



# THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### QUADRATIC FBSDE WITH GENERALIZED BURGERS' TYPE NONLINEARITIES, PERTURBATIONS AND LARGE DEVIATIONS

**Citation for published version:**

Frei, C & Dos Reis, G 2013, 'QUADRATIC FBSDE WITH GENERALIZED BURGERS' TYPE NONLINEARITIES, PERTURBATIONS AND LARGE DEVIATIONS' *Stochastics and dynamics*, vol 13, no. 2, 1250015., 10.1142/S0219493712500153

**Digital Object Identifier (DOI):**

[10.1142/S0219493712500153](https://doi.org/10.1142/S0219493712500153)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Author final version (often known as postprint)

**Published In:**

*Stochastics and dynamics*

**Publisher Rights Statement:**

Electronic version of an article published as [Stochastics and Dynamics, Volume 13, Issue 02, June 2013, 37 Pages]  
[Article DOI 10.1142/S0219493712500153] © [copyright World Scientific Publishing Company]  
[<http://www.worldscientific.com/doi/abs/10.1142/S0219493712500153?queryID=31%2F574902>]

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# Quadratic FBSDE with generalized Burgers' type nonlinearities, perturbations and large deviations\*

Christoph Frei

University of Alberta  
Dept. of Math. and Stat. Sciences  
Edmonton AB T6G 2G1  
Canada

cfrei@ualberta.ca

Gonçalo Dos Reis

Technische Universität Berlin  
10623 Berlin, Germany  
and  
CMA/FCT/UNL  
2829-516 Caparica, Portugal  
dosreis@math.tu-berlin.de

15h32, 04/07/2012

## Abstract

We discuss BSDE with drivers containing nonlinearities of the type  $p(y)|z|$  and  $p(y)|z|^2$  with  $p$  a polynomial of any degree. Sufficient conditions are given for existence and uniqueness of solutions as well as comparison results. We then connect the results to the Markovian FBSDE setting, discussing applications in the theory of PDE perturbation and stating a result concerning a large deviations principle for the first component of the solution to the BSDE.

**2010 AMS subject classifications:** Primary: 60H10; Secondary: 60J60, 60F10, 35B20.

**Key words and phrases:** FBSDE, quadratic growth BSDE, generalized Burgers' PDE, PDE perturbation, large deviations principle, Navier-Stokes.

## 1 Introduction

In the past 10 years there has been an explosion of publications in the field of backward stochastic differential equations (BSDE) mainly due to their significance in optimization problems in stochastic control theory. BSDE provide

---

\*This is a preprint-postreview version. Electronic version of an article to be published as [Stochastics and Dynamics (SD), Vol. ?, Issue ?, Year?, 37 Pages] [doi:10.1142/S0219493712500153] © [copyright World Scientific Publishing Company] [http://www.worldscinet.com/sd/]

a full stochastic approach to control problems which are usually found in the Hamilton-Jacobi-Bellman formalism. Moreover, in a Markovian setting they provide a connection to certain classes of parabolic PDE via the generalized Feynman-Kac formula. An application of this falls into the realm of numerical analysis, where it may be better to employ a probabilistic solver for a PDE than a standard deterministic one (because of dimensionality for instance). BSDE with driver that grow quadratically in the control variable (introduced in [Kob00] and referred to as qgBSDE) are of particular importance as they appear naturally in many problems from mathematical finance, for example in the context of utility optimization with exponential or power utility functions in incomplete markets.

In this work we study an extended class of qgBSDE where we allow for cross nonlinearities in the driver function. The canonical setting for qgBSDE, apart from a bounded terminal condition, is a driver satisfying a growth condition of the type

$$|f(t, y, z)| \leq C(1 + |y| + |z|^2) \quad \text{with } C > 0.$$

Here we extend, for the first time, the framework to a growth condition of the type

$$|f(t, y, z)| \leq C(1 + |y| + (1 + |y|^k)|z| + (1 + |y|^k)|z|^2), \quad k \in \mathbb{N}, C > 0,$$

which coupled with the corresponding modulus of continuity condition allows us to obtain existence, uniqueness and comparison results. In a second step we associate with a BSDE of this type a standard forward diffusion process (these systems are called forward-backward SDE and denoted as FBSDE) and make the link to the PDE framework via a nonlinear Feynman-Kac formula.

The results we present in this work concerning this class of BSDE serve as a stepping stone for future research, for instance, under certain conditions this type of BSDE can be seen as cash sub-additive risk measures as described in Section 8 of [EKR09]. Moreover, they also appear in problems of optimal investment and consumption when payments of taxes take place (see [EFKY11]).

In the PDE framework this type of nonlinearities are also of importance. These improvements on existence and uniqueness results for BSDE include, as particular cases in the PDE framework, results for the Burgers' and generalized Burgers' equations. For example, [CS09, CS10] consider a problem in fluid mechanics that links the Navier-Stokes equation for incompressible fluid flow and FBSDE with a driver of the type  $f(y, z) = yz$ . But they only study that specific FBSDE.

Another example is the equation  $u_t = \Delta u + u|\nabla_x u|^2$  for  $x \in D$ ,  $t \geq 0$  and where  $D$  is a smooth bounded domain of  $\mathbb{R}^n$ . This PDE relates to an FBSDE with a driver of the type  $f(y, z) = y|z|^2$  that clearly falls in our

setting. This type of equation belongs to a class of equations used in a wide range of applications: in geometry, as a tool to build a harmonic map that is homotopy equivalent to another given map (see e.g. [Str96]), in the theory of ferromagnetic materials (the Landau-Lifschitz-Gilbert equation or in models of magnetostriction) and the theory of liquid crystals (see [BDPvH02] and references therein).

The third message of this paper is concerned with the theory of PDE perturbation and the same concept for FBSDE. A rough intuition is as follows: given a first-order transport PDE for which one is uncertain whether it has a solution or one is unable to ascertain its regularity, then one usually adds a vanishing regularizing term (typically a vanishing Laplacian) transforming the first order PDE into a second order parabolic PDE which has better properties (or so one hopes). Then as the regularizing term vanishes one aims at proving that part of those better properties carry to the limit.

We discuss the PDE equivalent to a perturbation technique in the FBSDE framework. Here we look at a FBSDE system where the forward diffusion equation is perturbed by a  $\sqrt{\varepsilon}\sigma dW$  term and investigate what happens to the FBSDE solution as  $\varepsilon$  vanishes. The scope of the results we are able to obtain depend on the type of perturbation we work with. With suitable assumptions we are able to go indeed far and even show a large deviations principle (LDP) for the first component of the BSDE's solution as the SDE's diffusion coefficient vanishes (LDPs for SDEs are well known, see e.g. [FW98], [DZ09] or [FK06]). In a subsequent iteration, we give two particular applications of the theory we have just developed to emphasize the results and their limitations. The relevance of this type of results in finance is related to the work of [SS11], where the authors look at smile asymptotics derived from a model using FBSDE as risk measures. Although the authors start with a FBSDE formulation, they quickly change to a PDE formulation in order to obtain the said asymptotics. The results we propose in this work may prove useful to show such asymptotics without the need of such PDE results.

Concerning the techniques we use to show the large deviations principle, based on the works of [DR07] and [Rai06], they differ quite substantially from the techniques commonly used to prove such a result for diffusion processes. In the FBSDE case one is able to directly use the contraction principle while in the SDE framework, when the diffusion coefficient depends on the state process, one is not. The reason is that the SDE's solution process when interpreted in the space of continuous functions, although measurable, needs not be continuous in that space. A rigorous derivation of an LDP for SDEs can be found for example in [DZ09] or [FK06]. In our case we are able to take advantage of the Markovian framework in which the FBSDE is immersed and its inherent PDE characterization. This lightens considerably the complexity of the proof and, under certain conditions, allows a direct use of the contraction principle.

The work is organized as follows: in Section 2 we introduce the nota-

tion we will work with and recall for completeness some results on bounded mean oscillation (BMO) martingales. Section 3 contains the main theorem concerning existence, uniqueness, comparison and estimates for the class of BSDE we are introducing. In Section 4 we apply the results to the Markovian FBSDE setting, link the existence of solution of the FBSDE to the viscosity solution of the corresponding PDE and discuss differentiability of the FBSDE with relation to the initial condition of the forward diffusion. We close with Section 5 and the results on the perturbation of FBSDE and the corresponding PDEs, a large deviations results and two applications.

## 2 Preliminaries

### 2.1 Spaces and Notation

Throughout fix  $T > 0$ . We work on a canonical Wiener space  $(\Omega, \mathcal{F}, \mathbb{P})$  carrying a  $d$ -dimensional Wiener process  $W = (W^1, \dots, W^d)$  restricted to the time interval  $[0, T]$ , and we denote by  $\mathcal{F} = (\mathcal{F}_t)_{t \in [0, T]}$  its natural filtration enlarged in the usual way by the  $\mathbb{P}$ -zero sets. We shall need the following operators, and auxiliary spaces of functions and stochastic processes: let  $p \geq 2, m, n, d \in \mathbb{N}$ ,  $\mathbb{Q}$  a probability measure on  $(\Omega, \mathcal{F})$ . We use the symbol  $\mathbb{E}^{\mathbb{Q}}$  for the expectation with respect to  $\mathbb{Q}$ , and omit the superscript for the canonical measure  $\mathbb{P}$ . For vectors  $x = (x^1, \dots, x^m)$  in Euclidean space  $\mathbb{R}^m$  we write  $|x| = (\sum_{i=1}^m (x^i)^2)^{\frac{1}{2}}$ , we also define  $\mathbf{1} = (1, \dots, 1)$  for the  $m$ -dimensional vector with all entries equal to 1. We define by  $\text{sgn}(x) : \mathbb{R} \rightarrow \mathbb{R}$  the sign function yielding 1 if  $x \geq 0$  and  $-1$  for  $x < 0$ . Let  $B_R(x_0)$  denote the ball of radius  $R \geq 0$  centered around the point  $x_0$ . By  $1_A$  we denote the indicator function of a set  $A$ .

For a map  $b : \mathbb{R}^m \rightarrow \mathbb{R}^d$ , we denote by  $\nabla b$  its Jacobian matrix whenever it exists. To denote the  $j$ -th first derivative of the function  $b(x)$  with  $x \in \mathbb{R}^m$  we write  $\nabla_{x_j} b$ . For a function  $h(x, y) : \mathbb{R}^m \times \mathbb{R}^d \rightarrow \mathbb{R}$  we write  $\nabla_x h$  or  $\nabla_y h$  to refer to the first derivatives with relation to  $x$  and  $y$  respectively.  $\Delta$  denotes the canonical Laplacian operator. We will use the operator  $\partial_t$  to denote the temporal partial derivative.

We also introduce the following spaces:

- $C_b^k(\mathbb{R}^m, \mathbb{R}^n)$  the set of  $k$ -times differentiable real valued maps defined on  $\mathbb{R}^m$  mapping onto  $\mathbb{R}^n$  with bounded partial derivatives up to order  $k$ , and  $C_b^\infty(\mathbb{R}^m, \mathbb{R}^n) = \cap_{k \geq 1} C_b^k(\mathbb{R}^m, \mathbb{R}^n)$ ; We omit the subscript  $b$  to denote the same set but without the boundedness assumptions.
- $L^p(\mathbb{R}^m; \mathbb{Q})$  the space of  $\mathcal{F}_T$ -measurable random variables  $X : \Omega \rightarrow \mathbb{R}^m$ , normed by  $\|X\|_{L^p} = \mathbb{E}^{\mathbb{Q}}[|X|^p]^{\frac{1}{p}}$ ;  $L^\infty$  the space of bounded random variables;

- $\mathcal{S}^p(\mathbb{R}^m)$  the space of all  $\mathcal{F}$ -adapted processes  $(Y_t)_{t \in [0, T]}$  with values in  $\mathbb{R}^m$  normed by  $\|Y\|_{\mathcal{S}^p} = \mathbb{E}[\sup_{t \in [0, T]} |Y_t|^p]^{\frac{1}{p}}$ ;  $\mathcal{S}^\infty(\mathbb{R}^m)$  the space of bounded measurable processes;
- $\mathcal{H}^p(\mathbb{R}^m, \mathbb{Q})$  the space of all  $\mathbb{R}^m$ -valued predictable processes  $(Z_t)_{t \in [0, T]}$  normed by  $\|Z\|_{\mathcal{H}^p} = \mathbb{E}^{\mathbb{Q}}[(\int_0^T |Z_s|^2 ds)^{p/2}]^{\frac{1}{p}}$ ;
- $BMO(\mathbb{Q})$  or  $BMO_2(\mathbb{Q})$  the space of square integrable martingales  $\Phi$  with  $\Phi_0 = 0$  and satisfying

$$\|\Phi\|_{BMO(\mathbb{Q})} = \sup_{\tau} \left\| \mathbb{E}^{\mathbb{Q}}[\langle \Phi \rangle_T - \langle \Phi \rangle_{\tau} | \mathcal{F}_{\tau}] \right\|_{\infty}^{1/2} < \infty,$$

where the supremum is taken over all stopping times  $\tau \in [0, T]$ .

- $\mathcal{H}_{BMO}$  the space of  $\mathbb{R}^m$ -valued  $\mathcal{H}^p$ -integrable processes  $(Z_t)_{t \in [0, T]}$  for all  $p \geq 2$  such that  $\int_0^T Z_s dW_s \in BMO$ . We define<sup>1</sup>  $\|Z\|_{\mathcal{H}_{BMO}} = \|\int Z dW\|_{BMO}$ .

If there is no ambiguity about the underlying spaces or measures, we also omit them as arguments in the function spaces defined above.

Constants appearing in inequalities of our proofs will for simplicity be denoted by  $C$ , although they may change from line to line.

## 2.2 BMO processes and their properties

The BMO space is an interesting space of stochastic processes satisfying

$$L^\infty \subsetneq BMO \subsetneq \bigcap_{p \geq 1} L^p.$$

For more details on BMO spaces we refer the reader to [Kaz94].

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

**Lemma 2.1** (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  $d\mathbb{Q} = \mathcal{E}(M)_T d\mathbb{P}$  is a probability measure.
- 2) For every BMO martingale  $M$ , there exists  $p^* > 1$  such that<sup>2</sup>  $\mathcal{E}(M) \in L^{p^*}$ . Moreover, there exists a constant  $C_{p^*}$  depending only on  $p^*$  and the BMO norm of  $M$  such that for any stopping time  $\tau \in [0, T]$  it holds that

$$\mathbb{E}[\mathcal{E}(M)_T^{p^*} | \mathcal{F}_t] \leq C_{p^*} (\mathcal{E}(M)_t)^{p^*} \quad (2.1)$$

<sup>1</sup>If  $Z * W \in BMO$  then one has automatically that  $Z \in \mathcal{H}^p$  for all  $p \geq 2$ . This is a consequence of the BMO spaces, for more see the BMO results subsection below.

<sup>2</sup>The number  $p^*$  can be found through the function  $\Psi(x) = \{1 + \frac{1}{x^2} \log \frac{2x-1}{2(x-1)}\}^{1/2} - 1$  defined for all  $1 < x < \infty$  and verifying  $\lim_{x \rightarrow 1^+} \Psi(x) = \infty$  and  $\lim_{x \rightarrow \infty} \Psi(x) = 0$ . In other words, if  $\|M\|_{BMO_2} < \Psi(p^*)$ , then  $\mathcal{E}(M) \in L^{p^*}$ , see Theorem 3.1 [Kaz94].

3) If  $\|M\|_{BMO_2} < 1$ , then<sup>3</sup> for every stopping time  $\tau \in [0, T]$

$$\mathbb{E} \left[ \exp \{ \langle M \rangle_T - \langle M \rangle_\tau \} | \mathcal{F}_\tau \right] < \frac{1}{1 - \|M\|_{BMO_2}^2}. \quad (2.2)$$

In particular, if  $Z * W \in BMO$ , then for every  $p \geq 1$  it holds that

$$\mathbb{E} \left[ \left( \int_0^T |Z_s|^2 ds \right)^p \right] \leq p! \|Z * W\|_{BMO}^{2p} \quad \Rightarrow \quad Z \in \mathcal{H}^{2p}.$$

Moreover, for any  $p \geq 1$  and any<sup>4</sup>  $\varepsilon \in (0, 2)$

$$\mathbb{E} \left[ \exp \left\{ p \int_0^T |Z_s|^\varepsilon ds \right\} \right] \leq \widehat{C} < \infty, \quad (2.3)$$

where  $\widehat{C}$  depends on  $p$ ,  $\varepsilon$  and  $\|Z * W\|_{BMO_2}^2$ .

### 3 Existence, uniqueness and comparison

We now introduce the class of BSDE we will work with. The main novelty is a setting that allows for nonlinearities of the type  $y^k z$  or  $y^k |z|^2$  for some positive power  $k \geq 1$ .

#### 3.1 Assumptions

Take a BSDE satisfying the following dynamics

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s dW_s. \quad (3.1)$$

We refer to this stochastic equation as BSDE( $\xi, f$ ). We also introduce the set of assumption under which we will be working.

**Assumption 3.1.**  $\xi$  is an  $\mathcal{F}_T$ -measurable uniformly bounded random variable, i.e. for some  $M > 0$  we have  $\|\xi\|_{L^\infty} \leq M$ ;

**Assumption 3.2.**  $f : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$  is an  $\mathcal{F}$ -predictable continuous function. There exist  $k \in \mathbb{N}$  and positive constants  $K, \delta$  such that for all  $(\omega, t, y, z) \in \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^d$

$$|f(\omega, t, y, z)| \leq K(1 + |y| + (1 + |y|^k)|z| + (1 + \gamma|y|^k)|z|^2). \quad (3.2)$$

<sup>3</sup>This result is known as the John-Nirenberg inequality, see Theorem 2.2 in [Kaz94].

<sup>4</sup>This inequality follows from (2.2), since for every  $\varepsilon \in (0, 2)$  and  $\delta > 0$  there exists a constant  $C_{\delta, \varepsilon}$  such that  $|z|^\varepsilon \leq C_{\delta, \varepsilon} + \delta|z|^2$  for all  $z$ .

**Assumption 3.3.** *There exist  $k \in \mathbb{N}$  and a positive constant  $K$  such that for all  $t \in [0, T]$ ,  $y, y' \in \mathbb{R}$  and  $z, z' \in \mathbb{R}^d$  it holds  $\mathbb{P}$ -a.s. that*

$$\begin{aligned} & |f(\cdot, t, y, z) - f(\cdot, t, y', z')| \\ & \leq K \left\{ (1 + (1 + \gamma|z| + \gamma|z'|)(|y|^{k-1} + |y'|^{k-1})(|z| + |z'|))|y - y'| \right. \\ & \quad \left. + (1 + |y|^k + |y'|^k + (1 + \gamma|y|^k + \gamma|y'|^k)(|z| + |z'|))|z - z'| \right\}. \end{aligned} \quad (3.3)$$

**Remark 3.4.** *Drivers like  $f(y, z) = \nu yz$  or  $f(y, z) = y|z|^2$  can be found in applications in physics (the first relates to Burger's PDE, see the introduction for remarks on the second). Drivers like  $f(t, y, z) = \theta_t z + \gamma|z|^2$  or  $f(y, z) = -(ay^+ - by^-)|z|^2$  with  $0 < a < b$  are found in applications in finance (the last one relates to cash subadditive risk measures, see e.g. [EKR09]).*

**Remark 3.5.** *In the above assumptions we write a domination in terms of a power  $|y|^k$  but since  $|y|^q \leq C_{p,q} + |y|^p$  for any  $q \leq p$  and some constant  $C_{p,q}$  it is clear that this includes any polynomial dependence of  $y$  up to power  $k$ .*

**Remark 3.6.** *In [Tev08], the main theorem holds under his assumption (B) that imposes " $|\nabla_y f| \leq \text{Const.}$ ", a condition which does not allow for  $yz$  type non-linearities. However, Proposition 1 of [Tev08] partially covers our setting by allowing  $y^2$  terms if an extra smallness assumption of the involved data is taken, namely that  $\|\xi\|_{L^\infty}$  and  $\|f(\cdot, 0, 0)\|_{L^\infty}$  are very small.*

For completeness we quote Theorem 2.3 from [Kob00]. This result plays a crucial role in proving that a BSDE under Assumption 3.2 has a solution.

**Proposition 3.7** (Theorem 2.3 of [Kob00]). *Let Assumption 3.1 hold and assume a continuous  $\mathcal{F}$ -predictable function  $f : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$  satisfies for any  $(\omega, t) \in \Omega \times [0, T]$*

$$|f(\omega, t, y, z)| \leq A + B|y| + F|z|^2, \quad \text{with } A, B, F \in [0, +\infty). \quad (3.4)$$

*Then the BSDE  $(\xi, f)$  (3.1) has a solution  $(Y, Z) \in \mathcal{S}^\infty \times \mathcal{H}^2$ . The process  $Y$  has continuous paths.*

*Moreover, there exists a unique minimal solution  $(Y_*, Z_*)$  (respectively a unique maximal solution  $(Y^*, Z^*)$ ) in the sense that the solution  $(U, V)$  of BSDE  $(\eta, h)$  where  $f \leq h$  and  $\xi \leq \eta$  (respectively  $f \geq h$  and  $\xi \geq \eta$ ) satisfies  $Y_* \leq U$  (respectively  $Y^* \geq U$ ).*

As is expected from the theory of quadratic BSDE (see [Mor09], [IDR10] or [DR11]), we can conclude that<sup>5</sup>  $Z * W \in \text{BMO}$ .

**Lemma 3.8.** *Under the conditions of Proposition 3.7, we have  $Z \in \mathcal{H}_{\text{BMO}}$ .*

---

<sup>5</sup>Here we cannot refer to Lemma 3.1 of [Mor09] with her Assumption H1 instead of just (3.4). Her extra restriction  $\gamma \geq b$  restricts the methods we use in this work. Please compare her assumption H1 with our Assumption 3.2.



*Proof.* We only sketch the proof since this kind of argument is known. Take  $t \in [0, T]$ . Assume that  $\|Y\|_{\mathcal{S}^\infty}$  and  $\|\xi\|_{L^\infty}$  are bounded by the same constant  $M$  and take a constant  $\alpha$  satisfying  $\alpha > 2F$ . Let  $\tau \in [0, T]$  be a stopping time. Applying Itô's formula to the process  $\widehat{Y}_t = \exp\{\alpha Y_t\}$  between  $[\tau, T]$ , (3.4) and the conditional expectation on  $\mathcal{F}_\tau$ , we obtain:

$$\begin{aligned} \widehat{Y}_\tau &\leq \mathbb{E}\left[e^{\alpha\xi} + \int_\tau^T \widehat{Y}_s[\alpha A + \alpha B|Y_s| + (\alpha F - \frac{\alpha^2}{2})|Z_s|^2] ds \mid \mathcal{F}_\tau\right] \\ &\Leftrightarrow (\frac{\alpha^2}{2} - \alpha F)e^{-\alpha M}\mathbb{E}\left[\int_\tau^T |Z_s|^2 ds \mid \mathcal{F}_\tau\right] \leq e^{\alpha M}(1 + (\alpha A + \alpha BM)T). \end{aligned}$$

Since  $\alpha > 2F$  we have that  $2\alpha(\alpha - 2F) > 0$  and hence we easily get from the definition of the BMO-norm that  $Z * W \in BMO$ . Moreover, from the calculations we just did, the BMO norm of  $Z * W$  is bounded from above by a universal constant depending only on  $\|\xi\|_{L^\infty}$ ,  $\|Y\|_{\mathcal{S}^\infty}$  and the constants  $A$ ,  $B$  and  $F$ .

We remark as well that as  $F$  decreases the upper bound for the  $\mathcal{H}_{BMO}$ -norm of  $Z$  also decreases and vice-versa.  $\square$

### 3.2 The main results - Abstract BSDE setting

We now state the main results of this section. We start with an existence and uniqueness result.

**Theorem 3.9** (Existence). *Let Assumptions 3.1 and 3.2 hold. Then the BSDE (3.1) has a solution  $(Y, Z)$  in  $\mathcal{S}^\infty \times \mathcal{H}_{BMO}$ . There exists an upper bound for  $\|Y\|_{\mathcal{S}^\infty}$  independent of  $\gamma$ .*

*Moreover, there exists a unique maximal solution  $(Y, Z)$  in the sense that  $\widehat{Y} \leq Y$  for any other possible solution  $(\widehat{Y}, \widehat{Z})$ .*

**Theorem 3.10** (Uniqueness). *Let Assumptions 3.1, 3.2 and 3.3 hold and take  $(Y, Z)$  to be the solution of BSDE (3.1). Let  $\|Y\|_{\mathcal{S}^\infty} \leq R$ ,  $\|Z\|_{\mathcal{H}_{BMO}} \leq R_\gamma^Z$  for  $R, R_\gamma^Z \geq 0$ , where  $R$  is independent of  $\gamma$  and  $R_\gamma^Z$  is not. Define the number<sup>6</sup>  $p^* = \Psi^{-1}(2K(1 + 2R^k)\sqrt{T} + 4K(1 + 2\gamma R^k)R_\gamma^Z)$  with  $q^*$  its Hölder conjugate. Assume further that  $\gamma$  satisfies  $(2q^*K8\gamma R^{k-1})^{\frac{1}{2}}R_\gamma^Z < 1$ .*

*Then BSDE (3.1) has a unique solution in  $\mathcal{S}^\infty \times \mathcal{H}_{BMO}$ .*

**Remark 3.11.** 1) A small clarification is necessary concerning the existence of  $\gamma$  in Theorem 3.10, since  $q^*$  and  $R_\gamma^Z$  both depend on  $\gamma$  and there is the possibility of a circle argument. One argues as follows, as  $\gamma$  decreases so does the upper bound for  $\|Z\|_{\mathcal{H}_{BMO}}$  (this is intuitive but see e.g. Lemma 3.8). From the definition of  $\Psi$  in Lemma 2.1, the smaller the upper bound for  $\|Z\|_{\mathcal{H}_{BMO}}$  is, the greater  $p^*$  is and hence the smaller  $q^*$  is. So, as  $\gamma$

<sup>6</sup>The function  $\Psi$  is defined in part 2) of Lemma 2.1.

decreases so do  $q^*$  and  $R_\gamma^Z$ , making it even easier for condition (3.13) to be satisfied.

2) The proof of Theorem 3.10 hints that it is possible to obtain the same result but under an assumption of the type  $f(Y, Z) - f(U, Z) \leq -\alpha|Z|^2(1 + \phi(Y, U))(Y - U)$  with  $\phi$  a continuous positive function. This condition would then replace the smallness assumption of  $\gamma$ . We do not explore this direction.

Finally, we state a comparison result:

**Theorem 3.12** (Comparison). *Let  $i \in \{1, 2\}$ . Define  $(Y^i, Z^i) \in \mathcal{S}^\infty \times \mathcal{H}_{BMO}$  as the solution of BSDE (3.1) with terminal condition  $\xi^i$  and driver  $f_i$ . Assume that  $\xi^i$  and  $f_i$  satisfy Assumptions 3.1, 3.2 and 3.3. Assume  $\gamma_i$  satisfies the condition of Theorem 3.10 and that<sup>7</sup>*

$$\xi^1 \leq \xi^2, \quad \text{and} \quad f_1(t, Y_t^2, Z_t^2) \leq f_2(t, Y_t^2, Z_t^2) \quad dt \otimes d\mathbb{P}\text{-a.s.},$$

Then we have for all  $t \in [0, T]$  that  $Y_t^1 \leq Y_t^2$   $\mathbb{P}$ -a.s.

Moreover, if either  $\xi^1 < \xi^2$  or  $f_1(t, Y_t^2, Z_t^2) < f_2(t, Y_t^2, Z_t^2)$  in a set of positive  $dt \otimes d\mathbb{P}$ -measure then  $Y_0^1 < Y_0^2$ .

**Corollary 3.13** (A priori estimate). *Under the assumptions of Theorem 3.12 with  $\gamma = 0$ , it holds for  $\delta f := f_1(\cdot, Y^2, Z^2) - f_2(\cdot, Y^2, Z^2)$  that for  $p \geq 2$*

$$\begin{aligned} \|Y^1 - Y^2\|_{\mathcal{S}^p}^p &\leq C_{r'}^p C_p C_{pr', q'} \mathbb{E} \left[ \left( |\xi^1 - \xi^2| + \int_0^T |\delta f_s| ds \right)^{prq} \right]^{\frac{1}{qr}}, \\ \|Z^1 - Z^2\|_{\mathcal{H}^p}^p &\leq C \left\{ \|\delta Y\|_{\mathcal{S}^p}^p + \|\delta Y\|_{\mathcal{S}^{2p}}^p + \|\delta Y\|_{\mathcal{S}^p}^{\frac{p}{2}} \mathbb{E} \left[ \left( \int_0^T |\delta f_s| ds \right)^{p1} \right]^{\frac{1}{2}} \right\}, \end{aligned} \quad (3.5)$$

where the constants  $C_{pr', q'}$ ,  $C_p$ ,  $C_{r'}^p$  and  $C$  depend only on the constants appearing in Assumptions 3.1, 3.2 and 3.3 and the upper bounds of the BMO norms of  $Z^1 * W$  and  $Z^2 * W$ . Furthermore, the numbers  $r, q > 1$  are related as well to the BMO norms of  $Z^1 * W$  and  $Z^2 * W$  via point 3 of Lemma 2.1.

The following subsections contain the proofs of the above theorems.

### 3.2.1 Proof of the Theorem 3.9 - Existence

As we mentioned earlier, we now need to state and prove a sequence of results on a certain family of BSDEs that approximates the BSDE (3.1).

---

<sup>7</sup>The comparison theorem still holds if instead of  $f_1(t, Y_t^2, Z_t^2) \leq f_2(t, Y_t^2, Z_t^2)$  one assumes  $f_1(t, Y_t^1, Z_t^1) \leq f_2(t, Y_t^1, Z_t^1)$ . The proof is a straightforward modification of the proof we are giving.

### A “truncation of the identity” function family

We start by defining a family of smooth functions that truncate the identity function, namely

**Definition 3.14.** *For each  $n \in \mathbb{N}$  let  $h_n : \mathbb{R} \rightarrow \mathbb{R}$  be a continuously differentiable function with the following properties:*

- $|h_n(x)| \leq n$ ,  $|h_n(x)| \leq |x|$  and  $|\nabla_x h_n(x)| \leq 1$  for all  $x \in \mathbb{R}$ ;
- $h_n(x) = x$  for any  $x \in [-(n-1), (n-1)]$  and  $h_n(x) = n$  for any  $x$  outside  $(-n, n)$ ;

The sequence  $(h_n)_{n \in \mathbb{N}}$  converges locally uniformly to the identity function; the sequence  $(\nabla_x h_n)_{n \in \mathbb{N}}$  converges to 1 locally uniformly.

We call the family of functions  $(h_n)_{n \in \mathbb{N}}$  a (differentiable) truncation of the identity<sup>8</sup>.

### A family of truncated drivers and their properties

With the above defined family of functions we take a driver  $f$  satisfying Assumptions 3.2 and 3.3 and define the sequence  $(f_n)_{n \in \mathbb{N}}$  through

$$f_n(t, y, z) = f(t, h_n(y), z), \quad \text{for } (\omega, t, y, z) \in \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^d, n \in \mathbb{N}.$$

Using the properties of  $h_n$  (along with  $|z| \leq 1 + |z|^2$ ) and in view of (3.2), it is clear that there exists a sequence  $(C_n)_{n \in \mathbb{N}}$  such that for any  $(y, z) \in \mathbb{R} \times \mathbb{R}^d$ ,  $n \in \mathbb{N}$

$$|f_n(\cdot, y, z)| \leq K(1 + C_n(1 + |z|^2)) \quad \text{with } \lim_{n \rightarrow \infty} C_n = \infty, \quad (3.6)$$

$$|f_n(\cdot, y, z)| \leq K(1 + |y| + (1 + |y|^k)|z| + (1 + \gamma|y|^k)|z|^2), \quad (3.7)$$

where the latter inequality corresponds to (3.2) for  $f_n$ .

It is also easy to see that for any  $(y, z), (y', z') \in \mathbb{R} \times \mathbb{R}^d$  and  $n \in \mathbb{N}$  the difference  $|f_n(\cdot, y, z) - f_n(\cdot, y', z')|$  satisfies condition (3.3) with constants independent of  $n$ : just use that  $|h_n(x)| \leq |x|$  along with the mean value theorem on  $h_n$  and the fact that  $|\nabla h_n| \leq 1$ .

### A family of truncated BSDE and results on them

With the family of drivers  $(f_n)_{n \in \mathbb{N}}$  we define a family of approximating BSDE obtained by replacing  $f$  in (3.1) by  $f_n$  so that

$$Y_t^n = \xi + \int_t^T f_n(s, Y_s^n, Z_s^n) ds - \int_t^T Z_s^n dW_s. \quad (3.8)$$

---

<sup>8</sup>Such a family of functions clearly exists, for an explicit example we point to Subsection 3.2.1 of [DR11].

The next result states the existence and uniqueness of the solution to (3.8). The argumentation here follows two steps: in the first we show existence of a solution to (3.8) where we obtain upper bounds for the solution that depend on  $n \in \mathbb{N}$ . In a second step, we twist the arguments and obtain upper bounds for the norms of  $\mathcal{S}^\infty$  and  $\mathcal{H}_{BMO}$  independent of the truncation height  $n \in \mathbb{N}$ .

**Lemma 3.15.** *Let Assumptions 3.1 and 3.2 hold. For each  $n \in \mathbb{N}$ , BSDE (3.8) has a (unique minimal/maximal) solution  $(Y^n, Z^n)$  in  $\mathcal{S}^\infty \times \mathcal{H}_{BMO}$ .*

*Proof.* This follows immediately from Proposition 3.7 and Lemma 3.8 applied to (3.8), using that  $f_n$  satisfies (3.6).  $\square$

The next lemma shows that it is possible to estimate the norms of  $(Y^n, Z^n)$  solution to (3.8) in  $\mathcal{S}^\infty \times \mathcal{H}_{BMO}$  by universal constants independent of  $n \in \mathbb{N}$  imposing *only* Assumptions 3.1 and 3.2. We use (3.7) instead of (3.6).

**Lemma 3.16.** *Let Assumptions 3.1 and 3.2 hold. Let  $n \in \mathbb{N}$  and take  $(Y^n, Z^n)$  to be the solution of BSDE (3.8) as given in the previous lemma, then*

$$\sup_{n \in \mathbb{N}} \{ \|Y^n\|_{\mathcal{S}^\infty} + \|Z^n\|_{\mathcal{H}_{BMO}} \} < \infty.$$

*In particular  $\sup_n \|Y^n\|_{\mathcal{S}^\infty} \leq e^{KT}(M + KT)$ , which is independent of  $\gamma$ .*

*Proof.* Let  $t \in [0, T]$  and  $n \in \mathbb{N}$ .

**Step 1 - preparation:** We start by going back to BSDE (3.8). Let  $(Y^n, Z^n)$  solve BSDE (3.8), then we can decompose the growth of driver  $f_n(t, Y_t^n, Z_t^n)$  by using the growth assumption (3.7) in the following way:

$$|f_n(t, Y_t^n, Z_t^n)| \leq K + \beta_t^n Y_t^n + b_t^n \cdot Z_t^n,$$

where  $\beta^n$  and  $b^n$  are two processes valued in  $\mathbb{R}$  and  $\mathbb{R}^d$  respectively defined by: (the number  $k$  follows from Assumption 3.2)

$$\beta_t^n := K \operatorname{sgn}(Y_t^n), \quad \text{and} \quad b_t^n := K \frac{1 + |Y_t^n|^k + (1 + \gamma|Y_t^n|^k)|Z_t^n|}{|Z_t^n|} Z_t^n \mathbf{1}_{\{|Z_t^n| \neq 0\}}.$$

It is clear that  $\sup_{n \in \mathbb{N}, t \in [0, T]} \|\beta^n\|_{\mathcal{S}^\infty} < \infty$  and  $(\beta^n Y^n) = K|Y^n|$ . Moreover,  $b^n \cdot Z^n = (1 + |Y^n|^k + (1 + \gamma|Y^n|^k)|Z^n|)|Z^n|$  and since  $Y^n \in \mathcal{S}^\infty$ ,  $Z^n \in \mathcal{H}_{BMO}$  and (3.2) we have

$$|b^n| \leq K \{1 + |Y^n|^k + (1 + \gamma|Y^n|^k)|Z^n|\} \quad \Rightarrow \quad b^n * W \in BMO.$$

Hence the probability measure  $\mathbb{Q}^n$  with Radon-Nikodym density  $d\mathbb{Q}^n/d\mathbb{P} = \mathcal{E}(b^n * W)_T$  is well defined and with relation to which  $W^{\mathbb{Q}^n} = W - \int_0^\cdot b_s^n ds$  is a Brownian motion.

**Step 2 - uniform  $\mathcal{S}^\infty$  bound for  $Y^n$ :** We set  $e_t^n := \exp\{\int_0^t \beta_s^n ds\}$ , which satisfies  $e^{-KT} \leq e^n \leq e^{KT}$ , and use a change of measure to show that an upper bound for  $\|Y^n\|_{\mathcal{S}^\infty}$  can be obtained independently of  $n$  (and  $\gamma$ ).

We take  $(Y^n, Z^n)$  as the solution of BSDE (3.8) and with the help of processes  $\beta^n$ ,  $b^n$  and  $e^n$  defined above we write the BSDE for  $e_t^n Y_t^n$  under  $\mathbb{Q}^n$  via Itô's formula. We have then

$$\begin{aligned} e_t^n Y_t^n &\leq e_T^n \xi - \int_t^T e_s^n Z_s^n dW_s + \int_t^T e_s^n [-\beta_s^n Y_s^n + |f_n(s, Y_s^n, Z_s^n)|] ds \\ &\leq e_T^n \xi - \int_t^T e_s^n Z_s^n dW_s + \int_t^T e_s^n [-\beta_s^n Y_s^n + K + \beta_s^n Y_s^n + b_s^n \cdot Z_s^n] ds \\ &\leq e^{KT} M - \int_t^T e_s^n Z_s^n dW_s^{\mathbb{Q}^n} + e^{KT} K(T-t). \end{aligned} \quad (3.9)$$

where we used the properties of the process  $e^n$ , the measure change to  $\mathbb{Q}^n$  and that  $\xi \in L^\infty$ . We obtain by taking  $\mathbb{Q}^n$ -conditional expectation that for some positive constant  $C$  and for all  $t \in [0, T]$

$$Y_t^n \leq e^{KT}(M + KT) \Rightarrow \sup_{n \in \mathbb{N}} \|Y^n\|_{\mathcal{S}^\infty} \leq e^{KT}(M + KT).$$

where the last conclusion follows from using similar arguments to obtain a lower bound for  $Y^n$ .

**Step 3 - uniform  $\mathcal{H}_{BMO}$  bound for  $Z^n$ :** The argument uses that we already know that  $\sup_{n \in \mathbb{N}} \|Y^n\|_{\mathcal{S}^\infty} < \infty$ . This yields the existence of  $n^* \in \mathbb{N}$  such that for any  $n \in \mathbb{N}$  with  $n \geq n^*$

$$f_n(t, Y^n, Z^n) = f_{n^*}(t, Y^n, Z^n) = f(t, Y^n, Z^n). \quad (3.10)$$

The arguments used to prove Lemma 3.8 imply that the upper bound for the BMO norm of  $Z * W$  depends only on the problem data and an upper bound for  $\|Y\|_{\mathcal{S}^\infty}$ . Hence we get here that  $\sup_{n \in \mathbb{N}} \|Z^n\|_{\mathcal{H}_{BMO}} < \infty$ .  $\square$

### The proof of the existence theorem

We are now ready to prove the existence theorem. Lemma 3.16 yields a sequence  $(Y^n, Z^n)_{n \in \mathbb{N}}$  of processes belonging to  $\mathcal{S}^\infty \times \mathcal{H}_{BMO}$  that solve BSDE (3.8) for each  $n \in \mathbb{N}$ . Moreover, the processes' respective norms are bounded uniformly in  $n$ . We need only to conclude that a limit for the sequence  $(Y^n, Z^n)_{n \in \mathbb{N}}$  exists and that the said limit solves BSDE (3.1) under Assumption 3.2.

*Proof of Theorem 3.9.* In view of (3.10) it is clear that for each  $n > n^*$  pairs  $(Y^n, Z^n)$  solve both the BSDE (3.8) and (3.1). And hence we obtain the “existence of solution” result.

Moreover, since for each  $n > n^*$  there exists a maximal solution to (3.8), we have the existence of a (infinite amount of) maximal solutions to (3.1).

It is clear that the maximal solution is unique, in the sense given in the theorem's statement. If  $(Y, Z)$  and  $(\widehat{Y}, \widehat{Z})$  are two maximal solutions then we have that  $\widehat{Y} \leq Y$  and also  $Y \leq \widehat{Y}$ . This implies also the uniqueness of a maximal solution of BSDE (3.1).  $\square$

### 3.2.2 Proof of Theorem 3.10 - Uniqueness

We now present the proof of the uniqueness result. As usual to prove existence, one needs only bounds on the growth of the involved functions but to prove uniqueness one needs to control the modulus of continuity of the involved functions, hence for uniqueness one additionally needs Assumption 3.3.

Unfortunately, the arguments we present here do not allow for general choices of  $\gamma$  (from Assumption 3.2 or 3.3). Here  $\gamma$  has to be small enough. This smallness enters in play in (3.14) below. In the later Subsection 3.3 we discuss a particular situation where it is possible to solve the BSDE for any value of  $\gamma$ .

*Proof of Theorem 3.10.* Let  $t \in [0, T]$  and assume that  $(Y, Z)$  and  $(U, V)$ , belonging to the space  $\mathcal{S}^\infty \times \mathcal{H}_{BMO}$ , are two solutions to the BSDE (3.1). The  $\mathcal{S}^\infty$ -norms of  $Y, U$  are universally bounded by a constant, say  $R$ , independently of  $\gamma$  (see Lemma 3.16); the  $\mathcal{H}_{BMO}$ -norm of  $Z, U$  are bounded by a constant  $R_\gamma^Z$  that depends on  $\gamma$  and decreases as  $\gamma$  decreases (see the proof of Lemma 3.8).

Now, define  $\delta Y = Y - U$ ,  $\delta Z = Z - V$  and the processes

$$\Gamma_t := \frac{f(t, Y_t, Z_t) - f(t, U_t, Z_t)}{Y_t - U_t} 1_{\{Y_t - U_t \neq 0\}}, \quad \hat{e}_t := \exp \left\{ \int_0^t \Gamma_s ds \right\}, \quad (3.11)$$

$$\text{and } \hat{b}_t = \frac{f(t, U_t, Z_t) - f(t, U_t, V_t)}{|Z_t - V_t|^2} (Z_t - V_t) 1_{\{|Z_t - V_t| \neq 0\}}. \quad (3.12)$$

Concerning the above defined processes, we remark that for  $\hat{b}$  we have from (3.3) (and the triangular inequality of the  $\mathcal{H}_{BMO}$ -norm) that

$$\begin{aligned} |\hat{b}| &\leq K \{1 + 2R^k + (1 + 2\gamma R^k)(|Z| + |V|)\}, \\ \|\hat{b}\|_{\mathcal{H}_{BMO}} &\leq \widehat{C} := 2K \{(1 + 2R^k)\sqrt{T} + 2(1 + 2\gamma R^k)R_\gamma^Z\}. \end{aligned}$$

Since  $Z, V \in \mathcal{H}_{BMO}$ , it is easy to see that the process  $\hat{b} * W \in BMO$  and hence the probability measure  $\widehat{\mathbb{Q}}$  with Radon-Nykodim density  $d\widehat{\mathbb{Q}}/d\mathbb{P} = \mathcal{E}(\hat{b} * W)$  is well defined and  $W^{\widehat{\mathbb{Q}}} = W - \int_0^\cdot \hat{b}_s ds$  is a  $\widehat{\mathbb{Q}}$ -Brownian motion.

Moreover, from Lemma 2.1 and defining  $p^* = \Psi^{-1}(\widehat{C})$  with  $q^*$  its Hölder conjugate we have that  $\mathcal{E}(\hat{b} * W) \in L^{p^*}$ .

In view of (3.3) we have that  $|\Gamma| \leq K(1 + 4R^{k-1}|Z| + 8\gamma R^{k-1}|Z|^2)$ . Since  $Z \in \mathcal{H}_{BMO}$ , the properties of BMO martingales, namely (2.3), combined

with the behavior of  $R_\gamma^Z$  (described in the proof of Lemma 3.8) imply that it is possible to find a  $\gamma$  small enough such that<sup>9</sup>:

$$(2q^*K8\gamma R^{k-1})^{\frac{1}{2}} < \frac{1}{R_\gamma^Z} \leq \frac{1}{\|Z * W\|_{BMO}}. \quad (3.13)$$

In view of (2.2) and the strictness of (3.13) it is always possible to find a small enough  $\varepsilon > 0$  such that

$$\hat{e}_T^{2q^*} \in L^{1+\varepsilon}. \quad (3.14)$$

It is also clear that  $\hat{e}$  has continuous paths.

**Step 1 - Uniqueness of the solution's first component:** With the help of the processes  $\hat{b}$  and  $\hat{e}$  defined above we proceed as in step 2 of the proof of Theorem 3.9 (we skip some details and point the reader to (3.9)). Using Itô's formula, we write a BSDE for  $(\hat{e}_t \delta Y_t)$  under the  $\hat{\mathbb{Q}}$ -measure:

$$\begin{aligned} \hat{e}_t \delta Y_t + \int_t^T \hat{e}_s \delta Z_s dW_s^{\hat{\mathbb{Q}}} \\ = 0 + \int_t^T \hat{e}_s [-\Gamma_s \delta Y_s + f(s, Y_s, Z_s) - f(s, U_s, Z_s)] ds = 0 \end{aligned} \quad (3.15)$$

We obtain  $\hat{e}_t \delta Y_t = 0$   $\hat{\mathbb{Q}}$ - and  $\mathbb{P}$ -a.s. for all  $t \in [0, T]$  once we justify that the stochastic integral  $\int_0^T \hat{e}_s \delta Z_s dW_s^{\hat{\mathbb{Q}}}$  is indeed a true  $\hat{\mathbb{Q}}$ -martingale. For this we have to essentially prove the square integrability of the process, i.e. the finiteness of

$$\begin{aligned} \mathbb{E}^{\hat{\mathbb{Q}}} \left[ \int_0^T |\hat{e}_s|^2 |\delta Z_s|^2 ds \right] &\leq \mathbb{E} \left[ \mathcal{E}(\hat{b} * W) |\hat{e}_T|^2 \int_0^T |\delta Z_s|^2 ds \right] \\ &\leq \mathbb{E} \left[ \left( \mathcal{E}(\hat{b} * W) \right)^{p^*} \right]^{\frac{1}{p^*}} \mathbb{E} \left[ |\hat{e}_T|^{2q^*} \left( \int_0^T |\delta Z_s|^2 ds \right)^{q^*} \right]^{\frac{1}{q^*}}. \end{aligned}$$

Combining now with the fact that  $\mathcal{E}(\hat{b} * W) \in L^{p^*}$ , that  $\delta Z_s \in \mathcal{H}_{BMO}$  and hence  $\delta Z \in \mathcal{H}^p$  for all  $p \geq 2$  and finally (3.14), we can obtain the sought conclusion.

At this point it is clear that  $\hat{e}_t \delta Y_t = 0$  for every  $t \in [0, T]$   $\mathbb{P}$ -a.s. We now want to conclude that we also have  $\delta Y = 0$ . The continuity of  $\hat{e}$  (and  $\delta Y$ ) yield that there exists a set  $A \in \Omega$  satisfying  $\mathbb{P}[A] = 0$  such that

$$\text{for all } t \in [0, T], \omega \in \Omega \setminus A \text{ we have } \hat{e}_t(\omega) > 0.$$

Now, given the positivity of  $(\hat{e}_t)_{t \in [0, T]}$  we can conclude that  $Y_t = U_t$  for any  $t \in [0, T]$   $\mathbb{P}$ -a.s.

<sup>9</sup>In short  $|\Gamma| \leq \gamma |Z|^2$  and we are faced with the question of the integrability of  $\hat{e}$ . From the John-Nirenberg inequality in (2.2) we are only able to ensure the integrability of  $\hat{e}$  when the constant  $\gamma$  is sufficiently small. In general the quadratic variation of a BMO process  $Z * W$  is not exponentially integrable.

**Step 2 - Uniqueness of the solution's second component:** We are missing only the uniqueness proof for the control component of the BSDE. Here we return to (3.15) and take advantage of the fact that we already know that  $\delta Y = 0$ . We have, by using Itô's Isometry, that  $\mathbb{E}^{\widehat{\mathbb{Q}}}\left[\int_0^T |\widehat{e}_s|^2 |\delta Z_s|^2 ds\right] = 0$ , from here using the positivity of  $(\widehat{e}_t)_{t \in [0, T]}$  (see the above step), we conclude that  $|\delta Z_t| = |Z_t - V_t| = 0 \, dt \otimes \mathbb{P}$ -a.s. This concludes the proof.  $\square$

### 3.2.3 Proof of Theorem 3.12 and Corollary 3.13 - Comparison

We now prove Theorem 3.12 and Corollary 3.13. The first parts of these proofs are very similar to that of Theorem 3.10 and so we do not give the full details.

*Proof of Theorem 3.12.* Let  $t \in [0, T]$ . Define  $\delta Y = Y^1 - Y^2$ ,  $\delta Z = Z^1 - Z^2$ ,  $\delta \xi = \xi^1 - \xi^2$  and  $\delta f = f_1(\cdot, Y^2, Z^2) - f_2(\cdot, Y^2, Z^2)$ .

We define as well, processes similar to those of (3.11) and (3.12), namely:

$$\tilde{\Gamma}_t := \frac{f_1(t, Y_t^1, Z_t^1) - f_1(t, Y_t^2, Z_t^1)}{Y_t^1 - Y_t^2} 1_{\{Y_t^1 - Y_t^2 \neq 0\}}, \quad \tilde{e}_t := \exp \left\{ \int_0^t \tilde{\Gamma}_s ds \right\}, \quad (3.16)$$

$$\text{and } \tilde{b}_t = \frac{f_1(t, Y_t^2, Z_t^1) - f_1(t, Y_t^2, Z_t^2)}{|Z_t^1 - Z_t^2|^2} (Z_t^1 - Z_t^2) 1_{\{|Z_t^1 - Z_t^2| \neq 0\}}. \quad (3.17)$$

And as before, with  $\tilde{b} * W$  we define a new probability measure  $\tilde{\mathbb{Q}}$  with Radon-Nikodym density  $d\tilde{\mathbb{Q}}/d\mathbb{P} = \mathcal{E}(\tilde{b} * W)$  and  $W \cdot^{\tilde{\mathbb{Q}}} = W - \int_0^\cdot \tilde{b}_s ds$  is a  $\tilde{\mathbb{Q}}$ -Brownian motion.

From the BSDE for  $\delta Y$ , and Itô's formula we can write a BSDE for the process  $\tilde{e}_t \delta Y_t$  as

$$\begin{aligned} \tilde{e}_t \delta Y_t &= \tilde{e}_T \delta \xi - \int_t^T \tilde{e}_s \delta Z_s dW_s^{\tilde{\mathbb{Q}}} + \int_t^T \tilde{e}_s \delta f_s ds \\ &= \mathbb{E}^{\tilde{\mathbb{Q}}} \left[ \tilde{e}_T \delta \xi + \int_t^T \tilde{e}_s \delta f_s ds \mid \mathcal{F}_t \right]. \end{aligned} \quad (3.18)$$

Using that  $\tilde{e}_t > 0$  and the theorem's assumptions,  $\delta \xi \leq 0$  and  $\delta f_t \leq 0$ , we conclude that  $\tilde{e}_t \delta Y_t \leq 0$  and hence that for all  $t \in [0, T]$  it holds that  $\delta Y_t = Y_t^1 - Y_t^2 \leq 0$   $\widehat{\mathbb{Q}}$ -a.s. and hence also  $\mathbb{P}$ -a.s.

We see also that at  $t = 0$  if  $\delta \xi < 0$  or if  $\delta f < 0$  in a set of positive  $dt \otimes d\mathbb{P}$ -measure, then the inequality is strict and we conclude that  $\delta Y_0 = Y_0^1 - Y_0^2 < 0$ , which proves the result.  $\square$

We finish this subsection with the proof of Corollary 3.13. This proof is close to that of Lemma 3.2 in [IDR10] but with a different argumentation in what the linearization trick is concerned.



*Proof of Corollary 3.13.*

**Step 1 - The estimate for  $Y^1 - Y^2$ :** To prove this corollary we stop at the point of inequality (3.18) in the proof of Theorem 3.12 and continue it from there but in a different fashion. Recall that  $\tilde{e}$  and  $\tilde{\Gamma}$  were defined in (3.16) and (3.17), and from the corollary's statement, it is assumed that  $\gamma = 0$ , this means that (3.14) now holds without any restriction on  $p$ .

We remark as well that for any  $0 \leq t \leq s \leq T$

$$\tilde{e}_s \tilde{e}_t^{-1} = \exp \left\{ \int_t^s \tilde{\Gamma}_r dr \right\} \leq A_T := \exp \left\{ \int_0^T K(1 + 4R^{k-1}|Z_s^1|) ds \right\}.$$

Since  $Z^1 \in \mathcal{H}_{BMO}$ , the random variable  $A_T$  is integrable in view of (2.3). To shorten the notation, we define  $\mathcal{E}_T := \mathcal{E}(\tilde{b} * W)_T$  as the density of the probability measure  $\tilde{\mathbb{Q}}$ . We have then

$$\begin{aligned} \delta Y_t &\leq \mathbb{E}^{\tilde{\mathbb{Q}}} \left[ \tilde{e}_T (\tilde{e}_t)^{-1} |\delta \xi| + (\tilde{e}_t)^{-1} \int_t^T \tilde{e}_s |\delta f_s| ds \mid \mathcal{F}_t \right] \\ &\leq \mathbb{E}^{\tilde{\mathbb{Q}}} \left[ A_T \left( |\delta \xi| + \int_0^T |\delta f_s| ds \right) \mid \mathcal{F}_t \right]. \end{aligned}$$

From the above inequality, it follows that

$$\begin{aligned} \delta Y_t &\leq \mathcal{E}_t^{-1} \mathbb{E}^{\mathbb{P}} \left[ \mathcal{E}_T A_T \left( |\delta \xi| + \int_0^T |\delta f_s| ds \right) \mid \mathcal{F}_t \right] \\ &\leq \mathcal{E}_t^{-1} \mathbb{E}^{\mathbb{P}} \left[ \mathcal{E}_T^{r'} \right]^{\frac{1}{r'}} \mathbb{E}^{\mathbb{P}} \left[ A_T^r \left( |\delta \xi| + \int_0^T |\delta f_s| ds \right)^r \mid \mathcal{F}_t \right]^{\frac{1}{r}} \\ &\leq C_{r'} \mathbb{E}^{\mathbb{P}} \left[ A_T^r \left( |\delta \xi| + \int_0^T |\delta f_s| ds \right)^r \mid \mathcal{F}_t \right]^{\frac{1}{r}}, \end{aligned}$$

where the last lines follow from  $\tilde{b} * W \in BMO$  combined with (2.1) and Hölder's inequality (with  $1/r + 1/r' = 1$ ) for  $r' > 1$  with  $\mathcal{E}_T \in L^{r'}(\mathbb{P})$ .

Let now  $p \geq 1$  and apply Doob's inequality (and concave Jensen's inequality) to conclude that

$$\begin{aligned} \mathbb{E} \left[ \sup_{t \in [0, T]} |\delta Y_t|^p \right] &\leq C_{r'}^p C_p \mathbb{E} \left[ A_T^{pr} \left( |\delta \xi| + \int_0^T |\delta f_s| ds \right)^{pr} \right]^{\frac{1}{r}} \\ &\leq C_{r'}^p C_p C_{pr', q'} \mathbb{E} \left[ \left( |\delta \xi| + \int_0^T |\delta f_s| ds \right)^{prq} \right]^{\frac{1}{qr}}, \end{aligned}$$

where we applied once again Hölder's inequality with exponents  $1/q + 1/q' = 1$  and version (2.3) of (2.2) for  $q' > 1$  with  $A_T^{pr} \in L^{q'}(\mathbb{P})$ .

**Step 2 - The estimate for  $Z^1 - Z^2$ :** We follow the notation introduced above. The estimate here is obtained using standard BSDE techniques, so we shorten the proof a little. We first remark that

$$|f_1(\cdot, Y^1, Z^1) - f_1(\cdot, Y^2, Z^2)| \leq L^y |\delta Y| + L^z |\delta Z|,$$

where, using (3.3) with  $\gamma = 0$  and that  $\|Y^i\|_{\mathcal{S}}^\infty < R$  for  $i \in \{1, 2\}$ ,

$$L^y := K(1 + 2R^{k-1}(|Z^1| + |Z^2|)) \quad \text{and} \quad L^z := K(1 + 2R^k + |Z^1| + |Z^2|).$$

Applying Itô's formula applied to  $|\delta Y_t|^2$  over the interval  $[0, T]$ , using the "Lipschitz" assumptions of  $f$ , Young's inequality (with parameter 2) and observe that with  $\delta f$  defined as  $\delta f = f_1(\cdot, Y^2, Z^2) - f_2(\cdot, Y^2, Z^2)$ , we obtain

$$\begin{aligned} |\delta Y_0|^2 &+ \int_0^T 2\delta Y_s \delta Z_s dW_s \\ &= |\delta \xi|^2 + \int_0^T [2\delta Y_s (f_1(s, Y_s^1, Z_s^1) - f_2(s, Y_s^2, Z_s^2)) - |\delta Z_s|^2] ds \\ &\leq |\delta \xi|^2 + \int_0^T [2(L_s^y + (L_s^z)^2)|\delta Y_s|^2 + 2|\delta Y_s| |\delta f_s| - \frac{1}{2}|\delta Z_s|^2] ds \\ &\Rightarrow \frac{1}{2} \|\delta Z\|_{\mathcal{H}^{2p}}^{2p} \leq C \left\{ \|\delta Y\|_{\mathcal{S}^{2p}}^{2p} + \|\delta Y\|_{\mathcal{S}^{4p}}^{2p} \mathbb{E} \left[ \left( \int_0^T [L_s^y + (L_s^z)^2] ds \right)^{2p} \right]^{\frac{1}{2}} \right. \\ &\quad \left. + \|\delta Y\|_{\mathcal{S}^{2p}}^p \mathbb{E} \left[ \left( \int_0^T |\delta f_s| ds \right)^{2p} \right]^{\frac{1}{2}} + \varepsilon \|\delta Y\|_{\mathcal{S}^{4p}}^{2p} + \frac{1}{\varepsilon} \|\delta Z\|_{\mathcal{H}^{2p}}^{2p} \right\}, \end{aligned}$$

where the last line follows from: reordering the terms, taking absolute values, exponentiating to power  $p$  and taking expectation, then use Itô's Isometry, the fact that  $\delta Y \in \mathcal{S}^\infty$  and Young's inequality for some  $\varepsilon > 0$ . Since  $Z^i \in \mathcal{H}_{BMO}$  for  $i \in \{1, 2\}$  the processes  $L^y, L^z \in \mathcal{H}_{BMO}$ , now with the appropriate choice of  $\varepsilon$ , estimate (3.5) follows.  $\square$

### 3.3 A remark on Theorems 3.9 and 3.10

Although we managed to show that a solution to BSDE (3.1) exists under Assumption 3.2 for any  $\gamma \geq 0$ , we were only able to show a comparison result when  $\gamma$  was small enough (see (3.14)). This is due to a limitation of the mathematical tools available, namely that a small  $\gamma$  ensured that the  $\Gamma$ 's in (3.11) and (3.16) had the required integrability properties.

But there are situations where one is able to show uniqueness for any value of  $\gamma$ . We next discuss about the possibility of linearizing a driver, e.g.  $f(y, z) = g(y)|z|^2$  with a convenient  $g$  function, through an invertible transformation. The idea is similar to the usual quadratic driver case,  $f(z) = |z|^2/2$  and the transformation  $P_t = \exp Y_t$ .

#### A linearization transformation

Linearization for standard quadratic BSDE is well understood: through the transformation  $P_t = \exp\{\gamma Y_t\}$  and  $Q_t = Z_t P_t$ , the BSDE

$$Y_t = \xi + \int_t^T \frac{\gamma}{2} |Z_s|^2 ds - \int_t^T Z_s dW_s$$

becomes

$$P_t = e^{\gamma\xi} - \int_t^T Q_s dW_s \Rightarrow Y_t = \frac{1}{\gamma} \log \mathbb{E}[e^{\gamma\xi} | \mathcal{F}_t].$$

It is possible to use the same type of arguments and find linearizing transformations for BSDE with drivers of the form  $f(y, z) = g(y)|z|^2$  for an integrable function  $g$ , by solving the ODE<sup>10</sup>

$$g(y)\Phi'(y) - \frac{1}{2}\Phi''(y) = 0, \quad \Phi(0) = 0, \quad \Phi'(0) = 1.$$

The above ODE is solved by

$$\Phi'(y) = \exp \left\{ \int_0^y 2g(x)dx \right\} \quad \text{and} \quad \Phi(y) = \int_0^y \Phi'(x)dx.$$

Since  $g$  is integrable by assumption, we have  $\Phi'(y) \neq 0$  for any  $y \in \mathbb{R}$  so that  $\Phi$  is invertible and hence the unique solution of the transformed BSDE yields the unique solution of the original BSDE.

We also point out that this type of transformation works well in the one dimensional setting. In the multidimensional case one quickly finds problems, see the Remark 4.9 below.

## 4 FBSDEs and PDEs

In this section we discuss the framework of Markovian FBSDE and their connection to PDEs. The results here open way to the later results on PDE perturbation and large deviations.

Here we particularize the previous results to the framework of *decoupled* Forward-Backward SDE (FBSDE), where we assume that the randomness of the terminal condition and driver have origin in a diffusion process. Namely, we define an FBSDE, with solution  $(X, Y, Z)$ , as the following system of stochastic equations: let  $m \in \mathbb{N}$  and take  $(t, x) \in [0, T] \times \mathbb{R}^m$  and for  $s \in [t, T]$

$$X_s^{t,x} = x + \int_t^s b(r, X_r^{t,x})dr + \int_t^s \sigma(r, X_r^{t,x})dW_r, \quad (4.1)$$

$$Y_s^{t,x} = g(X_T^{t,x}) + \int_s^T f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x})dr - \int_s^T Z_r^{t,x}dW_r. \quad (4.2)$$

The involved functions  $b, \sigma, g$  and  $f$  satisfy:

**Assumption 4.1.**  $b : [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}^m$  and  $\sigma : [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}^{m \times d}$  are continuous functions for which there exists a constant  $K > 0$  such that  $\sup_{t \in [0, T]} \{ |b(t, 0)| + |\sigma(t, 0)| \} \leq K$ . Moreover  $b$  and  $\sigma$  satisfy standard Lipschitz conditions in the spatial variables.

<sup>10</sup>Apply Itô's formula to  $\Phi(Y_t)$  and organize the terms inside the Lebesgue integral in a convenient way. Note that we already know that  $Y \in \mathcal{S}^\infty$  and can easily verify that  $\int \Phi'(Y)ZdW$  is a true martingale.

**Assumption 4.2.**  $g : \mathbb{R}^m \rightarrow \mathbb{R}$  and  $f : [0, T] \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$  are continuous functions.  $g$  satisfies a standard Lipschitz condition and is uniformly bounded by a constant  $M$ . Let  $k \in \mathbb{N}$  and a positive constant  $K$  such that for all  $(t, x, y, z) \in [0, T] \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R}^d$

$$|f(t, x, y, z)| \leq K(1 + |y| + (1 + |y|^k)|z| + |z|^2). \quad (4.3)$$

Let  $t \in [0, T]$ ,  $x, x' \in \mathbb{R}^m$ ,  $y, y' \in \mathbb{R}$  and  $z, z' \in \mathbb{R}^d$  it holds that

$$\begin{aligned} & |f(t, x, y, z) - f(t, x', y', z')| \\ & \leq K \{ (1 + |y| + |y'| + |y|^k|z| + |y'|^k|z'| + |z|^2 + |z'|^2) |x - x'| \\ & \quad + (1 + (|y|^{k-1} + |y'|^{k-1})(|z| + |z'|)) |y - y'| \\ & \quad + (1 + |y|^k + |y'|^k + (|z| + |z'|)) |z - z'| \}. \end{aligned} \quad (4.4)$$

#### 4.1 Particularizing the main results - Decoupled FBSDE

We now state a particularization of the main theorems in the previous section to the framework of FBSDE.

**Proposition 4.3.** Take  $(t, x) \in [0, T] \times \mathbb{R}^m$  and let Assumptions 4.1 and 4.2 hold. Then there exists a unique triple  $(X^{t,x}, Y^{t,x}, Z^{t,x}) \in \mathcal{S}^p \times \mathcal{S}^\infty \times \mathcal{H}_{BMO}$  for any  $p \geq 2$  that solves FBSDE (4.1)-(4.2). Moreover,

$$\sup_{(t,x) \in [0,T] \times \mathbb{R}^m} \|Y^{t,x}\|_{\mathcal{S}^\infty} + \|Z^{t,x}\|_{\mathcal{H}_{BMO}} < \infty. \quad (4.5)$$

Furthermore,  $Y^{t,x}$  and  $Z^{t,x}$  are Markovian: there exist two Borel-measurable functions  $u : [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}$  and  $v : [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}^d$  such that

$$Y_s^{t,x} = u(s, X_s^{t,x}), \quad s \in [t, T], \quad \mathbb{P}\text{-a.s.} \quad Z_s^{t,x} = v(s, X_s^{t,x}), \quad ds \otimes \mathbb{P}\text{-a.s.}$$

*Proof.* Existence and uniqueness follow from standard SDE results and the results of the previous section.

Let  $(t, x) \in [0, T] \times \mathbb{R}^m$ . The uniform bounds of the  $\mathcal{S}^\infty$ - and  $\mathcal{H}_{BMO}$ -norms, with relation to  $t$  and  $x$ , follow from the uniform bounds of the growth assumption in the driver  $f$  and the arguments used to prove Theorem 3.9.

The constant  $K$  appearing in the domination (4.3) is independent of  $x$  and  $t$  and it is shown in the proof of Theorem 3.9 that the uniform bounds for the norms of  $Y$  and  $Z$  depend only on the constants  $K$  and  $M$  appearing in Assumptions 3.1 and 3.2, which are essentially the same constants appearing in Assumption 4.3.

The Markov property of  $(Y, Z)$  is also trivial. It follows from arguments similar to those in Lemma 4.1 in [EKPQ97] or Theorem 4.1.1 in [DR11].  $\square$

The next result states the continuous dependence of the solution in the Euclidean parameter  $x$ .

**Corollary 4.4.** *Take  $t \in [0, T]$  and  $x, x' \in \mathbb{R}^m$  and let Assumptions 4.1 and 4.2 hold. Then for any  $p \geq 2$  there exists a constant  $C > 0$  (independent of  $x$  and  $x'$ ) such that*

$$\mathbb{E} \left[ \sup_{s \in [t, T]} |Y_s^{t,x} - Y_s^{t,x'}|^p \right] + \mathbb{E} \left[ \left( \int_0^T |Z_s^{t,x} - Z_s^{t,x'}|^2 ds \right)^{\frac{p}{2}} \right] \leq C |x - x'|^p$$

For  $0 \leq t \leq s \leq T$  the mapping  $\mathbb{R}^m \ni x \mapsto Y_s^{t,x}$  admits a continuous modification where almost all sample paths are  $\alpha$ -Hölder continuous in  $\mathbb{R}^m$  for any  $\alpha \in (0, 1)$ . For  $(t, x) \in [0, T] \times \mathbb{R}^m$  the mapping  $s \mapsto Y_s^{t,x}(\omega)$  is continuous for  $\mathbb{P}$ -a.s.  $\omega \in \Omega$ .

*Proof.* For simplicity we take  $t = 0$ . The bounds (4.5) play a crucial role because we can now apply the results of Lemma 3.13 where the involved constants  $C$  are independent of the parameters  $x$  and  $x'$ . Then, for  $p \geq 2$  we have (recall that  $\|Y^x\|_{\mathcal{S}^\infty}$  is uniformly bounded in  $x$ )

$$\begin{aligned} & \|Y^x - Y^{x'}\|_{\mathcal{S}^p}^p \\ & \leq C \mathbb{E} \left[ \left( |g(X_T^x) - g(X_T^{x'})| \right. \right. \\ & \quad \left. \left. + \int_0^T |f(s, X^x, Y^x, Z^x) - f(s, X^{x'}, Y^x, Z^x)| ds \right)^{prq} \right]^{\frac{1}{qr}} \\ & \leq C \left\{ \|X^x - X^{x'}\|_{\mathcal{S}^{pqr}}^{pqr} + \right. \\ & \quad \left. + \mathbb{E} \left[ \left( \sup_{t \in [0, T]} |X_t^x - X_t^{x'}| \int_0^T [1 + |Z_s^x| + |Z_s^x|^2] ds \right)^{prq} \right]^{\frac{1}{qr}} \right\}. \end{aligned}$$

Using  $|z| \leq 1 + |z|^2$  along with the fact that  $\|Z^x\|_{\mathcal{H}_{BMO}}$  is uniformly bounded in  $x$ , the result follows from standard results of SDE<sup>11</sup>, namely that under Assumption 4.1 we have for any  $x, x' \in \mathbb{R}^m$  and  $p \geq 2$  that

$$\|X^x - X^{x'}\|_{\mathcal{S}^p}^p = \mathbb{E} \left[ \sup_{t \in [0, T]} |X_t^x - X_t^{x'}|^p \right] \leq C_p |x - x'|^p,$$

with  $C_p$  a constant independent of  $x$  and  $x'$ . The continuity of the map in  $t$  and  $x$  follows from the properties of the solution of the BSDE and Kolmogorov's continuity criterion for  $p$  large enough (see Theorem 2 in [Sch00]). So we can deduce the result concerning  $\|Z^x - Z^{x'}\|_{\mathcal{H}^p}^p$  from (3.5).  $\square$

The previous result hints that it is perhaps possible to make sense of the variation of  $Y^x$ , i.e. that  $x \mapsto Y^x$  is differentiable (in some sense).

Using a particular form of the results in [BC08] we can conclude the existence of all partial derivatives (in  $x \in \mathbb{R}^m$ ) of  $(Y^x, Z^x)$ . What is then left to argue is the existence of the total derivative of  $Y$ . We will work under the following assumption

<sup>11</sup>See for instance Theorem 1.2.5 of [DR11]

**Assumption 4.5.** *Let Assumptions 4.1 and 4.2 hold.  $b$ ,  $\sigma$  and  $g$  have continuous, uniformly bounded first order spatial derivatives.  $f$  is differentiable with continuous derivatives satisfying for some  $K > 0$  for all  $(t, x, y, z) \in [0, T] \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R}^d$*

$$\begin{aligned} |\nabla_x f(t, x, y, z)| &\leq K(1 + |y| + |y|^k |z| + |z|^2), \\ |\nabla_y f(t, x, y, z)| &\leq K(1 + |y|^{k-1} |z|), \\ |\nabla_z f(t, x, y, z)| &\leq K(1 + |y|^k + |z|). \end{aligned}$$

Let  $t \in [0, T]$ ,  $x \in \mathbb{R}^m$ ,  $\Theta^x = (X^x, Y^x, Z^x)$  be the solution of (4.1)-(4.2), and define, at least formally, the dynamics<sup>12</sup>

$$\nabla_x X_t^x = I_d + \int_0^t \nabla_x b(s, X_s^x) \nabla_x X_s^x ds + \int_0^t \nabla_x \sigma(s, X_s^x) \nabla_x X_s^x dW_s, \quad (4.6)$$

$$\begin{aligned} \nabla_x Y_t^x &= (\nabla_x g)(X_T^x) \nabla_x X_T^x - \int_t^T \nabla_x Z_s^x dW_s \\ &\quad + \int_t^T \langle (\nabla f)(s, \Theta_s^x), (\nabla_x X_s^x, \nabla_x Y_s^x, \nabla_x Z_s^x) \rangle ds. \end{aligned} \quad (4.7)$$

**Proposition 4.6.** *Let  $x \in \mathbb{R}^m$ ,  $p \geq 2$  and let Assumption 4.5 hold. Then the mapping  $\mathbb{R}^m \rightarrow \mathcal{S}^p(\mathbb{R}^m) \times \mathcal{S}^p(\mathbb{R}) \times \mathcal{H}^p(\mathbb{R}^d)$ ,  $x \mapsto (X^x, Y^x, Z^x)$  is differentiable in the norm topology and the derivative is the unique solution of FBSDE (4.6)-(4.7) in  $\mathcal{S}^p \times \mathcal{S}^p \times \mathcal{H}^p$  for any  $p \geq 2$ .*

*In particular, for  $x', x \in \mathbb{R}^m$  we have*

$$\lim_{x' \rightarrow x} \{ \|\nabla_x Y^{t,x} - \nabla_x Y^{t,x'}\|_{\mathcal{S}^p}^p + \|\nabla_x Z^{t,x} - \nabla_x Z^{t,x'}\|_{\mathcal{H}^p}^p \} = 0.$$

**Remark 4.7.** *As in Theorem 2.2 of [AIdR07] (or Section 3.1 of [DR11]) it is possible to show that there exists a function  $\Omega \times [t, T] \times \mathbb{R}^m \rightarrow \mathbb{R}$  mapping  $(\omega, s, x) \mapsto Y_s^{t,x}$  such that for almost all  $\omega$ ,  $Y_s^{t,x}$  is continuously differentiable in  $x$ , i.e. understanding  $x \mapsto Y_s^{t,x}(\omega)$  as a function in  $C^1(\mathbb{R}^m, \mathbb{R})$ !*

*For such a result one would need to ask that  $\nabla g$  satisfies a Lipschitz condition as well a Lipschitz like condition for the derivatives of  $f$ . Since the argumentation for such a result follows from a straightforward adaptation of the proofs in [AIdR07] or [DR11], we refrain from giving the details.*

The proof of Proposition 4.6 follows a methodology rather standard for this type setting and so it is given in the appendix.

## 4.2 Connection to PDEs

In the classical theory of FBSDE one is able to connect the solution of FBSDE (4.1)-(4.2) to the solution of a certain PDE. Formally, for  $0 \leq t \leq$

<sup>12</sup>For  $m \in \mathbb{N}$ ,  $I_m$  represents the  $m \times m$ -dimensional identity matrix.

$s \leq T$ ,  $x \in \mathbb{R}^m$  the FBSDE (4.1)-(4.2) is related to the PDE:

$$\begin{aligned} \partial_t u(t, x) + \mathcal{L}u(t, x) + f(t, x, u(t, x), (\nabla_x u \sigma)(t, x)) &= 0, \\ u(T, x) &= g(x), \end{aligned} \quad (4.8)$$

with the second order differential operator  $\mathcal{L}$  being the infinitesimal semi-group generator of the Markov process  $(X_s^{t,x})_{s \in [t, T]}$  given by

$$\mathcal{L} := \sum_{i=1}^m b_i(t, x) \partial_{x_i} + \frac{1}{2} \sum_{i,j=1}^m (\sigma \sigma^T)_{i,j}(t, x) \partial_{x_i x_j}.$$

The relation between the FBSDE and the PDE is given by the identities:

$$Y_s^{t,x} = u(s, X_s^{t,x}) \quad \text{and} \quad Z_s^{t,x} = (\nabla_x u \sigma)(s, X_s^{t,x}), \quad (4.9)$$

where the second relation holds provided the derivatives of  $u$  are appropriately defined. Indeed, under Assumptions 4.1 and 4.2, we get from Corollary 4.4 that

$$|u(t, x) - u(t, x')|^p = |Y_t^{t,x} - Y_t^{t,x'}|^p \leq C|x - x'|^p, \quad (4.10)$$

and hence  $u(t, x)$  is a Lipschitz function in its spatial variables. With the results from Proposition 4.6 we have also that  $\nabla_x u$  exists and is continuous (in the spatial variable).

**Remark 4.8.** *Particularizing the forward diffusion by assuming  $b = 0$ ,  $\sigma = \sqrt{2}$  and  $f(t, y, z) = \nu y z$ , one obtains the known Burgers' PDE.*

**Remark 4.9** (Multidimensional case of  $f(y, z) = y|z|^2$ ). *Related to a particular case of the Theorem 3.9, we make a small remark concerning the PDE representation of such BSDE (when paired with a SDE) in higher dimensions. There exists a phenomenon of finite-time gradient blow-up in PDEs: [CDY92] (also presented as Theorem III.6.14 in [Str08]) consider mappings  $u$  from the closed unit disk in  $\mathbb{R}^2$  into the unit sphere in  $\mathbb{R}^3$  which satisfy*

$$\partial_t u = \Delta u + |\nabla_x u|^2 u, \quad u(0, x) = u_0(x), \quad u(t, \cdot)|_{\partial D^2} = u_0|_{\partial D^2}. \quad (4.11)$$

*They show that for some smooth and bounded boundary condition  $u_0$ , the solution of (4.11) blows up in finite time, i.e., the maximal existence interval  $[0, T)$  has a finite  $T$ .*

*Moreover, in [BDPvH02] for the same PDE*

$$\begin{aligned} \partial_t u &= \Delta u + |\nabla_x u|^2 u \text{ in } \mathcal{Q} := \Omega \times \mathbb{R}^+ \\ u &= u_0 \text{ on } \partial\Omega \times \mathbb{R}^+ \quad \text{and} \quad u(0, x) = u_0(x) \text{ for } x \in \Omega, \end{aligned}$$

*with  $\Omega$  a bounded domain in  $\mathbb{R}^n$  with smooth boundary, the authors show that  $u_0 \in C^\infty$  exists such that an infinite number of solutions to the PDE exist.*

### 4.3 Viscosity solution

As far as the relation between PDE and FBSDE go, the natural counterpart of the FBSDE solution is a PDE solution in the viscosity sense via (4.9). Observe that we have not mentioned in any way an ellipticity assumption for  $\sigma$ , which is consistent with viscosity theory. Although we have a unique solution for the FBSDE, the FBSDE is only to provide existence for the PDE while uniqueness of the PDE has to be shown via a comparison result for PDE. It is unclear whether such a result holds in this framework that allows for cross term nonlinearities like  $yz$  since this type of nonlinearities implies that the Hamiltonian arising from (4.8) is not a *proper* function. Nonetheless, such a comparison result exists in the probabilistic framework (see our Theorem 3.12).

We do not give many details below, but as expected, one is able to prove that the  $Y$  component of the solution to the FBSDE (4.1)-(4.2) is indeed a viscosity solution to PDE (4.8).

**Theorem 4.10.** *Take  $(t, x) \in [0, T] \times \mathbb{R}^m$  and let Assumptions 4.1 and 4.2 hold. Then the mapping  $u : [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}$  defined by  $u(t, x) := Y_t^{t, x}$  is a uniformly bounded continuous function of its temporal and spatial variables and a viscosity solution of (4.8).*

The proof of this theorem is the same as that of Theorem 3.2 in [Par99] except for the application of the comparison theorem in the last step of the proof (where we use our Theorem 3.12). We postpone the proof to the appendix.

## 5 PDE Perturbation and large deviations

In this section we look at the relation between FBSDE and PDE perturbation. In a first approach, the form of the perturbation is quite amenable to a in-depth analysis including the possibility of showing a large deviations principle for the solution of the FBSDE as the perturbation vanishes. The second step looks at another setting of PDE perturbation and only partial results, within the framework of the previous sections, are obtained. We then stretch the theory to recover some other known results on a certain perturbed transport PDE and the Navier-Stokes equations for incompressible fluid flow.

In this section we establish a large deviations principle for the FBSDE (4.1), (4.2). This type of result has its relevance in the theory of perturbation of PDEs. We discuss two types of perturbations, while the first leads to classical results (and relates to the theory of  $\varepsilon$ -vanishing viscosity solutions) the second shows the limitations of the techniques used.



## 5.1 Canonical perturbation in FBSDE framework

Throughout let  $x \in \mathbb{R}^m$ ,  $0 \leq t \leq s \leq T$  and  $\varepsilon \in (0, 1]$ . Moreover, we denote by  $C$  a constant that may change from line to line but it is *always independent* of the parameters  $\varepsilon$ ,  $x$  or  $t$ .

Throughout we work with Assumptions 4.1 and 4.2 and from the context it is clear that the constants involved in the growth bounds are *independent* of  $\varepsilon$ ! At a later stage we will discuss the situation  $|f(y, z)| \leq K|y||z|/\varepsilon$ .

We work with the family of equations indexed in  $\varepsilon$  and following the dynamics

$$X_s^{t,x,\varepsilon} = x + \int_t^s b(r, X_r^{t,x,\varepsilon})dr + \sqrt{\varepsilon} \int_t^s \sigma(r, X_r^{t,x,\varepsilon})dW_r. \quad (5.1)$$

It is known that as  $\varepsilon$  vanishes, the solution  $X^{t,x,\varepsilon}$  converges to  $X^{t,x,0}$  (see Lemma 5.1 below), which solves the deterministic equation

$$X_s^{t,x,0} = x + \int_t^s b(r, X_r^{t,x,0})dr. \quad (5.2)$$

Moreover, under some extra assumptions, that the law induced by  $X^{t,x,\varepsilon}$  satisfies a large deviations principle (LDP) (see Definition 5.4 below).

We want to show that the FBSDE with dynamics

$$Y_s^{t,x,\varepsilon} = g(X_T^{t,x,\varepsilon}) - \int_s^T Z_r^{t,x,\varepsilon}dW_r + \int_s^T f(r, X_r^{t,x,\varepsilon}, Y_r^{t,x,\varepsilon}, Z_r^{t,x,\varepsilon})dr, \quad (5.3)$$

converges to the deterministic backward equation

$$Y_s^{t,x,0} = g(X_T^{t,x,0}) + \int_s^T f(r, X_r^{t,x,0}, Y_r^{t,x,0}, 0)dr, \quad (5.4)$$

and that the law induced by  $Y^{t,x,\varepsilon}$  satisfies an LDP as well.

We start with a small result on the properties of  $X^{t,x,\varepsilon}$ .

**Lemma 5.1.** *Under Assumption 4.1 and for any  $\varepsilon \in (0, 1]$  equations (5.1) and (5.2) have unique solutions,  $X^{t,x,\varepsilon} \in \mathcal{S}^p$  for any  $p \geq 2$  and  $X^{t,x,0} \in C^0([t, T] \times \mathbb{R}^m, \mathbb{R}^m)$ . Moreover, for any  $p \geq 2$ , for some  $C_p > 0$  it holds that*

$$\sup_{\varepsilon \in (0, 1]} \|X^{t,x,\varepsilon}\|_{\mathcal{S}^p}^p \leq C_p(1 + |x|^p) \quad \text{and} \quad \|X^{t,x,\varepsilon} - X^{t,x,0}\|_{\mathcal{S}^p}^p \leq C_p(1 + |x|^p)\varepsilon^p.$$

Furthermore,  $\lim_{\varepsilon \rightarrow 0} X^{t,x,\varepsilon} = X^{t,x,0}$  a.s.

*Proof.* These results are quite standard from either SDE or ODE theory and follow from the good properties of the  $b$  and  $\sigma$  functions. See e.g. [DZ09].  $\square$

We next establish properties of  $(Y^{t,x,\varepsilon}, Z^{t,x,\varepsilon})$  and  $Y^{t,x,0}$ .

**Proposition 5.2.** *Let Assumptions 4.1 and 4.2 hold. Then, a unique solution  $(X^{t,x,\varepsilon}, Y^{t,x,\varepsilon}, Z^{t,x,\varepsilon})$  of FBSDE (5.1), (5.3) in  $\mathcal{S}^p \times \mathcal{S}^\infty \times \mathcal{H}_{BMO}$  for any  $p \geq 2$  exists. Moreover,*

$$\sup_{\varepsilon \in (0,1]} \sup_{(t,x) \in [0,T] \times \mathbb{R}^m} \{ \|Y^{t,x,\varepsilon}\|_{\mathcal{S}^\infty} + \|Z^{t,x,\varepsilon}\|_{\mathcal{H}_{BMO}} \} < \infty.$$

The mapping  $(t, x) \mapsto Y^{t,x,\varepsilon}$  has a continuous modification.

Additionally, the ODE (5.4) has a unique solution  $Y^{t,x,0} \in C_b^0([t, T] \times \mathbb{R}^m, \mathbb{R})$ . Moreover, there exists a constant  $C > 0$  such that

$$|Y_t^{t,x,0} - Y_t^{t,x',0}| \leq C|x - x'|, \quad t \in [0, T], \quad x, x' \in \mathbb{R}^m. \quad (5.5)$$

*Proof.* Existence and uniqueness of a solution to (5.3) for  $\varepsilon \in (0, 1]$  follow from Theorems 3.9 and 3.10. An argumentation similar to that of the proof of (4.5) yields the uniform bounds in  $\varepsilon$ . The continuity result follows from Corollary 4.4.

Concerning  $Y^{t,x,0}$  from (5.4), we easily see that  $g$  is still a uniformly bounded function and  $F(s, y) = f(s, X_s^{t,x,0}, y, 0)$  for  $(s, y) \in [t, T] \times \mathbb{R}$  is an  $\mathcal{F}_s$ -adapted continuous function (in  $s$  and  $y$ ) satisfying (simply use (4.3) and (4.4)) a linear growth and Lipschitz condition (in the  $y$ -variable). After a time reversal, standard ODE theory yields (e.g. Picard-Lindelöf's theorem) the existence and uniqueness of a continuous solution  $Y^{t,x,0}$  to (5.4). One then obtains that  $Y^{t,x,0}$  is uniformly bounded.

The estimate for the difference  $(Y^{t,x,0} - Y^{t,x',0})$  follows easily from the same estimate for  $(X^{t,x,0} - X^{t,x',0})$ , Gronwall's inequality combined with the uniform boundedness of  $Y^{t,x,0}$  and (4.4).  $\square$

The next result states the convergence of  $(Y^{t,x,\varepsilon}, Z^{t,x,\varepsilon})$  to  $(Y^{t,x,0}, 0)$  as  $\varepsilon$  vanishes. Notice that we can always interpret the deterministic equation (5.4) as a FBSDE with solution  $(Y^{t,x,0}, 0)$ .

**Proposition 5.3.** *Let Assumptions 4.1 and 4.2 hold. Then for any  $p \geq 2$  there exists a constant  $C_p > 0$  such that*

$$\sup_{(t,x) \in [0,T] \times \mathbb{R}^m} \left\{ \mathbb{E} \left[ \sup_{t \leq s \leq T} |Y_s^{t,x,\varepsilon} - Y_s^{t,x,0}|^p \right] + \mathbb{E} \left[ \left( \int_t^T |Z_r^{t,x,\varepsilon}|^2 dr \right)^p \right] \right\} \leq C_p \varepsilon^p.$$

We emphasize that  $C_p$  is independent of  $t$ ,  $x$  or  $\varepsilon$ .

*Proof.* Let  $p \geq 2$ . We start by interpreting the pair  $(Y^{t,x,0}, 0)$  as the solution of a deterministic BSDE. This interpretation allows us to use Corollary 3.13 along with the uniform boundedness of  $Y^{t,x,\varepsilon}$  and the assumptions to obtain

$$\|Y^{t,x,\varepsilon} - Y^{t,x,0}\|_{\mathcal{S}^p}^p \leq C_\beta \|X^{t,x,\varepsilon} - X^{t,x,0}\|_{\mathcal{S}^{\beta p}}^p \leq C_p \varepsilon^p,$$

where  $\beta$  is related to the uniformly bounded (in  $\varepsilon$ ) BMO norm of  $Z^{t,x,\varepsilon} * W$  and the result follows from Lemma 5.1.

From (3.5) and the results already proved we easily get the sought result concerning  $\|Z^{t,x,\varepsilon} - Z^{t,x,0}\|_{\mathcal{H}^p}^p = \|Z^{t,x,\varepsilon}\|_{\mathcal{H}^p}^p$ .  $\square$

## The corresponding perturbed PDE problem

By solving (5.1) and (5.3) one is able to solve, at least in the viscosity sense, the corresponding PDE, namely in view of (4.8) and Theorem 4.10 one has for  $(t, x) \in [0, T] \times \mathbb{R}^m$

$$\begin{aligned} \partial_t u^\varepsilon(t, x) + \mathcal{L}^\varepsilon u^\varepsilon(t, x) + f(t, x, u^\varepsilon(t, x), \sqrt{\varepsilon}(\nabla_x u^\varepsilon \sigma)(t, x)) &= 0, \\ u^\varepsilon(T, x) &= g(x), \end{aligned}$$

with  $\mathcal{L}^\varepsilon := (b \cdot \nabla)(t, x) + \frac{\varepsilon}{2} \sum_{i,j=1}^m (\sigma \sigma^T)_{i,j}(t, x) \partial_{x_i x_j}$ . And as  $\varepsilon$  vanishes one is lead to the first order hyperbolic PDE for  $(t, x) \in [0, T] \times \mathbb{R}^m$

$$\partial_t u^0(t, x) + (b \cdot \nabla u^0)(t, x) + f(t, x, u^0(t, x), 0) = 0, \quad u^0(T, x) = g(x),$$

or after a time inversion,  $t \mapsto T - t$ , with  $(t, x) \in [0, T] \times \mathbb{R}^m$

$$\partial_t u^0(t, x) = (b \cdot \nabla u^0)(t, x) + f(t, x, u^0(t, x), 0), \quad u^0(0, x) = g(x).$$

## 5.2 The large deviations result

We now explain the large deviations principle (LDP). For a deeper analysis on large deviations we point the reader to [FW98], [DZ09], [FK06] and references therein. In this section we extend the results of [DR07] and [Rai06] to another class of FBSDE.

We denote  $C^0([a, b], \mathbb{R}^m)$  (for  $0 \leq a \leq b < +\infty$ ) the set of continuous functions on the interval  $[0, T]$  with values in  $\mathbb{R}^m$  and we consider in this space the uniform norm

$$\rho_{[a,b]}(\phi) := \sup_{t \in [a,b]} |\phi(t)|. \quad (5.6)$$

**Definition 5.4.** *The family of processes  $(X_t^\varepsilon)_{t \in [0, T]}$  depending on a parameter  $\varepsilon$  is said to satisfy a large deviations principle (LDP) with rate function  $S(\Psi)$  if the following conditions hold for every Borel set  $A \subset C([0, T])$*

$$\begin{aligned} \limsup_{\varepsilon \rightarrow 0} \varepsilon \log \mathbb{P}[X^\varepsilon \in A] &\leq \inf_{\Psi \in Cl(A)} S(\Psi) \\ \liminf_{\varepsilon \rightarrow 0} \varepsilon \log \mathbb{P}[X^\varepsilon \in A] &\geq - \inf_{\Psi \in Int(A)} S(\Psi), \end{aligned}$$

where  $Cl(A)$  denotes the closure of the set  $A$  and  $Int(A)$  denotes the interior of the set  $A$ .

Before we state that  $X^{t,x,\varepsilon}$  satisfies an LDP, we recall the Cameron-Martin space  $H$  of absolutely continuous functions.

**Definition 5.5** (Cameron-Martin space). *Let  $k \in \mathbb{N}$ , the space  $H([t, T], \mathbb{R}^k)$  is defined as the space of continuous functions  $\phi \in C^0([t, T], \mathbb{R}^k)$  for which there exists a square integrable function  $\dot{\phi}$  such that  $\phi_s = \phi_t + \int_t^s \dot{\phi}_r dr$  for  $0 \leq t \leq s \leq T$  (i.e.  $\phi$  is absolutely continuous).*

We now make a new assumption, stronger than Assumption 4.1, for the result stating that  $X^{t,x,\varepsilon}$  satisfies an LDP.

**Assumption 5.6.** *Let Assumption 4.1 hold.  $b$  and  $\sigma$  satisfy  $b(t,x) = b(x)$  and  $\sigma(t,x) = \sigma(x)$  for all  $(t,x) \in [0,T] \times \mathbb{R}^m$  and are uniformly bounded.*

**Theorem 5.7** (Theorem 5.6.7 of [DZ09]). *Let Assumption 5.6 hold. Then the solution  $X^{t,x,\varepsilon}$  of (5.1) satisfies, as  $\varepsilon$  goes to zero, a large deviations principle in the space  $C^0([t,T], \mathbb{R}^m)$  associated to the rate function  $I_x$  defined for any  $\phi \in C^0([t,T], \mathbb{R}^m)$  by*

$$I_x(\phi) := \inf \left\{ \frac{1}{2} \int_t^T |\dot{v}_r|^2 dr, v \in H([t,T], \mathbb{R}^d) \text{ such that} \right. \\ \left. \phi_s = x + \int_t^s b(\phi_r) dr + \int_t^s \sigma(\phi_r) \dot{v}_r dr, s \in [t,T] \right\},$$

with the convention that  $\inf \emptyset = +\infty$ .

**Remark 5.8.** *We point out that if  $\sigma\sigma^T$  is invertible, then the rate function functional can be written in a simplified manner, more precisely*

$$I_x(\phi) := \begin{cases} \frac{1}{2} \int_t^T \mathcal{Q}_{\phi_r}^* [\dot{\phi}_r - b(\phi_r)] dr & , \text{ if } \phi \in H([t,T], \mathbb{R}^m) \text{ and } \phi_t = x \\ +\infty & , \text{ otherwise,} \end{cases}$$

where  $\mathcal{Q}^*$  is defined by  $\mathcal{Q}_u^*[v] := \langle v, (\sigma\sigma^T)^{-1}(u)v \rangle$  for all  $u, v \in \mathbb{R}^m$ .

We now define a set of operators from and onto the space of continuous functions that, combined with the contraction principle<sup>13</sup> (see Theorem 2.4 in [Var84] or Theorems 4.2.1 and 4.2.21 of [DZ09]), will allow us to conclude that  $Y^{t,x,\varepsilon}$  satisfies an LDP.

**Definition 5.9.** *The operator  $F^\varepsilon$  is defined as*

$$F^\varepsilon : C^0([t,T], \mathbb{R}^m) \rightarrow C^0([t,T], \mathbb{R}) \\ \phi \mapsto F^\varepsilon(\phi) = u^\varepsilon(\cdot, \phi),$$

where for  $\varepsilon \in (0,1]$  the function  $u^\varepsilon$  is defined by  $u^\varepsilon(t,x) := Y_t^{t,x,\varepsilon}$  the first component of the solution to (5.3). And for  $\varepsilon = 0$ ,  $u^0(t,x)$  is given by  $Y_t^{t,x,0}$  solution to (5.4).

We observe that for all  $0 \leq t \leq s \leq T$ ,  $x \in \mathbb{R}^m$  and  $\varepsilon \in [0,1]$  we have  $Y_s^{t,x,\varepsilon} = F^\varepsilon(X^{t,x,\varepsilon})(s) = u^\varepsilon(s, X_s^{t,x,\varepsilon})$  (see Theorem 4.3). We now state and prove the result concerning the LDP satisfied by the law induced by  $Y^{t,x,\varepsilon}$ .

<sup>13</sup>In rough, let  $\Theta$  and  $\Xi$  be two complete separable metric spaces and  $f_\varepsilon : \Theta \mapsto \Xi$  a family of continuous functions in those metric spaces. If  $\lim_{\varepsilon \rightarrow 0} f_\varepsilon = f$  exists uniformly over compact sets, then the contraction principle states that if the  $\Theta$  valued process  $X_\varepsilon$ , satisfies an LDP with rate function  $I$  then  $Y_\varepsilon = f_\varepsilon(X_\varepsilon)$  also satisfies an LDP with the same rate and with a raw rate function  $J(y) = I(f^{(-1)}(y))$ . Moreover, if  $I$  is a good rate function then so is  $J$ . See Section 2 of [Var84] or Section 4.2 of [DZ09].

**Theorem 5.10.** *Let Assumption 5.6 and 4.2 hold. Then the process  $Y^{t,x,\varepsilon}$  satisfies an LDP in  $C^0([t, T], \mathbb{R})$  with rate function  $\widehat{I}_x$  defined for any  $\psi \in C^0([t, T], \mathbb{R})$  by*

$$\widehat{I}_x(\psi) := \inf \left\{ I_x(\phi) : \phi \in H([t, T], \mathbb{R}^m) \text{ such that } \psi = F^0(\phi) \right\},$$

with the convention that  $\inf \emptyset = +\infty$  and where  $I_x(\cdot)$  is the rate function from Theorem 5.7.

**Remark 5.11.** *If the mapping  $\sigma\sigma^T$  is invertible then the rate function  $\widehat{I}_x$  can be expressed in an alternative fashion using the operator  $\mathcal{Q}^*$  introduced in Remark 5.8, namely for  $\psi \in C^0([t, T], \mathbb{R})$  and  $x \in \mathbb{R}^m$*

$$\widehat{I}_x(\psi) = \inf \left\{ \frac{1}{2} \int_t^T \mathcal{Q}_{\phi_r}^* [\dot{\phi}_r - b(\phi_r)] dr, \text{ over} \right. \\ \left. \phi \in H([t, T], \mathbb{R}^m), \phi_t = x, \psi_s = u^0(s, \phi_s), s \in [t, T] \right\}.$$

*Proof of Theorem 5.10.* This proof uses the well known contraction principle, see Theorem 2.4 in [Var84]. To apply the contraction principle we need only to show that  $F^\varepsilon$  for  $\varepsilon \in [0, 1]$  are uniformly continuous operators from  $C^0([t, T], \mathbb{R}^m)$  onto  $C^0([t, T], \mathbb{R})$ , and moreover, that  $F^\varepsilon$  converges uniformly to  $F^0$  over all compact sets of  $C([t, T], \mathbb{R}^m)$  as  $\varepsilon$  vanishes. Notice from Corollary 4.4 that the mappings  $x \mapsto Y_s^{t,x,\varepsilon}$  and  $s \mapsto Y_s^{t,x,\varepsilon}$  are continuous ( $\mathbb{P}$ -a.s.).

**Step 1 - Continuity of  $F^\varepsilon$ :** Let  $\varepsilon \in [0, 1]$ . From (4.10) and (5.5) we have that  $x \mapsto u^\varepsilon(t, x)$  is Lipschitz continuous. Let  $\mathcal{K}^0$  be a compact subset of  $C^0([t, T], \mathbb{R}^m)$  and let  $(\phi^n)_{n \in \mathbb{N}}$  be a sequence of functions from  $\mathcal{K}^0$  that converge to some  $\phi \in \mathcal{K}^0$  in the uniform norm, i.e.  $\lim_{n \rightarrow \infty} \rho_{[t, T]}(\phi^n - \phi) = 0$ , where  $\rho$  is defined in (5.6). From the definition of  $F^\varepsilon$ , (4.10), (5.5) and the continuity of  $Y^{t,\cdot}$  we have

$$\begin{aligned} [\rho_{[t, T]}(F^\varepsilon(\phi^n) - F^\varepsilon(\phi))]^2 &= \sup_{s \in [t, T]} |u^\varepsilon(s, \phi_s^n) - u^\varepsilon(s, \phi_s)|^2 \\ &= \sup_{s \in [t, T]} |Y_s^{s, \phi_s^n, \varepsilon} - Y_s^{s, \phi_s, \varepsilon}|^2 \\ &\leq C \sup_{s \in [t, T]} |\phi_s^n - \phi_s|^2 = C [\rho_{[t, T]}(\phi^n - \phi)]^2 \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

Note that the constant  $C$  present in the estimate of Corollary 4.4 is independent of  $x$  or  $\phi$ .

**Step 2 - Convergence over compacts:** To prove that  $F^\varepsilon$  converges to  $F^0$  uniformly over the compact sets  $\mathcal{K}^0$  of  $C^0([t, T], \mathbb{R}^m)$  we start by defining the set  $\mathcal{I} := \{\phi_s : \phi \in \mathcal{K}^0, s \in [t, T]\}$ , as the set of points consisting of the images of the  $\phi \in \mathcal{K}^0$  in  $\mathbb{R}^m$  for all  $s \in [t, T]$ . Remark that  $\mathcal{I}$  is a compact set of  $\mathbb{R}^m$ .

Now, by Proposition 5.3 (in the particular case of  $s = t$ ) and taking the set  $\mathcal{I}$  into account, there exists a constant  $C > 0$  such that

$$\begin{aligned} & \sup_{\phi \in \mathcal{K}^0} \left[ \rho_{[t,T]}(F^\varepsilon(\phi) - F^0(\phi)) \right]^2 \\ &= \sup_{\phi \in \mathcal{K}^0} \sup_{s \in [t,T]} |u^\varepsilon(s, \phi_s) - u^0(s, \phi_s)|^2 = \sup_{\phi \in \mathcal{K}^0} \sup_{s \in [t,T]} |Y_s^{s, \phi_s, \varepsilon} - Y_s^{s, \phi_s, 0}|^2 \\ &\leq \sup_{x \in \mathcal{I}} \sup_{s \in [t,T]} |Y_s^{s, x, \varepsilon} - Y_s^{s, x, 0}|^2 \leq C \varepsilon^2 \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0. \end{aligned}$$

The result now follows from the already mentioned contraction principle.  $\square$

### 5.3 PDE perturbation and Burgers' equation with damping

Here we want to discuss the following PDE for  $\lambda > 0$ ,  $a \in \mathbb{R}$ ,  $x \in \mathbb{R}$  where  $u : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$  and  $g \in C_b^0$

$$\partial_t u^\varepsilon + a u^\varepsilon \nabla_x u^\varepsilon + \lambda u^\varepsilon = \varepsilon \Delta u^\varepsilon, \quad u^\varepsilon(0, x) = g(x). \quad (5.7)$$

Or, after a canonical time inversion  $t \mapsto T - t$  for  $t \in [0, T]$

$$\partial_t u^\varepsilon + \varepsilon \Delta u^\varepsilon = a u^\varepsilon \nabla_x u^\varepsilon + \lambda u^\varepsilon, \quad u^\varepsilon(T, x) = g(x). \quad (5.8)$$

Via the already discussed Feynman-Kac formula we can write (at least formally) the corresponding FBSDE system for  $x \in \mathbb{R}$  and  $0 \leq t \leq s \leq T$

$$\begin{aligned} X_s^{t,x,\varepsilon} &= x + \sqrt{2\varepsilon}(W_s - W_t), \\ Y_s^{t,x,\varepsilon} &= g(X_T^{t,x,\varepsilon}) - \int_s^T Z_r^{t,x,\varepsilon} dW_r - \int_s^T \left[ \frac{a}{\sqrt{2\varepsilon}} Y_r^{t,x,\varepsilon} Z_r^{t,x,\varepsilon} + \lambda Y_r^{t,x,\varepsilon} \right] dr. \end{aligned} \quad (5.9)$$

The appearance of the  $1/\sqrt{2\varepsilon}$  weight is due to the representation  $Z = (\nabla_x u \sigma)(\cdot, X)$  with  $\sigma = \sqrt{2\varepsilon}$ ; see (4.9). This type of FBSDE still fits in the setting we discussed in the first sections, but the presence of the term  $1/\varepsilon$  creates complications if we consider the FBSDE not just for one single fixed  $\varepsilon$ .

For each fixed  $\varepsilon \in (0, 1]$  we obtain from Theorem 4.3 existence and uniqueness of the triple  $(X^{t,x,\varepsilon}, Y^{t,x,\varepsilon}, Z^{t,x,\varepsilon})$  solving the FBSDE above in  $\mathcal{S}^p \times \mathcal{S}^\infty \times \mathcal{H}_{BMO}$  for any  $p \geq 2$ . The real interesting result is the uniform boundedness in  $\varepsilon$  of  $\|Y^{t,x,\varepsilon}\|_{\mathcal{S}^\infty}$ .

**Lemma 5.12.** *Let  $(t, x) \in [0, T] \times \mathbb{R}$  then in the above defined framework*

$$\sup_{x \in \mathbb{R}, \varepsilon \in (0, 1], t \in [0, T]} \|Y^{t,x,\varepsilon}\|_{\mathcal{S}^\infty} < \infty.$$

*Proof.* From the original existence Theorem 3.9 and the lemmata used in its proof, we easily obtain for each fixed  $\varepsilon \in (0, 1]$  that

$$\sup_{x \in \mathbb{R}, t \in [0, T]} \|Y^{t, x, \varepsilon}\|_{\mathcal{S}^\infty} < \infty.$$

We sketch the argument concerning the uniform boundedness on  $\varepsilon$  since it is based on an argument we have already seen in the proof of Lemma 3.16. We again use the linearization and measure change argument. Define processes  $e_t = \exp\{-\lambda t\}$  and  $b_t^\varepsilon := aY_t^\varepsilon/\sqrt{2\varepsilon}$ . Since  $Y^\varepsilon$  is a bounded process it is clear that  $b^\varepsilon * W \in BMO$  and we can define a new probability measure  $\mathbb{Q}^\varepsilon$  with Radon-Nikodym density  $d\mathbb{Q}^\varepsilon/d\mathbb{P} = \mathcal{E}(b^\varepsilon * W)$ , with relation to which  $W^\varepsilon = W - \int_0^\cdot b_r^\varepsilon dr$  is a Brownian motion.

Applying Itô's formula to  $e_t Y_t^\varepsilon$  and changing to the measure  $\mathbb{Q}^\varepsilon$  we obtain

$$\begin{aligned} e_s Y_s^{t, x, \varepsilon} &= e_T g(X_T^{t, x, \varepsilon}) - \int_s^T e_r Z_r^{t, x, \varepsilon} dW_r^\varepsilon \\ &\Rightarrow Y_s^{t, x, \varepsilon} = e^{-\lambda(T-s)} \mathbb{E}^{\mathbb{Q}^\varepsilon} [g(X_T^{t, x, \varepsilon}) | \mathcal{F}_s] \\ &\Rightarrow \sup_{x \in \mathbb{R}} \sup_{\varepsilon \in (0, 1]} \left\| \sup_{s \in [t, T]} |Y_s^{t, x, \varepsilon}| \right\|_{L^\infty} \leq \sup_{x \in \mathbb{R}} |g(x)| < \infty, \end{aligned}$$

where the last line follows from the uniform boundedness of  $g$  and from  $e^{-\lambda(T-s)} \leq 1$  for all  $s \in [0, T]$ .  $\square$

This result hints that the boundedness of the solution carries over to the limit (as  $\varepsilon \rightarrow 0$ ). On the other hand, obtaining a similar estimate for  $Z^{t, x, \varepsilon}$  in the  $\mathcal{H}_{BMO}$ -norm is impossible without extra assumptions.

First remark that the limiting PDE of (5.7) as  $\varepsilon \rightarrow 0$  is the first order equation:

$$\partial_t v + av \nabla_x v + \lambda v = 0, \quad v(0, x) = g(x), \quad (t, x) \in [0, T] \times \mathbb{R}.$$

For this type of transport PDE it is well known that even with smooth initial boundary data the solution can develop shocks in finite time which translates into an explosion of the spatial derivative in finite time, see [Ter10]. An estimate on  $\nabla_x u^\varepsilon$  (or  $Z$ ) independent of  $\varepsilon$  would mean that the derivative remains stable (non-explosive) as we decrease the *regularizing* impact of the perturbation.

### Getting help from fully coupled FBSDE

If one assumes that  $\lambda \geq K^2 a^2 T$  then it is possible to prove via *fully coupled* FBSDE arguments that  $\sup_{\varepsilon \in (0, 1]} \|\nabla_x u^\varepsilon\|_\infty \leq \sqrt{2} K$ . This condition is, in some cases, close to that of Theorem 1 in [Ter10] and it roughly means that if a strong enough damping is present in the PDE (via the term  $\lambda u^\varepsilon$ ) then the cross term  $u^\varepsilon \nabla_x u^\varepsilon$  does not lead to shocks up to time  $T$ . We remark though, that [Ter10] works with bounded domains.

Even with the mentioned condition these authors have not been able to show the result using the *decoupled FBSDE* theory alone. Notice that one could effectively use the theory developed in Proposition 4.6 (with the appropriate assumptions) to show that the solution  $(Y^{t,x,\varepsilon}, Z^{t,x,\varepsilon})$  (of (5.9)) is differentiable in  $x$ , but the mechanics employed does not lead to estimates on the moments of  $Z^{t,x,\varepsilon}$  or  $\nabla_x Y^{t,x,\varepsilon}$  that are independent of  $\varepsilon$ !

An alternative FBSDE representation of PDE (5.8) follows by interpreting the term  $au^\varepsilon \nabla_x u^\varepsilon$  not as a component of the BSDE's driver but as a part of the second order differential operator, i.e., as the drift of the forward diffusion (compare FBSDE (4.1)-(4.2) and PDE (4.8) along with the operator  $\mathcal{L}$ ). This leads then to a *fully coupled* FBSDE, namely take  $(t, x) \in [0, T] \times \mathbb{R}^m$  and  $s \in [t, T]$

$$\begin{aligned} X_s^{t,x,\varepsilon} &= x - \int_t^s a Y_r^{t,x,\varepsilon} dr + \sqrt{2\varepsilon} \int_t^s 1 dW_r, \\ Y_s^{t,x,\varepsilon} &= g(X_T^{t,x,\varepsilon}) - \int_s^T \lambda Y_r^{t,x,\varepsilon} dr - \int_s^T Z_s^{t,x,\varepsilon} dW_r. \end{aligned}$$

We give a quick sketch of the arguments for the *fully coupled case* since the theory differs a bit from what has been presented before. Although the argumentation is not complicated, a full justification would be lengthy (for this already long manuscript); for more on the theory of *fully coupled FBSDE* we point the reader to [MY99].

Let us assume that the above coupled FBSDE has a unique solution  $(X^{t,x,\varepsilon}, Y^{t,x,\varepsilon}, Z^{t,x,\varepsilon}) \in \mathcal{S}^p \times \mathcal{S}^\infty \times \mathcal{H}_{BMO}$ , and rather trivially (to see) that  $\|Y^{t,x,\varepsilon}\|_{\mathcal{S}^\infty}$  and  $\|Z^{t,x,\varepsilon}\|_{\mathcal{H}_{BMO}}$  are uniformly bounded in their parameters. For  $x, x' \in \mathbb{R}^m$ ,  $t \in [0, T]$  and  $\delta Y = Y^{t,x,\varepsilon} - Y^{t,x',\varepsilon}$  and  $\delta Z = Z^{t,x,\varepsilon} - Z^{t,x',\varepsilon}$ , applying Itô's formula to  $|\delta Y_s|^2$  for  $s \in [t, T]$  and noticing that

$$|g(X_T^{t,x,\varepsilon}) - g(X_T^{t,x',\varepsilon})|^2 \leq 2K^2 \left( |x - x'|^2 + a^2(T-t) \int_t^T |\delta Y_s|^2 ds \right),$$

leads to

$$\begin{aligned} & \int_t^T |\delta Z_s|^2 ds + |\delta Y_t|^2 \\ & \leq 2K^2 |x - x'|^2 + 2(K^2 a^2(T-t) - \lambda) \int_t^T |\delta Y_s|^2 ds - \int_t^T 2\delta Y_s \delta Z_s dW_s. \end{aligned}$$

Assuming that  $\lambda \geq K^2 a^2 T$  and taking  $\mathcal{F}_t$ -conditional expectation yields

$$|\delta Y_t|^2 = |u^\varepsilon(t, x) - u^\varepsilon(t, x')|^2 \leq 2K^2 |x - x'|^2.$$

Surely, this result does not imply differentiability in  $x$ , but with some extra arguments (mollification of the terminal condition then passing to the limit) one easily obtains the spatial differentiability of  $Y^{t,x,\varepsilon}$  (and hence of  $u(t, x)$ ). From the above estimate, one would obtain  $|\nabla_x u^\varepsilon| \leq \sqrt{2}K$  uniformly in  $\varepsilon$ . This again hints that the estimate would hold in the limit.



## 5.4 Incompressible Navier-Stokes

While in the previous subsection we were forced in the end to work with fully coupled FBSDE in order to obtain the extra result we were searching for, we now give an example which simplifies greatly if one uses the decoupled FBSDE instead of the fully coupled ones. In this section we look at the works by [CS09, CS10] where they present a connection between FBSDE and the incompressible Navier-Stokes equation.

More than being a standard PDE, the Navier-Stokes equation describes a conservation law (see the equations below) and it is this conservation law that is the hard part in the FBSDE representation. We saw in Section 4 that the FBSDE yields solutions to PDE but the conservation of volume condition (i.e.  $\operatorname{div}(u) = \langle \mathbf{1}, \nabla_x u \rangle = 0$ , see (5.10)) has to be verified. One does not obtain it as a side product of the existence of solution.

Following [CS09, CS10] we look at the incompressible Navier-Stokes equation with  $u : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$  satisfying for  $(t, x) \in [0, T] \times \mathbb{R}^n$

$$\begin{aligned} \partial_t u &= -u \langle \mathbf{1}, \nabla_x u \rangle + \nu \Delta u - \nabla_x p, \quad \operatorname{div} u = 0, \\ u(0, x) &= h(x), \quad \operatorname{div} h = 0, \end{aligned} \quad (5.10)$$

where  $\nu > 0$ ,  $\operatorname{div}(u) := \langle \mathbf{1}, \nabla_x u \rangle$  and  $h \in C_b^1(\mathbb{R}^n)$ . The function  $p$  denotes the pressure field which is assumed to satisfy  $\nabla_x p = K$  for some constant  $K$ . Throughout we work with  $n = 2$  but the results presented here are easily extended to the general case.

In view of the calculations carried out in Subsection 5.3 the FBSDE corresponding to the PDE (5.10) (after a time reversal) is for  $(t, x) \in [0, T] \times \mathbb{R}^2$

$$\begin{aligned} X_s^{t,x} &= x + \sqrt{2\nu}(W_s - W_t), \quad s \in [t, T] \\ Y_s^{t,x} &= h(X_T^{t,x}) - \int_s^T Z_r^{t,x} dr - \int_s^T \left[ K + \frac{1}{\sqrt{2\nu}} Y_r^{t,x} \langle \mathbf{1}, Z_r^{t,x} \rangle \right] dr, \end{aligned}$$

where  $W = (W^1, W^2)$ ,  $X = (X^1, X^2)$  and  $Z = (Z^1, Z^2)$ .

From Theorems 3.9, 3.10 and 4.10 easily follows existence and uniqueness of a solution  $(X, Y, Z)$  to the FBSDE in  $\mathcal{S}^p \times \mathcal{S}^\infty \times \mathcal{H}_{BMO}$  (for any  $p \geq 1$ ) as well as its connection to the PDE. Moreover since  $h \in C_b^1$ , Proposition 4.6 also holds, yielding the existence of  $(\nabla_x Y^{t,x}, \nabla_x Z^{t,x}) \in \mathcal{S}^p \times \mathcal{H}^p$  for  $p \geq 2$ .

One of the difficulties in [CS09] was showing that  $Y$  satisfied the free divergence condition, namely that for  $u(t, x) := Y_t^{t,x}$  the function  $u$  satisfies  $\operatorname{div}(u) = 0$  for all  $(t, x) \in [0, T] \times \mathbb{R}^2$ . For the case we present here, this is considerably easier than [CS09]. We have

$$\begin{aligned} \nabla_x Y_s^{t,x} &= (\nabla_x h)(X_T^{t,x}) \nabla_x X_T^{t,x} - \int_s^T \nabla Z_r^{t,x} dW_r \\ &\quad - \int_s^T \frac{1}{\sqrt{2\nu}} \left[ \nabla_x Y_r^{t,x} \langle \mathbf{1}, Z_r^{t,x} \rangle + Y_r^{t,x} \nabla_x (\langle \mathbf{1}, Z_r^{t,x} \rangle) \right] dr, \end{aligned}$$

where  $\nabla_x X_s = (1, 1)$ . Then,  $\operatorname{div}(u) = 0$  is found by looking at the processes  $U := \langle \mathbf{1}, \nabla_x Y^{t,x} \rangle$  and  $V^i := \langle \mathbf{1}, \nabla_x Z^{t,x,i} \rangle$  for  $i \in \{1, 2\}$ . We have then

$$U_s = (\operatorname{div}(h))(X_T^{t,x}) - \int_s^T V_r dW_r - \int_s^T \left[ \frac{\langle \mathbf{1}, Z_r^{t,x} \rangle}{\sqrt{2\nu}} U_r + \frac{Y_r^{t,x}}{\sqrt{2\nu}} \langle \mathbf{1}, V_r \rangle \right] dr.$$

Now since  $\operatorname{div}(h) = 0$  and using the linearization and measure change argument, i.e. for  $s \in [t, T]$  define  $b_s := Y_s^{t,x}/\sqrt{2\nu}$ , a probability measure  $\mathbb{Q}$  with Radon-Nikodym density  $\mathcal{E}(b * W)$  and  $e_s := \exp\{\int_t^s \langle \mathbf{1}, Z_r^{t,x} \rangle / \sqrt{2\nu} dr\}$ . Note that since  $Y \in \mathcal{S}^\infty$  and  $Z \in \mathcal{H}_{BMO}$  the processes  $b$  and  $e$  are well defined, see for instance the arguments used in the proof of Theorem 3.10.

We obtain then

$$U_s = -e_s^{-1} \int_s^T e_r V_r dW_r^{\mathbb{Q}} \Rightarrow U_s = 0 \quad \mathbb{Q}\text{-a.s. for all } s \in [t, T].$$

In particular,  $U_t = \operatorname{div}(Y_t^{t,x}) = 0$  as we sought. This proves that the solution of the FBSDE provides a mild solution to the Navier-Stokes equation (5.10).

**Acknowledgment:** We thank Peter Friz and Peter Imkeller for their helpful comments. Christoph Frei gratefully acknowledges financial support by the Natural Sciences and Engineering Research Council of Canada through grant 402585. Gonalo dos Reis was partially supported by the CMA/FCT/UNL, under the project PEst-OE/MAT/UI0297/2011.

## A Proof of Proposition 4.6

*Proof of Proposition 4.6.* The results concerning the forward diffusion are standard, see for instance Section 1.2.4 of [DR11].

This proof has 3 steps. We first establish a candidate for the said derivative process, then we prove that the candidate is the correct one and we finish by showing continuity of the derivatives in the spatial component. Throughout let  $j \in \{1, \dots, m\}$  and  $t \in [0, T]$ .

**Step 1 - The candidate for the partial derivatives:** A quick look at the BSDEs composing the lines of (4.7) shows that the lines are independent of each other. So, in view of the assumptions and the already obtained results concerning  $Y^x$  and  $Z^x$  it easily follows that for *each line* of the system (4.7) Assumptions A1–A4 of [BC08] hold. Hence we can apply their results to conclude that (4.7) has a unique solution  $(\nabla_{x_j} Y^x, \nabla_{x_j} Z^x) \in \mathcal{S}^p \times \mathcal{H}^p$  for all  $p \geq 2$  and  $j = \{1, \dots, m\}$ . These processes are the natural candidates for the  $j$ -th partial derivatives of  $Y^x$  and  $Z^x$ .

**Step 2 - The candidates are the correct ones:** We now need only to prove that the solution of (4.7) corresponds indeed to the partial derivative of

$(Y^x, Z^x)$ . To this end we need only to show that for any  $h > (0, 1]$  such that we have for all  $p \geq 2$ :

$$\lim_{h \rightarrow 0} \left\{ \left\| \frac{Y^{x+he_j} - Y^x}{h} - \nabla_{x_j} Y^x \right\|_{\mathcal{S}^p} + \left\| \frac{Z^{x+he_j} - Z^x}{h} - \nabla_{x_j} Z^x \right\|_{\mathcal{H}^p} \right\} = 0.$$

Again we do not give the details, the proof of this result follows from a straightforward combination of the mechanics employed in the proof of Corollary 3.13 and the techniques used to prove Theorem 2.1 of [AIdR07] (alternatively see Section 3 of [DR11]).

**Step 3 - Total differentiability and the continuity estimate:** After having established that the partial derivatives exist and satisfy good integrability properties, it remains only to argue in favor of total differentiability of the map  $x \mapsto Y^x$ . This will be done by showing that the partial derivatives are indeed continuous at any point  $x$  (in the topology of  $\mathcal{S}^p \times \mathcal{H}^p$ ). We do this by proving the limit result on the difference of gradients.

This estimate is easy to obtain as it follows from the techniques used throughout Section 3. We do not give the full details. Let  $t \in [0, T]$ ,  $x, x' \in \mathbb{R}^m$  then the BSDE for the difference  $\delta \nabla Y = \nabla_x Y^{t,x} - \nabla_x Y^{t,x'}$  (we define as well  $\delta \nabla Z = \nabla_x Z^{t,x} - \nabla_x Z^{t,x'}$ ) is given by

$$\delta \nabla Y_s = \delta \eta - \int_s^T \delta \nabla Z_r dW_r + \int_s^T [F_s + K_s^{\mathcal{Y}} \delta \nabla Y_r + K_s^{\mathcal{Z}} \delta \nabla Z_r] dr, \quad (\text{A.1})$$

with  $\delta \eta = (\nabla_x g)(X_T^x) \nabla_x X_T^x - (\nabla_x g)(X_T^{x'}) \nabla_x X_T^{x'}$ ,  $K^{\mathcal{Y}} = \nabla_y f(\cdot, \Theta^x)$ ,  $K^{\mathcal{Z}} = \nabla_z f(\cdot, \Theta^x)$  and  $\nabla_x \Theta^{x'} = (\nabla_x X^{x'}, \nabla_x Y^{x'}, \nabla_x Z^{x'})$  we define the process  $F := \nabla_x f(\cdot, \Theta^x) (\nabla_x X^x - \nabla_x X^{x'}) + \langle \nabla_x f(\cdot, \Theta^x) - \nabla_x f(\cdot, \Theta^{x'}), \nabla_x \Theta^{x'} \rangle$ .

The results we have seen so far easily imply (see Theorem 4.3) that  $K^{\mathcal{Y}}$  and  $K^{\mathcal{Z}}$  belong to  $\mathcal{H}_{BMO}$  with norm uniformly bounded in  $x$ . From this structure it is easy to see that the technique used to prove Corollary 3.13 can be applied to this situation, yielding for a  $\beta > 1$  related to the BMO norms of the relevant processes and a constant  $C$  independent of  $x, x'$  that

$$\|\delta \nabla Y\|_{\mathcal{S}^p}^p \leq C \mathbb{E} \left[ |\delta \eta|^{p\beta} + \left( \int_0^T |F_s| ds \right)^{p\beta} \right]^{\frac{1}{\beta}}.$$

We point out that although (A.1) is a multidimensional equation and our trick consists of a measure change, the measure change is uniform across all entries of  $\nabla Y$ , namely for each entry it is always the process  $K^{\mathcal{Z}} * W$  that is at the origin of the measure change.

Given the already proved (see Corollary 4.4) continuity of  $X^x, Y^x$  and  $Z^x$  in  $x$ , the continuity of the derivatives of  $f$  (by assumption), one can easily conclude via the dominated convergence theorem the sought result. This type of arguments have been used in Section 3.1 of [DR11] or [IDR10].  $\square$

## B Proof of Theorem 4.10

*Proof of Theorem 4.10.* Let  $(t, x) \in [0, T] \times \mathbb{R}^m$ . We have seen in previous results that the process  $Y^{t,x}$  has continuous path as well as a continuous dependency in  $x$  (see Corollary 4.4). Uniform boundedness of  $u$  follows from that of  $Y$  implied by Proposition 4.3.

To prove that  $u$  is a viscosity solution we need to prove that it is a viscosity sub- and supersolution. We prove only that it is a subsolution since the supersolution part of the proof is very similar.

To prove that  $u$  is a viscosity subsolution, let  $\varphi \in C^{1,2}([0, T] \times \mathbb{R}^m, \mathbb{R}^m)$  and take  $(\bar{t}, \bar{x}) \in [0, T] \times \mathbb{R}^m$  such that  $(\bar{t}, \bar{x})$  is a local maximum of  $(u - \varphi)$ . We assume that without loss of generality that they touch, i.e. that  $u(\bar{t}, \bar{x}) = \varphi(\bar{t}, \bar{x})$ . Suppose that<sup>14</sup>

$$\partial_t \varphi(\bar{t}, \bar{x}) + \mathcal{L}\varphi(\bar{t}, \bar{x}) + f(\bar{t}, \bar{x}, u(\bar{t}, \bar{x}), (\nabla \varphi \sigma)(\bar{t}, \bar{x})) < 0,$$

and we now argue by contradiction.

Since  $[0, T] \times \mathbb{R}^m$  is a separable space we can define a neighborhood of  $(\bar{t}, \bar{x})$  in the following way: let  $0 < \alpha \leq T - \bar{t}$  be such that for all  $\bar{t} \leq s \leq \bar{t} + \alpha$ ,  $|y - \bar{x}| \leq \alpha$ ,  $u(s, y) \leq \varphi(s, y)$  and

$$\partial_t \varphi(s, y) + \mathcal{L}\varphi(s, y) + f(s, y, u(s, y), (\nabla \varphi \sigma)(s, y)) < 0, \quad (\text{B.1})$$

and define

$$\tau = \inf\{s \geq \bar{t} : |X_s^{\bar{t}, \bar{x}} - \bar{x}| \geq \alpha\} \wedge (\bar{t} + \alpha).$$

The pair

$$(\tilde{Y}_s, \tilde{Z}_s) := (Y_{s \wedge \tau}^{\bar{t}, \bar{x}}, 1_{[\bar{t}, \tau]}(s) Z_{s \wedge \tau}^{\bar{t}, \bar{x}}), \quad s \in [\bar{t}, \bar{t} + \alpha]$$

clearly solves the BSDE on  $[\bar{t}, \bar{t} + \alpha]$

$$\tilde{Y}_s = u(\tau, X_\tau^{\bar{t}, \bar{x}}) + \int_s^{\bar{t} + \alpha} 1_{[\bar{t}, \tau]}(r) f(r, X_r^{\bar{t}, \bar{x}}, u(r, X_r^{\bar{t}, \bar{x}}), \tilde{Z}_r) dr - \int_s^{\bar{t} + \alpha} \tilde{Z}_r dW_r.$$

On the other hand, since  $\varphi \in C^{1,2}$  we have by Itô's formula that

$$(\hat{Y}_s, \hat{Z}_s) := (\varphi(s, X_{s \wedge \tau}^{\bar{t}, \bar{x}}, 1_{[\bar{t}, \tau]}(s) (\nabla \varphi \sigma)(s, X_{s \wedge \tau}^{\bar{t}, \bar{x}})), \quad s \in [\bar{t}, \bar{t} + \alpha]$$

solves the BSDE for  $s \in [\bar{t}, \bar{t} + \alpha]$

$$\hat{Y}_s = \varphi(\tau, X_\tau^{\bar{t}, \bar{x}}) - \int_s^{\bar{t} + \alpha} 1_{[\bar{t}, \tau]}(r) (\partial_t \varphi + \mathcal{L}\varphi)(r, X_r^{\bar{t}, \bar{x}}) dr - \int_s^{\bar{t} + \alpha} \hat{Z}_r dW_r.$$

<sup>14</sup>Compare with the definition of viscosity subsolution: Take  $u \in C([0, T] \times \mathbb{R}^m, \mathbb{R})$ ,  $\varphi \in C^{1,2}([0, T] \times \mathbb{R}^d)$  and  $(t, x) \in [0, T] \times \mathbb{R}^m$ . Then  $u$  is a viscosity *subsolution* of (4.8) if for all  $x$  it holds that  $u(T, x) \leq g(x)$  and if for an  $\varphi$  and a pair  $(t, x)$  which is a local maximum of  $u - \varphi$ , it holds that  $-\partial_t \varphi(t, x) - \mathcal{L}\varphi(t, x) - f(t, x, u(t, x), (\nabla \varphi \sigma)(t, x)) \leq 0$ .

At this point we need only to use a comparison result that allows us to conclude that  $\tilde{Y} < \widehat{Y}$  and hence that  $u(\bar{t}, \bar{x}) < \varphi(\bar{t}, \bar{x})$  to find the contradiction with our assumption that  $u(\bar{t}, \bar{x}) = \varphi(\bar{t}, \bar{x})$ . Unfortunately we cannot apply directly our comparison Theorem 3.12, because we do not know if the BSDE for  $(\tilde{Y}, \tilde{Z})$  satisfies its assumptions.

We start by remarking that since  $\varphi \in C^{1,2}$  and within the neighborhood around  $(\bar{t}, \bar{x})$  defined above,  $\varphi(r, X_r^{\bar{t}, \bar{x}})$  and  $(\partial_t \varphi + \mathcal{L}\varphi)(r, X_r^{\bar{t}, \bar{x}})$  are continuous functions on a compact set and hence bounded! This allows us to trivially conclude (e.g. Theorem 3.7) that  $(\tilde{Y}, \tilde{Z}) \in \mathcal{S}^\infty \times \mathcal{H}_{BMO}$  on the time interval  $[\bar{t}, \bar{t} + \alpha]$ . It is also clear that  $(\tilde{Y}, \tilde{Z})$  belongs to  $\mathcal{S}^\infty \times \mathcal{H}_{BMO}$  simply because it is the solution of (4.1), (4.2). The BSDE for the difference  $\tilde{Y} - \widehat{Y}$  is

$$\begin{aligned} \tilde{Y}_s - \widehat{Y}_s &= u(\tau, X_\tau^{\bar{t}, \bar{x}}) - \varphi(\tau, X_\tau^{\bar{t}, \bar{x}}) - \int_s^{\bar{t}+\alpha} \tilde{Z}_r - \widehat{Z}_r dW_r \\ &\quad + \int_s^{\bar{t}+\alpha} 1_{[\bar{t}, \tau]}(r) [f(r, X_r^{\bar{t}, \bar{x}}, u(r, X_r^{\bar{t}, \bar{x}}), \tilde{Z}_r) + (\partial_t \varphi + \mathcal{L}\varphi)(r, X_r^{\bar{t}, \bar{x}})] dr. \end{aligned}$$

Adding and subtracting a term  $f(r, X_r^{\bar{t}, \bar{x}}, u(r, X_r^{\bar{t}, \bar{x}}), \widehat{Z}_r)$  to the driver of the BSDE for the difference  $\tilde{Y} - \widehat{Y}$  we can apply arguments similar to those of the proof of Theorem 3.12 and conclude from the fact that  $u \leq \varphi$  and the strictness of (B.1) that  $\tilde{Y} < \widehat{Y}$ , i.e.  $u(\bar{t}, \bar{x}) < \varphi(\bar{t}, \bar{x})$ . This contradicts the assumption and hence we can conclude that  $u(t, x) = Y_t^{t,x}$  is indeed a viscosity subsolution of the PDE.

The arguments to show that  $u$  is a supersolution are similar. A combination of the two results yields that  $u$  is a viscosity solution of (4.8).1  $\square$

## References

- [AIdR07] S. Ankirchner, P. Imkeller, and G. dos Reis. Classical and variational differentiability of BSDEs with quadratic growth. *Electron. J. Probab.*, 12(53):1418–1453, 2007.
- [BC08] P. Briand and F. Confortola. BSDEs with stochastic Lipschitz condition and quadratic PDEs in Hilbert spaces. *Stochastic Process. Appl.*, 118(5):818–838, 2008.
- [BDPvH02] M. Bertsch, R. Dal Passo, and R. van der Hout. Nonuniqueness for the heat flow of harmonic maps on the disk. *Arch. Ration. Mech. Anal.*, 161(2):93–112, 2002.
- [CDY92] K.-C. Chang, W. Y. Ding, and R. Ye. Finite-time blow-up of the heat flow of harmonic maps from surfaces. *J. Differential Geom.*, 36(2):507–515, 1992.

- [CS09] A. B. Cruzeiro and E. Shamarova. Navier-Stokes equations and forward-backward SDEs on the group of diffeomorphisms of a torus. *Stochastic Process. Appl.*, 119(12):4034–4060, 2009.
- [CS10] A. B. Cruzeiro and E. Shamarova. On a forward backward stochastic system associated to the Burgers equation. *Stochastic Analysis with Financial Applications* (A. Kohatsu-Higa, N. Privault, S.J. Sheu, eds.), Progress in Probability, Vol. 65, 43-59, Birkhäuser Verlag, 2011.
- [DR07] H. Doss and S. Rainero. Sur l’existence, l’unicité, la stabilité et les propriétés de grandes déviations des solutions d’équations différentielles stochastiques rétrogrades à horizon aléatoire. Application à des problèmes de perturbations singulières. *Bull. Sci. Math.*, 131(2):99–174, 2007.
- [DR11] G. Dos Reis. *Some advances on quadratic BSDE: Theory - Numerics - Applications*. LAP LAMBERT Academic Publishing, May 2011. ISBN: 978-3-844333077.
- [DZ09] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38 of *Applications of Mathematics*. Springer-Verlag, Berlin, 2009.
- [EFKY11] N. Englezos, N. Frangos, X.-I. Kartala, and A. Yannacopoulos. Stochastic Burgers PDEs with random coefficients and a generalization of the Cole-Hopf transformation. arXiv:1106.1976, June 2011.
- [EKPQ97] N. El Karoui, S. Peng, and M. C. Quenez. Backward stochastic differential equations in finance. *Math. Finance*, 7(1):1–71, 1997.
- [EKR09] N. El Karoui and C. Ravanelli. Cash subadditive risk measures and interest rate ambiguity. *Math. Finance*, 19(4):561–590, 2009.
- [FK06] J. Feng and T. G. Kurtz. *Large deviations for stochastic processes*, volume 131 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2006.
- [FW98] M. I. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems*, volume 260. Springer-Verlag, New York, second edition, 1998.
- [IDR10] P. Imkeller and G. Dos Reis. Path regularity and explicit convergence rate for BSDE with truncated quadratic growth. *Stochastic Process. Appl.*, 120(3):348–379, 2010.
- [Kaz94] N. Kazamaki. *Continuous exponential martingales and BMO*, volume 1579 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1994.

- [Kob00] M. Kobylanski. Backward stochastic differential equations and partial differential equations with quadratic growth. *Ann. Probab.*, 28(2):558–602, 2000.
- [Mor09] M.-A. Morlais. Quadratic BSDEs driven by a continuous martingale and applications to the utility maximization problem. *Finance Stoch.*, 13(1):121–150, 2009.
- [MY99] J. Ma and J. Yong. *Forward-backward stochastic differential equations and their applications*, volume 1702 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1999. ISBN: 3-540-65960-9.
- [Par99] É. Pardoux. BSDEs, weak convergence and homogenization of semi-linear PDEs. In *Nonlinear analysis, differential equations and control (Montreal, QC, 1998)*, volume 528, pages 503–549. Kluwer Acad. Publ., Dordrecht, 1999.
- [Rai06] S. Rainero. Un principe de grandes déviations pour une équation différentielle stochastique progressive rétrograde. *C. R. Math. Acad. Sci. Paris*, 343(2):141–144, 2006.
- [Sch00] R. L. Schilling. Sobolev embedding for stochastic processes. *Expo. Math.*, 18(3):239–242, 2000.
- [SS11] R. Sircar and S. Sturm. From smile asymptotics to market risk measures. arXiv:1107.4632, July 2011.
- [Str96] M. Struwe. Geometric evolution problems. In *Nonlinear partial differential equations in differential geometry (Park City, UT, 1992)*, volume 2 of *IAS/Park City Math. Ser.*, pages 257–339. Amer. Math. Soc., Providence, RI, 1996.
- [Str08] M. Struwe. *Variational methods*, volume 34. Springer-Verlag, Berlin, fourth edition, 2008.
- [Ter10] A. S. Tersenov. On the generalized Burgers equation. *NoDEA Nonlinear Differential Equations Appl.*, 17(4):437–452, 2010.
- [Tev08] R. Tevzadze. Solvability of backward stochastic differential equations with quadratic growth. *Stochastic Process. Appl.*, 118(3):503–515, 2008.
- [Var84] S. R. S. Varadhan. *Large deviations and applications*, volume 46 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1984.