}$ such that

i) For a.a. $t \in [0,T]$, $u(t,\cdot) \in (\mathbb{D}^{1,p})^n$;

ii) $(t,\omega) \to D_\theta u(t,\omega) \in (L^2([0,T]))^{d\times n}$ admits a progressively measurable version;

iii) $\|u\|^a_{1,p}= \mathbb{E}[\Big(\int_0^T |u(t)|^2 \mathrm{d}t\Big)^{\frac{p}{2}} + \Big(\int_0^T\int_0^T |D_\theta u(t)|^2 \mathrm{d}\theta \mathrm{d}t\Big)^{\frac{p}{2}}]^{\frac{1}{p}} < \infty$.

Here, for $y \in \mathbb{R}^{d\times n}$ we use the norm $|y|^2 = \sum_{i,j}(y_{i,j})^2$.

We also consider the space

$$\mathbb{D}^{1,\infty} = \cap_{p>1}\mathbb{D}^{1,p}.$$

We cite for completeness a result from [12] that we will use in the next section.

Lemma 8.1 (Lemma 1.2.3 in [12]). *Let $\{F_n,\, n \geq 1\}$ be a sequence of random variables in $\mathbb{D}^{1,2}$ that converges to F in $L^2(\Omega)$ and such that*

$$\sup_{n\in\mathbb{N}^*} \mathbb{E}[\|DF_n\|_{L^2}] < \infty.$$

Then F belongs to $\mathbb{D}^{1,2}$, and the sequence of derivatives $\{DF_n,\, n \geq 1\}$ converges to DF in the weak topology of $L^2(\Omega \times [0,T])$.

Let us now consider the BSDE

$$Y_t = \xi - \int_t^T Z_s \mathrm{d}W_s + \int_t^T f(s, Y_s, Z_s)ds. \tag{34}$$

Our assumptions on driver and terminal condition this time amount to

(E1) $f : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ is an adapted measurable function such that $f(\omega, t, y, z) = l(\omega, t, y, z) + \alpha|z|^2$, where $l(\omega, t, y, z)$ is globally Lipschitz and continuously differentiable in (y, z); for all $p > 1$ we have $\mathbb{E}^P[(\int_0^T |l(\omega, t, 0, 0)|^2 \mathrm{d}s)^{2p}] < \infty$;

(E2) for all (t, y, z), the mapping $\Omega \to \mathbb{R}$, $\omega \mapsto l(\omega, t, y, z)$ is Malliavin differentiable and belongs to $\mathbb{L}^a_{1,p}(\mathbb{R})$ for all $p > 1$.

For any (ω, t, y, z) and $\theta \in [0,T]$, the (a.e. valid) inequality holds true

$$|D_\theta l(\omega, t, y, z))| \leq \tilde{K}_\theta(\omega, t) + K_\theta(\omega, t)(|y| + |z|)$$

where K_θ and $\tilde{K}_\theta$ are positive adapted processes satisfying for all $p \geq 1$

$$\mathbb{E}[\Big(\sup_{t\in[0,T]}\int_0^T |K_\theta(t,\omega)|^2 \mathrm{d}\theta\Big)^p] < \infty \quad \text{and} \quad \mathbb{E}[\Big(\int_0^T\int_0^T |\tilde{K}_\theta(t,\omega)|^2 \mathrm{d}\theta \mathrm{d}t\Big)^p] < \infty$$

(E3) the random variable ξ is bounded and belongs to $\mathbb{D}^{1,\infty}$.

We first consider the case where the terminal variable has no further structural properties, such as depending on the terminal value of a forward equation. For notational simplicity we shall treat the case of one dimensional z and Wiener process and so may omit the superscript i in D^i etc. We will this time use the typical Sobolev space approach, hidden in Lemma 8.1, to describe Malliavin derivatives, which are in fact derivatives in the distributional sense. In this approach we shall employ an approximation of the driver of our BSDE by a sequence of globally Lipschitz continuous ones, for which the properties we want to derive are known.

Let us therefore introduce a family of truncated functions starting with describing their derivatives by

$$g_n'(z) = \begin{cases} -2n & , z < -n \\ 2z & , |z| \le n \\ 2n & , z > n. \end{cases}$$

Then we have $g_n(z) = z^2$ for $|z| \le n$, $g_n(z) = 2n|z| - n^2$ for $|z| > n$, and thus $|g_n(z)| \le z^2$ and $g_n(z) \to z^2$ locally uniformly on $\mathbb{R}$ for $n \to \infty$. A similar statement holds for the derivative of $g_n(z)$: $|g_n'(z)| \le 2|z|$ and $g_n'(z) \to 2z$ locally uniformly on $\mathbb{R}$ for $n \to \infty$.

With these truncation functions we obtain the following family of BSDEs:

$$Y_t^n = \xi - \int_t^T Z_s^n dW_s + \int_t^T [l(s, Y_s^n, Z_s^n) + \alpha g_n(Z_s^n)] ds, \quad n \in \mathbb{N}^*. \tag{35}$$

From Proposition 2.4 of [9] we obtain that there exists $(Y_s, Z_s) \in \mathcal{R}^\infty(\mathbb{R}) \times \mathbb{L}^2(\mathbb{R})$ such that $Y_s^n \to Y_s$ uniformly in $[0, T]$ and $Z_s^n \to Z_s$ in $\mathbb{L}^2(\mathbb{R})$.

Since the truncated equations have Lipschitz continuous drivers, Proposition 5.3 of [6] guarantees that $(Y_t^n, Z_t^n) \in \mathbb{D}^{1,2} \times \mathbb{D}^{1,2}$ with the following Malliavin derivative

$$\begin{aligned} D_\theta Y_t^n &= 0 \quad \text{and} \quad D_\theta Z_t^n = 0, \text{ if } t \in [0, \theta), \\ D_\theta Y_t^n &= D_\theta \xi + \int_t^T \Big[\partial_y l(Y_s^n, Z_s^n) D_\theta Y_s^n + \partial_z l(Y_s^n, Z_s^n) D_\theta Z_s^n \\ &\quad + D_\theta l(s, Y_s^n, Z_s^n) + \alpha g_n'(Z_s^n) D_\theta Z_s^n \Big] \mathrm{d}s - \int_t^T D_\theta Z_s^n \mathrm{d}W_s, \qquad \text{if } t \in [\theta, T]. \end{aligned} \tag{36}$$

Now we aim at showing that the sequences DY^n and DZ^n are bounded in $\mathbb{D}^{1,2}$, in order to use Lemma 8.1. This will be done by deriving a priori estimates in the style of the preceding sections, this time uniform in n. We therefore first show boundedness relative to the auxiliary measures $Q_n := \mathcal{E}\big(\alpha \int g_n'(Z^n) dW\big) \cdot P$, in the form of the following a priori inequality.

Lemma 8.2. *Let $p > 1$. If the driver and terminal condition satisfy hypotheses (E1), (E2) and (E3), then the following inequality holds for the BSDE (36):*

$$\begin{aligned} &\mathbb{E}^{Q_n}\Big[\Big(\sup_{t\in[0,T]} \int_0^T |D_\theta Y_t^n|^2 \mathrm{d}\theta\Big)^p\Big] + \mathbb{E}^{Q_n}\Big[\Big(\int_0^T \int_0^T |D_\theta Z_s^n|^2 \mathrm{d}\theta \mathrm{d}s\Big)^p\Big] \\ &\qquad \le C \mathbb{E}^{Q_n}\Big[\Big(\int_0^T |D_\theta \xi|^2 \mathrm{d}\theta\Big)^p + \Big(\int_0^T \int_0^T |D_\theta l(s, Y_s^n, Z_s^n)|^2 \mathrm{d}\theta \mathrm{d}s\Big)^p\Big] \end{aligned}$$

Proof. We will derive these estimates by proceeding in the same fashion as for Lemma 3.2. Again, $C_1, C_2, \ldots$ are constants depending on the coefficients and p.

Applying Itô's formula to $e^{\beta t}|D_\theta Y_t^n|^2$, using Equation (36) and simplifying as we did in the former sections we obtain (choosing $\beta = M^2 + 2M$)

$$\begin{aligned} &e^{\beta t}|D_\theta Y_t^n|^2 + \int_t^T e^{\beta s}(M|D_\theta Y_s^n| - |D_\theta Z_s^n|)^2 \mathrm{d}s \\ &\qquad \le e^{\beta T}|D_\theta \xi|^2 - 2\int_t^T e^{\beta s} D_\theta Y_s^n D_\theta Z_s^n \mathrm{d}\hat{W}_s + 2\int_t^T e^{\beta s}|D_\theta Y_s^n D_\theta l(s, Y_s^n, Z_s^n)| \mathrm{d}s, \end{aligned} \tag{37}$$

where $\hat{W}_t = W_t - \int_0^t \alpha g_n'(Z_s^n)\mathrm{d}s$, $t \in [0,T]$, is a Q_n- Brownian motion.
We remark that since $\mathbb{E}\sup_{t\in[0,T]} |D_\theta Y_t^n|^2 < \infty$ we have $\mathbb{E}\Big[(\int_0^T |D_\theta Y_s^n|^2 |D_\theta Z_s^n|^2 \mathrm{d}s)^{\frac{1}{2}}\Big] < \infty$ and hence the process $\int_t^T e^{\beta s} D_\theta Y_s^n D_\theta Z_s^n \mathrm{d}\hat{W}_s$ is well defined.
From (37) we obtain by taking conditional Q_n-expectations

$$|D_\theta Y_t^n|^2 \quad \leq \quad C\mathbb{E}^{Q_n}\Big[\,|D_\theta \xi|^2 + \int_0^T |D_\theta Y_s^n|\,|D_\theta l(s, Y_s^n, Z_s^n)|\mathrm{d}s\,\big|\mathcal{F}_t\Big].$$

Next, integrating in θ, using Fubini's Theorem and Doob's L^p- inequality, we get

$$\begin{aligned}
&\mathbb{E}^{Q_n}\Big[\Big(\sup_{t\in[0,T]}\int_0^T |D_\theta Y_t^n|^2 \mathrm{d}\theta\Big)^p\Big] \\
&\leq C\mathbb{E}^{Q_n}\Big[\sup_{t\in[0,T]}\Big(\mathbb{E}^{Q}\Big[\int_0^T |D_\theta\xi|^2\mathrm{d}\theta + \int_0^T\int_0^T |D_\theta Y_s^n|\,|D_\theta l(s, Y_s^n, Z_s^n)|\mathrm{d}\theta\mathrm{d}s\,\big|\mathcal{F}_t\Big]\Big)^p\Big] \\
&\leq C\mathbb{E}^{Q_n}\Big[\Big(\int_0^T |D_\theta\xi|^2\mathrm{d}\theta\Big)^p + \Big(\int_0^T\int_0^T |D_\theta Y_s^n|\,|D_\theta l(s, Y_s^n, Z_s^n)|\mathrm{d}\theta\mathrm{d}s\Big)^p\Big].
\end{aligned}$$

The last term on the right hand side of the preceding inequality can be simplified using Hölder's and Young's inequalities with the result

$$\begin{aligned}
&\int_0^T\int_0^T |D_\theta Y_s^n|\,|D_\theta l(s, Y_s^n, Z_s^n)|\mathrm{d}\theta\mathrm{d}s \\
&\quad\leq \int_0^T \Big[\Big(\int_0^T |D_\theta Y_s^n|^2\mathrm{d}\theta\Big)^{\frac{1}{2}}\Big(\int_0^T |D_\theta l(s, Y_s^n, Z_s^n)|^2\mathrm{d}\theta\Big)^{\frac{1}{2}}\Big]\mathrm{d}s \\
&\quad\leq \sup_{t\in[0,T]}\Big(\int_0^T |D_\theta Y_t^n|^2\mathrm{d}\theta\Big)^{\frac{1}{2}}\int_0^T\Big(\int_0^T |D_\theta l(s, Y_s^n, Z_s^n)|^2\mathrm{d}\theta\Big)^{\frac{1}{2}}\mathrm{d}s \\
&\quad\leq \frac{1}{C_1}\sup_{t\in[0,T]}\int_0^T |D_\theta Y_t^n|^2\mathrm{d}\theta + C_2\int_0^T\int_0^T |D_\theta l(s, Y_s^n, Z_s^n)|^2\mathrm{d}\theta\mathrm{d}s.
\end{aligned}$$

Since for $a, b \geq 0$ we have $(a+b)^p \leq C_3(a^p + b^p)$, by choosing C_1 conveniently we obtain

$$\begin{aligned}
&\mathbb{E}^{Q_n}\Big[\Big(\sup_{t\in[0,T]}\int_0^T |D_\theta Y_t^n|^2\mathrm{d}\theta\Big)^p\Big] \\
&\qquad\leq C_4\mathbb{E}^{Q_n}\Big[\Big(\int_0^T |D_\theta\xi|^2\mathrm{d}\theta\Big)^p + \Big(\int_0^T\int_0^T |D_\theta l(s, Y_s^n, Z_s^n)|^2\mathrm{d}s\mathrm{d}\theta\Big)^p\Big], \qquad (38)
\end{aligned}$$

which provides the desired bound for the part of the Malliavin derivatives of Y^n. Concerning the inequality for the Malliavin derivatives of the Z^n part, we consult again Equation (37), from which we derive

$$\begin{aligned}
\int_0^T e^{\beta s}|D_\theta Z_s^n|^2\mathrm{d}s \quad &\leq \quad e^{\beta T}|D_\theta\xi|^2 - 2\int_0^T e^{\beta s} D_\theta Y_s^n D_\theta Z_s^n \mathrm{d}\hat{W}_s \\
&\qquad +2\int_0^T e^{\beta s}|D_\theta Y_s^n||D_\theta l(s, Y_s^n, Z_s^n)|\mathrm{d}s + 2\int_0^T M e^{\beta s}|D_\theta Y_s^n||D_\theta Z_s^n|\mathrm{d}s.
\end{aligned}$$

Further estimate

$$2\int_0^T Me^{\beta s}|D_\theta Y_s^n||D_\theta Z_s^n|\mathrm{d}s \leq 4M^2\int_0^T e^{\beta s}|D_\theta Y_s^n|^2\mathrm{d}s + \frac{1}{2}\int_0^T e^{\beta s}|D_\theta Z_s^n|^2\mathrm{d}s,$$

$$2\int_0^T e^{\beta s}|D_\theta Y_s^n||D_\theta l(s,Y_s^n,Z_s^n)|\mathrm{d}s \leq \int_0^T e^{\beta s}|D_\theta Y_s^n|^2\mathrm{d}s + \int_0^T e^{\beta s}|D_\theta l(t,Y_s^n,Z_s^n)|^2\mathrm{d}s.$$

Hence the initial estimate leads to

$$\begin{aligned}\frac{1}{2}\int_0^T e^{\beta s}|D_\theta Z_s^n|^2\mathrm{d}s \leq\ & e^{\beta T}|D_\theta\xi|^2 - 2\int_0^T e^{\beta s}D_\theta Y_s^n D_\theta Z_s^n\mathrm{d}\hat{W}_s \\ & +(1+4M^2)\int_0^T e^{\beta s}|D_\theta Y_s^n|^2\mathrm{d}s + \int_0^T e^{\beta s}|D_\theta l(s,Y_s^n,Z_s^n)|^2\mathrm{d}s.\end{aligned}$$

Now for $p>1$ integrate in θ, take Q_n-expectations, using Fubini's Theorem as well as a stochastic version of it to estimate

$$\begin{aligned}&\mathbb{E}^{Q_n}\Big[\Big(\int_0^T\int_0^T|D_\theta Z_s^n|^2\mathrm{d}\theta\mathrm{d}s\Big)^p\Big] \\ &\leq C_5\Big\{\mathbb{E}^{Q_n}\Big[\Big(\int_0^T|D_\theta\xi|\mathrm{d}\theta\Big)^p + \Big(\sup_{t\in[0,T]}\int_0^T|D_\theta Y_t^n|^2\mathrm{d}\theta\Big)^p \\ &+\Big(\int_0^T\int_0^T|D_\theta l(s,Y_s^n,Z_s^n)|^2\mathrm{d}\theta\mathrm{d}s\Big)^p + \Big(\int_0^T\int_0^T D_\theta Y_s^n\, D_\theta Z_s^n\mathrm{d}\theta\mathrm{d}\hat{W}_s\Big)^p\Big]\Big\}. \qquad (39)\end{aligned}$$

We estimate the last term using Burkholder-Davis-Gundy's inequality, which results in

$$\mathbb{E}^{Q_n}\Big[\Big(\int_0^T\int_0^T D_\theta Y_s^n D_\theta Z_s^n\mathrm{d}\theta\mathrm{d}\hat{W}_s\Big)^p\Big] \leq C_6\mathbb{E}^{Q_n}\Big[\Big(\int_0^T\Big\{\int_0^T|D_\theta Y_s^n||D_\theta Z_s^n|\mathrm{d}\theta\Big\}^2\mathrm{d}s\Big)^{\frac{p}{2}}\Big].$$

Using Cauchy-Schwarz' inequality, we estimate further by

$$\int_0^T|D_\theta Y_s^n||D_\theta Z_s^n|\mathrm{d}\theta \leq \Big(\int_0^T|D_\theta Y_s^n|^2\mathrm{d}\theta\Big)^{\frac{1}{2}}\Big(\int_0^T|D_\theta Z_s^n|^2\mathrm{d}\theta\Big)^{\frac{1}{2}}.$$

Then, with another application of Young's inequality, we obtain

$$\begin{aligned}&\mathbb{E}^{Q_n}\Big[\Big(\int_0^T\Big\{\int_0^T|D_\theta Y_s^n||D_\theta Z_s^n|\mathrm{d}\theta\Big\}^2\mathrm{d}s\Big)^{\frac{p}{2}}\Big] \\ &\leq \mathbb{E}^{Q_n}\Big[\Big(\int_0^T\Big[\Big\{\int_0^T|D_\theta Y_s^n|^2\mathrm{d}\theta\Big\}^{\frac{1}{2}}\Big\{\int_0^T|D_\theta Z_s^n|^2\mathrm{d}\theta\Big\}^{\frac{1}{2}}\Big]^2\mathrm{d}s\Big)^{\frac{p}{2}}\Big] \\ &\leq \mathbb{E}^{Q_n}\Big[\Big\{\sup_{t\in[0,T]}\int_0^T|D_\theta Y_t^n|^2\mathrm{d}\theta\Big\}^{\frac{p}{2}}\Big\{\int_0^T\int_0^T|D_\theta Z_s^n|^2\mathrm{d}\theta\mathrm{d}s\Big\}^{\frac{p}{2}}\Big] \\ &\leq \frac{1}{C_7}\mathbb{E}^{Q_n}\Big[\Big\{\sup_{t\in[0,T]}\int_0^T|D_\theta Y_t^n|^2\mathrm{d}\theta\Big\}^p\Big] + C_7\mathbb{E}^{Q_n}\Big[\Big\{\int_0^T\int_0^T|D_\theta Z_s^n|^2\mathrm{d}\theta\mathrm{d}s\Big\}^p\Big].\end{aligned}$$

Using this last estimate with C_7 chosen properly in conjunction with (38) in (39), we obtain

$$\begin{aligned}&\mathbb{E}^{Q_n}\Big[\Big(\int_0^T\int_0^T|D_\theta Z_s^n|^2\mathrm{d}\theta\mathrm{d}s\Big)^p\Big] \\ &\quad\leq C_8\mathbb{E}^{Q_n}\Big[\Big(\int_0^T|D_\theta\xi|\mathrm{d}\theta\Big)^p + \Big(\int_0^T\int_0^T|D_\theta l(s,Y_s^n,Z_s^n)|^2\mathrm{d}\theta\mathrm{d}s\Big)^p\Big]. \qquad (40)\end{aligned}$$

Combining inequalities (38) and (40) yields the desired estimate and proves the Lemma. $\square$

In the same fashion as in Section 3, we can now combine the result of the a priori inequality under $Q_n, n \in \mathbb{N}^*$, with the inverse Hölder inequality in disguise of Lemma 1.2 to upgrade the a priori estimates to the following one. In fact, we observe $|g_n'(z)| \leq 2|z|$ for $z \in \mathbb{R}, n \in \mathbb{N}^*$. Moreover, a careful analysis of the demonstration of Lemma 1 of [11] shows that each $\int Z^n dW$ is also BMO and there exists a constant K such that

$$\sup_{n\in\mathbb{N}^*} \| \int g_n'(Z^n)dW \|_{BMO_2} \leq \sup_{n\in\mathbb{N}^*} \| \int Z^n dW \|_{BMO_2} + \| \int ZdW \|_{BMO_2} = K < \infty. \tag{41}$$

So by Lemma 1.2 there exists a $1 < r$ such that $\mathcal{E}(\int \alpha g_n'(Z^n)dW)_T$ and $\mathcal{E}(\int \alpha ZdW)_T$ are in $L^r(P)$ for all $n \in \mathbb{N}^*$ with r-norms bounded in n. So, again we may apply the argument based on the third statement of Lemma 1.2, uniformly in n. This, together with a similar argument applied to the sequence (Y^n, Z^n) leads to the following a priori estimate.

Lemma 8.3. *Let $p > 1$ and $r > 1$ such that $\mathcal{E}(\int \alpha g_n'(Z^n)dW)_T \in L^r(P)$ for all $n \in \mathbb{N}^*$ with a uniform bound. Then there exists a constant $C > 0$, depending only on p, T and K (from 41), such that with the conjugate exponent q of r we have*

$$\begin{aligned}
&\mathbb{E}^P\Big[\int_0^T\int_0^T |D_\theta Y_t^n|^2 \mathrm{d}\theta\mathrm{d}t\Big]^p + \mathbb{E}^P\Big[\int_0^T\int_0^T |D_\theta Z_s^n|^2\mathrm{d}\theta\mathrm{d}s\Big]^p \\
&\leq C\Big\{\mathbb{E}^P\Big[\Big(\int_0^T |D_\theta\xi|^2\mathrm{d}\theta + \int_0^T\int_0^T |\tilde{K}_\theta(\omega,t)|^2\mathrm{d}\theta\mathrm{d}t\Big)^{pq^2}\Big]^{\frac{1}{q^2}} \\
&+\mathbb{E}^P\Big[\Big(\sup_{t\in[0,T]}\int_0^T |K_\theta(\omega,t)|^2\mathrm{d}\theta\Big)^{2pq^2}\Big]^{\frac{1}{2q^2}}\mathbb{E}^P\Big[\Big(|\xi|^2 + \int_0^T |l(\omega,t,0,0)|^2\mathrm{d}s\Big)^{4pq^4}\Big]^{\frac{1}{2q^4}}\Big\} < \infty.
\end{aligned}$$

Proof. The proof of the lemma is achieved in three steps.

We start by applying the third statement of Lemma 1.2 as in the proof of Theorem 3.1 to the result of Lemma 8.2, from which we obtain with a constant C_1 not depending on n

$$\begin{aligned}
&\mathbb{E}^P\Big[\int_0^T\int_0^T |D_\theta Y_t^n|^2\mathrm{d}\theta\mathrm{d}t\Big]^p + \mathbb{E}^P\Big[\int_0^T\int_0^T |D_\theta Z_s^n|^2\mathrm{d}\theta\mathrm{d}s\Big]^p \\
&\quad\leq C_1\mathbb{E}^P\Big[\Big(\int_0^T |D_\theta\xi|^2\mathrm{d}\theta + \int_0^T\int_0^T |D_\theta l(s,Y_s^n,Z_s^n)|^2\mathrm{d}\theta\mathrm{d}s\Big)^{pq^2}\Big]^{\frac{1}{q^2}}.
\end{aligned}$$

In a second step, we have to estimate the last term of the preceding equation. From Condition (E2) we obtain with another universal constant

$$\begin{aligned}
&\int_0^T\int_0^T |D_\theta l(s,Y_s^n,Z_s^n)|^2\mathrm{d}\theta\mathrm{d}t \\
&\leq C_2\int_0^T\int_0^T |[\tilde{K}_\theta(\omega,t)|^2 + |K_\theta(\omega,t)|^2(|Y_t^n|^2 + |Z_t^n|^2)]\mathrm{d}\theta\mathrm{d}t \\
&\leq C_2\Big\{\int_0^T\int_0^T |\tilde{K}_\theta(\omega,t)|^2\mathrm{d}\theta\mathrm{d}t + \sup_{t\in[0,T]}\int_0^T |K_\theta(\omega,t)|^2\mathrm{d}\theta\int_0^T |Y_s^n|^2 + |Z_s^n|^2\mathrm{d}s\Big\}.
\end{aligned}$$

Hence an application of Hölder's inequality results in

$$\mathbb{E}[\Big(\int_0^T\int_0^T |D_\theta l(s,Y_s^n,Z_s^n)|^2\mathrm{d}\theta\mathrm{d}s\Big)^{pq^2}] \le C_3\mathbb{E}[\Big(\int_0^T\int_0^T |\tilde{K}_\theta(\omega,t)|^2\mathrm{d}\theta\mathrm{d}s\Big)^{pq^2}]$$
$$+C_4\mathbb{E}[\Big(\sup_{t\in[0,T]}\int_0^T |K_\theta(\omega,t)|^2\mathrm{d}\theta\Big)^{2pq^2}]^{\frac{1}{2}}\mathbb{E}[\Big(\int_0^T |Y_s^n|^2+|Z_s^n|^2\mathrm{d}s\Big)^{2pq^2}]^{\frac{1}{2}}. \qquad (42)$$

In a last step, we need to provide a bound for the $\mathbb{E}[\Big(\int_0^T |Y_s^n|^2+|Z_s^n|^2\mathrm{d}s\Big)^{2pq^2}]$ term. For this purpose, we shall use another application of Theorem 3.1, uniformly in n. It requires the intervention of a different family of measure changes depending on n, which can again be controlled by the BMO property of the intervening martingales and the third statement of Lemma 1.2. In fact, comparing (35) with (5), we see that the analogue of H has to be given by $h_n(Z^n)$, where

$$h_n(z) := \frac{g_n(z)}{z} = \begin{cases} z & ,\text{if } |z|\le n \\ \frac{2n|z|-n^2}{z} & ,\text{if } |z|>n, \end{cases}$$

which is obviously well defined for all z. In this situation, the stochastic integrals of $h_n(Z^n)$ generate BMO martingales with uniformly bounded norms. More precisely, since $g_n(z)\le z^2$, we have $\sup_{n\in\mathbb{N}^*}|h_n(z)|\le |z|$, $z\in\mathbb{R}$. A careful analysis of the demonstration of Lemma 1 of [11] shows that each Z^n is also BMO and there exists a constant K such that

$$\sup_{n\in\mathbb{N}^*}\|\int Z^n dW\|_{BMO_2}+\|\int Z dW\|_{BMO_2}= K<\infty.$$

Due to the definition of h_n, we may extend (41) to

$$\sup_{n\in\mathbb{N}^*}\|\int h_n(Z^n)\mathrm{d}W\|_{BMO_2}\le \sup_{n\in\mathbb{N}^*}\|\int Z^n\mathrm{d}W\|_{BMO_2}\le K.$$

By Lemma 1.2 we may assume that $\mathcal{E}(\int \alpha h_n(Z^n)\mathrm{d}W)_T$ and $\mathcal{E}(\int\alpha Z\mathrm{d}W)_T$ are in $L^r(P)$ for all $n\in\mathbb{N}^*$ with r-norms bounded in n. So, again we may apply the argument based on the third statement of Lemma 1.2, uniformly in n. We obtain from Theorem 3.1 with the settings $A_t = l(\omega,t,0,0)$, $\tilde{l}(\omega,t,Y_t^n,Z_t^n) := l(\omega,t,Y_t^n,Z_t^n)-l(\omega,t,0,0)$, $\zeta=\xi$ and $H_t = h_n(Z_t^n)$, for all $\gamma>1$ the inequality

$$\sup_{n\in\mathbb{N}^*}\mathbb{E}^P\Big[\Big(\sup_{t\in[0,T]}|Y_t^n|^2+\int_0^T |Z_s^n|^2\mathrm{d}s\Big)^\gamma\Big]\le C_4\mathbb{E}^P\Big[\Big(|\xi|^2+\int_0^T |l(\omega,s,0,0)|^2\mathrm{d}s\Big)^{\gamma q^2}\Big]^{\frac{1}{q^2}}. \qquad (43)$$

Plugging this inequality into (42) terminates the proof of the Lemma. □

Our main result can now be proved.

Theorem 8.4. *Assume that driver and terminal condition satisfy hypotheses (E1), (E2) and (E3). Then the solution processes* (Y_t,Z_t)*, of (34) belongs to* $\mathbb{D}^{1,2}\times(\mathbb{D}^{1,2})^d$ *and a version of* $(D_\theta Y_t, D_\theta Z_t)$ *satisfies for* $1\le i\le d$

$$\begin{aligned} D_\theta^i Y_t &= 0, \qquad D_\theta^i Z_t = 0, \qquad t\in[0,\theta), \\ D_\theta^i Y_t &= D_\theta^i\xi + \int_t^T\Big[\partial_y l(s,Y_s,Z_s)D_\theta^i Y_s + \partial_z l(s,Y_s,Z_s)D_\theta^i Z_s + D_\theta^i l(s,Y_s,Z_s) + 2\alpha Z_s D_\theta^i Z_s\Big]\mathrm{d}s \\ &\quad -\int_t^T D_\theta^i Z_s\mathrm{d}W_s, \qquad t\in[\theta,T]. \end{aligned} \qquad (44)$$

Moreover, $\{D_t Y_t : 0 \le t \le T\}$ *is a version of* $\{Z_t : 0 \le t \le T\}$.

Proof. Again, we simplify notation by just considering the case of one dimensional Z and Wiener process. We first apply Lemma 8.3 to obtain a bound of the L^2-norms of the processes $(D_\theta Y_t^n, D_\theta Z_t^n)$, which is uniform in n. Lemma 8.3 furthermore allows an appeal to a weak compactness result to deduce the existence of a pair of processes $(U_{\theta,t}, V_{\theta,t}), 0 \le \theta, t \le T$, and a subsequence (n_i) such that $(D_\theta Y_t^{n_i}(\omega), D_\theta Z_t^{n_i}(\omega))$ converges to $(U_{\theta,t}(\omega), V_{\theta,t}(\omega))$ in the weak topology of the L^2 space of random variables with values in $L^2([0,T] \times [0,T])$.

For almost all t, Lemma 8.1 implies that (Y_t, Z_t) is Malliavin differentiable and the equality $(D_\theta Y_t, D_\theta Z_t) = (U_{\theta,t}, V_{\theta,t})$ holds almost everywhere in $\Omega \times [0,T]$.

It remains to use these convergence properties to deduce convergence term by term in (36) to (44).

We first show that the stochastic integral terms converge weakly in $L^2(\Omega)$. To this end let $\Psi \in L^2(\Omega)$ be $\mathcal{F}_T$-measurable. Then there exists a predictable $\psi \in L^2(\Omega \times [0,T])$ with $\Psi = \mathbb{E}(\Psi) + \int_0^T \psi_s dW_s$, and hence

$$\begin{aligned} \lim_{i\to\infty} \mathbb{E}\left[\Psi \int_0^T D_\theta Z_s^{n_i} dW_s\right] &= \lim_{i\to\infty} \mathbb{E}\left[\int_0^T \psi_s D_\theta Z_s^{n_i} ds\right] = \mathbb{E}\left[\int_0^T \psi_s D_\theta Z_s ds\right] \\ &= \mathbb{E}\left[\Psi \int_0^T D_\theta Z_s dW_s\right], \end{aligned}$$

which shows that $\int_0^T D_\theta Z_s^{n_i} dW_s$ converges weakly to $\int_0^T D_\theta Z_s dW_s$ in $L^2(\Omega \times [0,T])$.

Next observe that for any bounded $\mathcal{F}_T$-measurable random variable B we have

$$\begin{aligned} &\mathbb{E}[B \int_0^T (g'_{n_i}(Z_t^{n_i}) D_\theta Z_t^{n_i} - g'(Z_t) D_\theta Z_t) \mathrm{d}t] \\ &\quad = \mathbb{E}[B \int_0^T (g'_{n_i}(Z_t^{n_i}) - g'(Z_t)) D_\theta Z_t^{n_i} \mathrm{d}t] + \mathbb{E}[B \int_0^T g'(Z_t)(D_\theta Z_t^{n_i} - D_\theta Z_t) \mathrm{d}t] \end{aligned} \tag{45}$$

The first summand on the RHS of Equation (45) is bounded by

$$(\text{esssup } |B|) \ \sup_i \left(\mathbb{E} \int_0^T D_\theta (Z_t^{n_i})^2 \mathrm{d}t\right)^{\frac{1}{2}} \left(\mathbb{E}[\int_0^T (g'_{n_i}(Z_t^{n_i}) - g'(Z_t))^2 \mathrm{d}t]\right)^{\frac{1}{2}},$$

which converges to 0 as $i \to \infty$. The second summand on the RHS of Equation (45) converges also to 0 since $Bg'(Z_t) \in L^2(\Omega \times [0,T])$ and $D_\theta Z_t^{n_i}$ converges weakly to $D_\theta Z_t$. Since B was arbitrary we have shown that $\int_0^T g'_{n_i}(Z_t^{n_i}) D_\theta Z_t^{n_i} \mathrm{d}t$ converges to $\int_0^T g'(Z_t) D_\theta Z_t \mathrm{d}t$ in the weak topology of $L^1(\Omega \times [0,T])$.

Finally we come to the various derivative terms of l. The boundedness of the partial derivatives of l as well as Condition (E2) and Inequality (43) imply

$$\begin{aligned}
& \sup_n \mathbb{E}\int_0^T \int_0^T [D_\theta(l(s,Y_s^n,Z_s^n))]^2 \mathrm{d}s\,\mathrm{d}\theta \\
\leq\ & C_1 \sup_n \mathbb{E}\int_0^T \int_0^T |\partial_y l(Y_s^n,Z_s^n)D_\theta Y_s^n|^2 + |\partial_z l(Y_s^n,Z_s^n)D_\theta Z^n|^2 + [(D_\theta l)(s,Y_s^n,Z_s^n)]^2 \mathrm{d}s\,\mathrm{d}\theta \\
\leq\ & C_2 \Big\{ \sup_n \mathbb{E}\int_0^T \int_0^T |D_\theta Y_s^n|^2 + |D_\theta Z_s^n|^2 + |\tilde{K}_\theta(s)|^2 \mathrm{d}s\,\mathrm{d}\theta \\
& \qquad + \sup_n \mathbb{E}\left[\int_0^T \sup_{s\in[0,T]} |K_\theta(s)|^2 \mathrm{d}\theta \int_0^T (|Y_s^n| + |Z_s^n|)^2 \mathrm{d}s\right] \Big\} < \infty.
\end{aligned}$$

Thus, by Lemma 8.1, for almost all $s \in [0,T]$, $l(s,Y_s,Z_s)$ belongs to $\mathbb{D}^{1,2}$, and $D_\theta(l(s,Y_s^n,Z_s^n))$ converges to $D_\theta(l(s,Y_s,Z_s))$ weakly in $L^2(\Omega \times [0,T])$. Since the partial derivatives of l are continuous we have $D_\theta(l(s,Y_s,Z_s)) = \partial_y l(Y_s,Z_s)D_\theta Y_s + \partial_z l(Y_s,Z_s)D_\theta Z_s + (D_\theta l)(s,Y_s,Z_s)$. □

We next assume more structural properties for the terminal variable. More precisely, we will turn to the framework of forward-backward systems. Given a $d-$dimensional Brownian motion W and an $x \in \mathbb{R}^n$, we denote by $X_t = (X_t^1, \cdots, X_t^n)$ the forward part given by

$$X_t = x + \int_0^t b(s,X_s)ds + \int_0^t [\sigma(s,X_s)]^* \mathrm{d}W_s. \tag{46}$$

The coefficients are supposed to satisfy

(P1) $b, \sigma_i : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$, are uniformly Lipschitz; $b(\cdot,0)$ and $\sigma_i(\cdot,0)$ are bounded for $1 \leq i \leq d$; $\sigma(t,X_t)$ is a $d \times n$ matrix and $[\sigma(t,X_t)]^*$ represents its transpose.

For the backward part we consider

$$Y_t = g(X_T) - \int_t^T Z_s dW_s + \int_t^T [l(s,X_s,Y_s,Z_s) + \alpha|Z_s|^2]ds, \qquad t \in [0,T] \tag{47}$$

where the driver and the terminal conditions are supposed to satisfy the following assumptions

(P2) $l : [0,T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ is globally Lipschitz and continuously differentiable in (x,y,z); $l(\cdot,0,0,0)$ and $\partial_x l(t,x,y,z)$ are bounded by a constant M.

(P3) $g : \mathbb{R}^n \to \mathbb{R}$ is a bounded differentiable function with bounded first partial derivatives.

Before stating the main theorem we recall that the processes X, Y and Z all depend on the variable x. For ease of notation we omit the corresponding superscripts. In this setting our main result is the following.

Theorem 8.5. *Suppose that the coefficients of the SDE (46) and the driver and terminal condition of the BSDE (47) satisfy conditions (P1), (P2) and (P3). Then the solution processes (X,Y,Z) possess the following properties.*

- *For any* $0 \le t \le T$, $x \in \mathbb{R}$, $(Y_t, Z_t) \in \mathbb{D}^{1,2} \times \left(\mathbb{D}^{1,2}\right)^d$, *and a version of* $\{(D^i_\theta Y_t, D^i_\theta Z_t); 0 \le \theta, t \le T\}$ *satisfies for* $1 \le i \le d$

$$\begin{aligned} D^i_\theta Y_t &= 0, \qquad D^i_\theta Z_t = 0, \qquad t \in [0,\theta), \\ D^i_\theta Y_t &= \partial_x g(X_T) D^i_\theta X_T + \int_t^T \Big[\partial_x l(s, Y_s, Z_s) D^i_\theta X_s + \partial_y l(Y_s, Z_s) D^i_\theta Y_s \\ &\quad + \partial_z l(Y_s, Z_s) D^i_\theta Z_s + 2\alpha Z_s D^i_\theta Z_s\Big] \mathrm{d}s - \int_t^T \langle D^i_\theta Z_s, \mathrm{d}W_s\rangle, \qquad t \in [\theta, T]. \end{aligned} \tag{48}$$

Moreover, $\{D_t Y_t; 0 \le t \le T\}$ *defined by the above equation is a version of* $\{Z_t; 0 \le t \le T\}$.

- *The following set of equations holds for any* $0 \le \theta \le t \le T$ *and* $x \in \mathbb{R}^n$, P*-almost surely,*

$$\begin{aligned} D_\theta X_t &= \partial_x X_t (\partial_x X_\theta)^{-1} \sigma(\theta, X_\theta) \\ D_\theta Y_t &= \partial_x Y_t (\partial_x X_\theta)^{-1} \sigma(\theta, X_\theta) \\ Z_t &= \partial_x Y_t (\partial_x X_t)^{-1} \sigma(s, X_t); \end{aligned}$$

and $D_\theta Z_t = \partial_x Z_t (\partial_x X_\theta)^{-1} \sigma(\theta, X_\theta)$ *for almost all* (ω, t).

Proof. Theorem 2.2.1 of [12] assures existence, uniqueness and Malliavin differentiability of solutions of SDE (46) under Hypothesis (P1). Moreover the solution processes satisfy $X_t \in (\mathbb{D}^{1,\infty})^n$ for any $t \in [0,T]$ and $1 \le i \le d$ and the following equation holds:

$$\begin{aligned} D^i_\theta X_t &= 0, && t \in [0,\theta), \\ D^i_\theta X_t &= \sigma(\theta, X_\theta) + \int_\theta^t \partial_x b(s, X_s) D^i_\theta X_s \mathrm{d}s + \int_\theta^t \partial_x \sigma(s, X_s) D^i_\theta X_s \mathrm{d}W_s, && t \in [\theta, T], \end{aligned}$$

(see f.ex. Theorem 2.2.1 of [12]).

Let us next check the validity of hypotheses (E1)-(E3) for the driver of our BSDE, for simplicity in the one dimensional case. From Condition (P3) it follows that the function g and its derivative are bounded. In combination with the fact that $X_t \in \mathbb{D}^{1,\infty}$ this implies that $g(X_T) \in \mathbb{D}^{1,\infty}$, i.e. in the setting of Theorem 8.4 Condition (E3) is verified. From Condition (P2), we have $|l(t, X_t, 0, 0)| \le M(1 + |X_t|)$ for $t \in [0,T]$. The fact that $X_t \in \mathbb{D}^{1,\infty}$ then entails $\mathbb{E}[\Big(\int_0^T |l(t, X_t, 0, 0)|^2 \mathrm{d}s\Big)^p] < \infty$ for all $p \ge 1$. Hence (E1) is satisfied. Condition (P2) includes the statement that $\partial_x l(\cdot, x, \cdot, \cdot)$ is bounded. Therefore we have

$$|D_\theta l(t, X_t, y, z)| = |\partial_x l(t, X_t, y, z) D_\theta X_t| \le M |D_\theta X_t|$$

with some constant M. Using the fact that $X_t \in \mathbb{D}^{1,\infty}$ we obtain $\mathbb{E}[\Big(\int_0^T \int_0^T |M\, D_\theta X_t|^2 \mathrm{d}\theta \mathrm{d}t\Big)^p] < \infty$ for $p > 1$, which means Condition (E2) also holds.

With conditions (E1), (E2) and (E3) verified we can apply Theorem 8.4, which implies the Malliavin differentiability of (Y_s, Z_s) and proves the first block of results.

For the second part of the theorem, the representation formula of DX is standard (see f.ex. Chapter 2.3 in [12]). The representation of Z by the trace of DY being granted, we only have to prove the representation formulas for DY and DZ. For this purpose, we apply Itô's formula to $\partial_x Y_t (\partial_x X_\theta)^{-1} \sigma(\theta, X_\theta)$, then use (4) to represent the $\partial_x Y_t$ term. We further use the representation of DX to account for the terminal condition. This way we obtain (48) with $D_\theta Y_t = \partial_x Y_t(\partial_x X_\theta)^{-1}\sigma(\theta, X_\theta)$ and $D_\theta Z_t = \partial_x Z_t(\partial_x X_\theta)^{-1}\sigma(\theta, X_\theta)$. The representation follows from uniqueness of solutions for the BSDE. $\square$

Example:
We finally study a specific setting of Theorem 8.5. We assume that ξ and $D\xi$ are bounded by M. Assume further that the driver does not depend on y. Then, choosing $\theta = t$, representation (48) can be simplified to

$$Z_t \;=\; D_t\xi + \int_t^T \Big[\partial_z l(Z_s)D_tZ_s + 2\alpha Z_s D_t Z_s\Big]\mathrm{d}s - \int_t^T D_tZ_s\mathrm{d}W_s.$$

Since $\int ZdW \in BMO$ and if we further assume $\int \partial_z l(Z)dW \in BMO$ we may change the measure to $Q = \mathcal{E}\Big(\int \partial_z l(Z) + 2\alpha Z dW_T\Big)\cdot P$. Hence we obtain, by applying conditional expectations

$$Z_t = \mathbb{E}^Q[D_t\xi|\mathcal{F}_t] \le M\mathbb{E}^Q[1|\mathcal{F}_t] \le M.$$

This means $Z \in L^\infty$. This way we recover the Malliavin differentiability results of [7] from our main result.

Remarks:

1. The methods of proof of this Section, building upon a truncated sequence of Lipschitz BSDEs, could also be used in the treatment of the differentiability problem in Section 6. This sequence would allow the use of the results in [6], which, combined with the a priori estimates of sections 3 and 4 would imply differentiability.

2. Our main results allow less restrictive hypotheses. For example in Section 3, we assume for our a priori estimations that $\zeta \in L^p$ for all $p \ge 1$. An analysis of the proof clearly reveals that to obtain estimates in $\mathcal{R}^p$ or $\mathbb{L}^p$ we only need that $\zeta \in L^p$ for all $p \in (2, 2pq^2]$. We chose to write $\zeta \in L^p$ for all $p \ge 1$ not to produce an overload of technicalities in a technically already rather complex text.

Acknowledgement: We thank the anonymous referee for a very careful reading and for many helpful remarks.

References

[1] S. Ankirchner, P. Imkeller, and A. Popier, *Optimal cross hedging of insurance derivatives*, 2005. Preprint.

[2] P. Briand and F. Confortola, *BSDEs with stochastic Lipschitz condition and quadratic PDEs in hilbert spaces*, 2007. To appear in Stochastic Process. Appl.

[3] S. Chaumont, P. Imkeller, M. Müller, and U. Horst, *A simple model for trading climate risk*, Vierteljahrshefte zur Wirtschaftsforschung / Quarterly Journal of Economic Research **74** (2005), no. 2, 175–195. available at http://ideas.repec.org/a/diw/diwvjh/74-2-6.html.

[4] S. Chaumont, P. Imkeller, and M. Müller, *Equilibrium trading of climate and weather risk and numerical simulation in a Markovian framework*, Stoch. Environ. Res. Risk Assess. **20** (2006), no. 3, 184–205.MR2297447

[5] P. Cheridito, H. M. Soner, N. Touzi, and N. Victoir, *Second-order backward stochastic differential equations and fully nonlinear parabolic PDEs*, Comm. Pure Appl. Math. **60** (2007), no. 7, 1081–1110.MR2319056

[6] N. El Karoui, S. Peng, and M. C. Quenez, *Backward stochastic differential equations in finance*, Math. Finance **7** (1997), no. 1, 1–71.MR1434407

[7] U. Horst and M. Müller, *On the spanning property of risk bonds priced by equilibrium*, 2006. To appear in Mathematics of Operations Research.

[8] N. Kazamaki, *Continuous exponential martingales and BMO*, Lecture Notes in Mathematics, vol. 1579, Springer-Verlag, Berlin, 1994.MR1299529

[9] M. Kobylanski, *Backward stochastic differential equations and partial differential equations with quadratic growth*, Ann. Probab. **28** (2000), no. 2, 558–602.MR1782267

[10] H. Kunita, *Stochastic flows and stochastic differential equations*, Cambridge Studies in Advanced Mathematics, vol. 24, Cambridge University Press, Cambridge, 1990.MR1070361

[11] M. A. Morlais, *Quadratic BSDEs driven by continuous martingale and application to maximization problem*, 2007. Available at http://arxiv.org/abs/math/0610749.

[12] D. Nualart, *The Malliavin calculus and related topics*, Probability and its Applications (New York), Springer-Verlag, New York, 1995.MR1344217

[13] P. E. Protter, *Stochastic integration and differential equations*, Second, Applications of Mathematics (New York), vol. 21, Springer-Verlag, Berlin, 2004. Stochastic Modelling and Applied Probability.MR2020294