

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Geiss, Stefan; Ylinen, Juha

Title: Weighted bounded mean oscillation applied to backward stochastic differential equations

Year: 2020

Version: Accepted version (Final draft)

Copyright: © 2019 Elsevier B.V. All rights reserved.

Rights: CC BY-NC-ND 4.0

Rights url: <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Please cite the original version:

Geiss, S., & Ylinen, J. (2020). Weighted bounded mean oscillation applied to backward stochastic differential equations. *Stochastic Processes and their Applications*, 130(6), 3711-3752.

<https://doi.org/10.1016/j.spa.2019.10.007>

Journal Pre-proof

Weighted bounded mean oscillation applied to backward stochastic differential equations

Stefan Geiss, Juha Ylinen



PII: S0304-4149(19)30599-X
DOI: <https://doi.org/10.1016/j.spa.2019.10.007>
Reference: SPA 3591

To appear in: *Stochastic Processes and their Applications*

Received date: 14 February 2017
Revised date: 1 August 2019
Accepted date: 14 October 2019

Please cite this article as: S. Geiss and J. Ylinen, Weighted bounded mean oscillation applied to backward stochastic differential equations, *Stochastic Processes and their Applications* (2019), doi: <https://doi.org/10.1016/j.spa.2019.10.007>.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2019 Published by Elsevier B.V.

Weighted Bounded Mean Oscillation applied to Backward Stochastic Differential Equations

Stefan Geiss¹, Juha Ylinen^{1,2}

Department of Mathematics and Statistics, University of Jyväskylä, P.O. Box 35, FIN-40014 University of Jyväskylä, Finland.

Abstract

We deduce conditional L_p -estimates for the variation of a solution of a BSDE. Both quadratic and sub-quadratic types of BSDEs are considered, and using the theory of weighted bounded mean oscillation we deduce new tail estimates for the solution (Y, Z) on subintervals of $[0, T]$. Some new results for the decoupling technique introduced in [18] are obtained as well and some applications of the tail estimates are given.

Keywords: BSDEs, Weighted Bounded Mean Oscillation, John-Nirenberg Theorem, Tail Estimates, Decoupling

Contents

1	Introduction	2
2	Preliminaries	5
3	Weighted BMO-estimates for BSDEs	8
3.1	Non-Markovian BSDEs	8
3.2	Decoupled FBSDEs	12
4	Decoupling operators	17
4.1	Setting	17
4.2	Decoupling operators	18
4.3	Basic properties	19
4.4	Conditional results	21
5	Proof of Theorem 3.7 and Example 3.11	23
5.1	Proof of Theorem 3.7	24
5.2	Proof of Example 3.11	27

Email address: stefan.geiss@jyu.fi (Stefan Geiss)

¹This work was supported by project "Stochastic and Harmonic Analysis, interactions, and applications" of the Academy of Finland [project number 133914].

²The author was supported by the Vilho, Yrjö and Kalle Väisälä foundation of the Finnish Academy of Science and Letters.

6	Some Applications	29
6.1	Uniform spline approximation of the process Y	29
6.2	Confidence interval for direct simulation	33
6.3	Change of measure	34
6.4	Outlook	34
7	Appendix A: General tools	35
8	Appendix B: Tools related to decoupling	36
9	Appendix C: A John-Nirenberg type theorem	41

1. Introduction

In this article we study backward stochastic differential equations (BSDEs from now on) of type

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s dW_s, \quad t \in [0, T], \quad (1.1)$$

where $T > 0$ is a fixed number and $(W_t)_{t \in [0, T]}$ is a d -dimensional Brownian motion. Roughly speaking, a BSDE is a map $(\xi, f) \mapsto (Y, Z)$, so that (ξ, f) is the data, and (Y, Z) is the solution. Here the *terminal value* $\xi \in L_2$ is a given random variable that is measurable with respect to the σ -algebra generated by the Brownian motion. In the present article, the *generator* $f : [0, T] \times \Omega \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$ is assumed to be such that

- (1) $(t, \omega) \mapsto f(t, \omega, y, z)$ is predictable for all $(y, z) \in \mathbb{R} \times \mathbb{R}^d$, and
- (2) there are $L_y, L_z \geq 0$ and a $\theta \in [0, 1]$ such that for all $(t, \omega, y_0, y_1, z_0, z_1)$ one has

$$|f(t, \omega, y_0, z_0) - f(t, \omega, y_1, z_1)| \leq L_y |y_0 - y_1| + L_z [1 + |z_0| + |z_1|]^\theta |z_0 - z_1|.$$

This means that the generator f can be random, is assumed to be uniformly Lipschitz in the y -variable, and locally Lipschitz in the z -variable. We will consider the uniformly Lipschitz case ($\theta = 0$), the quadratic case ($\theta = 1$), and the sub-quadratic case ($\theta \in (0, 1)$) at the same time. We say that (Y, Z) is a solution of BSDE (1.1) if Y is a continuous adapted process with $\mathbb{E} \sup_{t \in [0, T]} |Y_t|^2 < \infty$, if Z is a predictable process with $\mathbb{E} \int_0^T |Z_t|^2 dt < \infty$, and if (1.1) is satisfied almost surely.

BSDEs were first introduced by Bismut in [4], and the amount of research increased significantly after Pardoux and Peng showed in [24] that a BSDE with square-integrable terminal value ξ and a uniformly Lipschitz generator f has a unique solution. Concerning the Lipschitz-case, see also for example [25], [15], and [7]. More recently, the theory of BSDEs with a generator that grows quadratically in the z -variable has been developed, see for example [22], [19], [23], [8], [12], and the references therein. The original motivation of studying BSDEs comes from stochastic optimal control theory. In general, BSDEs have applications in stochastic differential games, stochastic finance in connection to option pricing and utility maximization, and they are closely connected to partial differential equations (PDEs).

The present article is a continuation of [18] and an application of [17], where [18] itself is a continuation of [16]. The main results of this article are

- THEOREM 3.7, that provides conditional variational estimates for the processes (Y, Z) , i.e. we bound the *mean oscillation* of the processes (Y, Z) from above by natural *weights* derived from the initial data (ξ, f) of the BSDE,
- THEOREM 3.8, that deduces from Theorem 3.7 *conditional tail estimates of John-Nirenberg type*,
- THEOREM 3.12, that is the version of Theorem 3.7 for decoupled FBSDEs,
- THEOREMS 3.13 AND 3.14, that are versions of Theorem 3.8 for decoupled FBSDEs.

Our strategy to prove the basic Theorems 3.7 and 3.8 consists in the following steps:

- STEP 1: We prove a conditional decoupling inequality for BSDEs in Proposition 5.3.
- STEP 2: We deduce conditional variational inequalities for (Y, Z) in Theorem 3.7.
- STEP 3: We deduce conditional tail estimates for (Y, Z) in Theorem 3.8.

In Steps 1 and 2 we extend and apply methods from [18], in Step 3 we use a result from [17]. To explain the role of [17] and [18] for the present article let us assume a stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \in [0, T]})$ as in Section 2.

Relation to [17]: In [17] a class of weighted BMO spaces BMO_p^Φ has been introduced. For a positive càdlàg and adapted weight process $\Phi = (\Phi_t)_{t \in [0, T]}$ and $p \in (0, \infty)$ we say that a continuous and adapted process $A = (A_t)_{t \in [0, T]}$ with $A_0 \equiv 0$ belongs to BMO_p^Φ provided that

$$\|A\|_{\text{BMO}_p^\Phi} := \sup \left\{ \left\| \mathbb{E} \left(\left| \frac{A_T - A_\tau}{\Phi_\tau} \right|^p \middle| \mathcal{F}_\tau \right) \right\|_{L^\infty(\mathbb{P})}^{\frac{1}{p}} \mid \tau : \Omega \rightarrow [0, T] \text{ stopping time} \right\} < \infty. \quad (1.2)$$

In the present article it is essential to use the concept of BMO_p^Φ *locally in time*. To explain this let us look at Theorem 3.7 where we have weight processes $(w_{p,s,u,t})_{u \in [s,t]}$ and $(w_{p,s,u,t}^{\xi,f})_{u \in [s,t]}$ for fixed $0 \leq s < t \leq T$. If we consider the special case $[s, t] = [0, T]$ and set $\Phi_u := w_{p,0,u,T} \vee \varepsilon$ and $\Phi'_u := w_{p,0,u,T}^{\xi,f} \vee \varepsilon$ for any $\varepsilon > 0$ (the parameter $\varepsilon > 0$ is only formal, to get the weights strictly positive to be in accordance with [17]), then part of Theorem 3.7 reads as

$$\|(Y_t - Y_0)_{t \in [0, T]}\|_{\text{BMO}_p^\Phi} \leq c_{(3.7)}, \quad (1.3)$$

$$\left\| \left(\int_0^t |Z_s|^2 ds \right)_{t \in [0, T]} \right\|_{\text{BMO}_{p/2}^{\Phi'^2}} \leq d_{(3.7)}^2. \quad (1.4)$$

However, this "global" setting, i.e. $[s, t] = [0, T]$, would not give us estimates on the distribution of $Y_t - Y_s$ that take the size of $t - s$ into the account. Therefore Theorem 3.7 provides local versions of (1.3) and (1.4) in the following sense: for an arbitrary sub-interval $[s, t] \subseteq [0, T]$ we show that, for any stopping time $\tau : \Omega \rightarrow [s, t]$,

$$\left(\mathbb{E}^{\mathcal{F}_\tau} |Y_t - Y_\tau|^p \right)^{\frac{1}{p}} \leq c_{(3.7)} w_{p,s,\tau,t}, \quad (1.5)$$

$$\left(\mathbb{E}^{\mathcal{F}_\tau} \left(\int_\tau^t |Z_r|^2 dr \right)^{\frac{p}{2}} \right)^{\frac{1}{p}} \leq d_{(3.7)} w_{p,s,\tau,t}^{\xi,f}. \quad (1.6)$$

The "main" weight process $(w_{p,s,u,t}^{\xi,f})_{u \in [s,t]}$ is obtained in Assumption 3.5. Because our approach requires a localization in $[s, t]$ the spaces BMO_p^Φ could not be used in the form they have been defined in [17] and we extracted the results from [17] in a form needed in Section 9. This made it possible to perform Step 2, i.e. to deduce Theorem 3.8 from Theorem 3.7.

The Assumption 3.5 measures the sensitivity of the initial data (ξ, f) of our BSDE with respect to a class of conditional expectations in a natural way that might be interpreted as a property related to directional Malliavin derivatives. But to prove Theorem 3.7 we have to translate Assumption 3.5 into the decoupling context and obtain the equivalent condition Assumption 5.1. So Assumption 3.5 and Theorem 3.7 combine [18] and [17]: the weights originate from the decoupling techniques in [18] and are used in a context that localizes the BMO spaces from [17]. It should be mentioned that Assumption 3.5 might be seen also from the point of view that we start with the initial data (ξ, f) of the BSDE and *then* look for good or even best possible weight processes $(w_{p,s,u,t}^{\xi})_{u \in [s,t]}$ and $(w_{p,s,u,t}^f)_{u \in [s,t]}$.

Let us explain the importance of the localized approach, i.e. to consider subintervals $[s, t] \subseteq [0, T]$, by the example of decoupled Forward Backward SDEs (FBSDEs) treated in Section 3.2. There we consider

$$\begin{aligned} X_t &= x + \int_0^t b(r, X_r) dr + \int_0^t \sigma(r, X_r) dW_r, \\ Y_t &= g(X_T) + \int_t^T h(r, X_r, Y_r, Z_r) dr - \int_t^T Z_r dW_r, \end{aligned}$$

for $t \in [0, T]$, where $x \in \mathbb{R}^d$ is fixed and the main assumption is that the functions b, σ, g, h are uniformly Lipschitz in the state variables (see Assumption 3.9 below). A consequence of Theorem 3.13 is, that there exists an absolute constant $c_0 > 0$ and constants $c, C > 0$, depending on the parameters of the FBSDE, such that for any $0 \leq s < t \leq T$ we have

$$\mathbb{P} \left(\sup_{u \in [s,t]} \frac{|Y_u - Y_t|}{\sqrt{t-s}} > c\mu\nu \mid \mathcal{F}_s \right) \leq e^{1-\mu} + c_0 \mathbb{P} \left(\sup_{u \in [s,t]} |X_u|^2 > \nu^2 - 1 \mid \mathcal{F}_s \right) \quad (1.7)$$

for all $\mu, \nu > 0$. In the case that σ is bounded, this improves to

$$\mathbb{P} \left(\sup_{u \in [s,t]} \frac{|Y_u - Y_t|}{\sqrt{t-s}} > c\mu\nu \mid \mathcal{F}_s \right) \leq e^{1-\mu} + c_0 \mathbb{P} \left(\sup_{u \in [s,t]} |X_u|^2 (t-u) > \nu^2 - 1 \mid \mathcal{F}_s \right). \quad (1.8)$$

Similar results are obtained for the process Z of the FBSDE in Theorem 3.14, and for the solution (Y, Z) to the general non-Markovian BSDE in Theorem 3.8. The idea behind the inequalities (1.7) and (1.8) is to minimize for a given $\lambda > 0$ the right hand sides over all decompositions $\lambda = \mu\nu$. This is used in Sections 6.1 and 6.2. Even though (1.7) and (1.8) concern a well-studied family of FBSDEs, the tail estimates we obtain in (1.7) and (1.8) (Theorems 3.13 and 3.14) seem to be new. Coming back to moment estimates there is another application that shows how the conditional approach can be used. Let $s \in [0, T]$ and $n_s \geq 1$ such that $s + \frac{1}{n} \leq T$ for $n \geq n_s$. Then Fatou's Lemma, the right-hand side continuity of the filtration, and Theorem 3.12 for $p = 2$ give

$$\begin{aligned} \liminf_{n \rightarrow \infty, n \geq n_s} \left(n \int_s^{s+\frac{1}{n}} |Z_r|^2 dr \right) &= \mathbb{E} \left(\liminf_{n \rightarrow \infty, n \geq n_s} n \int_s^{s+\frac{1}{n}} |Z_r|^2 dr \mid \mathcal{F}_s \right) \\ &\leq \liminf_{n \rightarrow \infty, n \geq n_s} \mathbb{E} \left(n \int_s^{s+\frac{1}{n}} |Z_r|^2 dr \mid \mathcal{F}_s \right) \\ &\leq \begin{cases} C_{(3.12)}^2 [1 + |X_s|^2] & \text{under condition } (A_{b,\sigma}) \\ D_{(3.12)}^2 & \text{under conditions } (A_{b,\sigma}) \text{ and } (A_\sigma), \end{cases} \end{aligned}$$

cf. [27, 11]. The above idea can also be applied directly to Theorem 3.7 instead of to Theorem 3.12.

Relation to [18]: To prove our basic Theorem 3.7 we use the decoupling technique from [18]. This technique has to be extended into two directions:

- Similarly as the concept of the expected value is extended to the conditional expectation, some results from [18] has to be extended to a conditional context, see Section 4.4 and Proposition 5.3.
- The above mentioned Assumption 3.5 we need to translate into Assumption 5.1 to apply the decoupling technique. This translation is based on Proposition 8.6. Having in mind that every separable Banach space can be isometrically embedded into $C([0, 1])$ by the Banach-Mazur Theorem and that $\mathbb{X} = [0, 1]$ is locally σ -compact, then Proposition 8.6 is also a statement about the conditional decoupling of random variables with values in separable Banach spaces. Therefore, Proposition 8.6 is an infinite-dimensional version of Lemma 4.6, where Lemma 4.6 is a conditional version of [18, Lemma 4.20].

The article is organized as follows. The main results are formulated in Section 3. We also include proofs in Section 3 as long as the decoupling technique from [18] is not required. This technique is introduced in Section 4. In Section 5 we complete the proofs of the results in Section 3 with the methods from Section 4. Some applications of the estimates we obtained are illustrated in Section 6. The Appendices A,B, and C contain some technical tools that were needed throughout the article.

2. Preliminaries

A constant with a subindex of the form (3.12) is a constant from the result that is numbered 3.12. For example, $c_{(3.12)}$, $d_{(3.12)}$, $C_{(3.12)}$ and $D_{(3.12)}$ are constants from Theorem 3.12. We fix a finite number $T > 0$ and work on the stochastic basis

$$(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_r)_{r \in [0, T]})$$

satisfying *the usual assumptions*. In particular, $(\Omega, \mathcal{F}, \mathbb{P})$ is complete and in our case $\mathbb{F} := (\mathcal{F}_r)_{r \in [0, T]}$ is the augmented filtration of a d -dimensional Brownian motion W , $\mathcal{F} = \mathcal{F}_T$, and we assume that all paths of W are continuous. If we give a statement or a definition that involves a filtration, but the filtration is not mentioned explicitly, then \mathbb{F} is used. Moreover, the following notation will be used:

2.1 Notation.

- (1) The Lebesgue-measure on $[0, T]$ is denoted by λ , and

$$\begin{aligned} (\Omega_0, \Sigma_0, \mathbb{P}_0) &:= (\Omega, \mathcal{F}, \mathbb{P}), \\ (\Omega_T, \Sigma_T, \mathbb{P}_T) &:= \left([0, T] \times \Omega, \mathcal{B}([0, T]) \otimes \mathcal{F}, \frac{\lambda}{T} \otimes \mathbb{P} \right). \end{aligned}$$

- (2) Given a σ -algebra $\mathcal{G} \subseteq \mathcal{F}$ and $X \in L_1(\Omega, \mathcal{F}, \mathbb{P})$, the conditional expectation of X given \mathcal{G} is denoted by $\mathbb{E}^{\mathcal{G}}X := \mathbb{E}[X | \mathcal{G}]$.

- (3) For any $B \in \mathcal{F}$ of positive measure and any $A \in \mathcal{F}$ we let

$$\mathbb{P}_B(A) := \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(B)}.$$

(4) For $0 \leq s < t \leq T$ we let

$$\mathcal{G}_s^t := \sigma(W_r, r \leq s) \vee \sigma(W_r - W_t, t < r \leq T) \quad \text{and} \quad \mathcal{H}_s^t := \mathcal{B}([0, T]) \otimes \mathcal{G}_s^t. \quad (2.1)$$

(5) The (predictable) σ -algebra on Ω_T generated by $(\mathcal{F}_t)_{t \in [0, T]}$ -adapted left-continuous processes is denoted by \mathcal{P} .

In general, inequalities concerning random variables, for example $\mathbb{E}^{\mathcal{G}} X \leq cY$, where $c > 0$ is a constant, hold only almost surely. If it is obvious what measure is used, we will just write $\mathbb{E}^{\mathcal{G}} X \leq cY$. If A is a subset of a metric space, then we denote the interior of A by $\overset{\circ}{A}$ and the closure of A by \bar{A} .

2.2 Definition. A complete metric space $\mathbb{X} \neq \emptyset$ is locally σ -compact, if there exist compact subsets $\emptyset \neq K_1 \subseteq K_2 \subseteq \dots$, such that $\overset{\circ}{K}_n = K_n$ and $\mathbb{X} = \bigcup_{n=1}^{\infty} \overset{\circ}{K}_n$.

2.3 Proposition. A locally σ -compact \mathbb{X} is separable. Moreover, if $(K_n)_{n \geq 1}$ are compact subsets as in Definition 2.2, and $\mathbb{A} \subseteq \mathbb{X}$ is a dense countable subset, then for any $n \geq 1$ the set $\mathbb{A}_n := \mathbb{A} \cap K_n$ is dense in K_n .

2.4 Definition. For $S \in \{0, T\}$ we use

$$L_0(\Omega_S, \Sigma_S, \mathbb{P}_S; C(\mathbb{X}))$$

to denote the equivalence-classes³ of $f : \Omega_S \times \mathbb{X} \rightarrow \mathbb{R}$ that satisfy:

- (1) $\eta \mapsto f(\eta, y)$ is Σ_S -measurable for all $y \in \mathbb{X}$,
- (2) $y \mapsto f(\eta, y)$ is continuous for all $\eta \in \Omega_S$.

We will need the Burkholder-Davis-Gundy-inequalities:

2.5 Proposition ([26, p.160], [3, Proposition 4.2]). Let $p \in (0, \infty)$. Then there exists $\alpha_p, \beta_p > 0$ such that for all (continuous) martingales $(M_t)_{t \in [0, T]}$ with $M_0 \equiv 0$ we have:

$$\alpha_p \left\| \langle M \rangle_t^{\frac{1}{2}} \right\|_{L_p} \leq \left\| \sup_{s \in [0, t]} |M_s| \right\|_{L_p} \leq \beta_p \left\| \langle M \rangle_t^{\frac{1}{2}} \right\|_{L_p}$$

for all $t \in [0, T]$, where $(\langle M \rangle_t)_{t \in [0, T]}$ is the quadratic variation process of M . For $p \in [2, \infty)$ the constant $\beta_p > 0$ can be chosen such that $\beta_p \leq c \sqrt{p}$ for some absolute $c > 0$.

Next we introduce the sliceable numbers.

2.6 Definition (cf. [18, Definition 5.2]). Assume that $(c_r)_{r \in [0, T]}$ is predictable, d -dimensional and such that

$$\|c\|_{\text{BMO}(S_2)} := \sup_{t \in [0, T]} \left\| \mathbb{E} \left(\int_t^T |c_s|^2 ds \middle| \mathcal{F}_t \right) \right\|_{L_{\infty}}^{\frac{1}{2}} < \infty.$$

Then we say $c \in \text{BMO}(S_2)$. This is quantified using, for any $N \geq 1$, $\text{sl}_N(c) := \inf \varepsilon$, where the infimum is taken over all $\varepsilon > 0$ such that there are stopping times $0 = \tau_0 \leq \tau_1 \leq \dots \leq \tau_N = T$ with

$$\sup_{k=1, \dots, N} \|\chi_{(\tau_{k-1}, \tau_k]} c\|_{\text{BMO}(S_2)} \leq \varepsilon.$$

Moreover, we let $\text{sl}_{\infty}(c) := \lim_{N \rightarrow \infty} \text{sl}_N(c)$.

³We identify f and g if $f(\eta, \cdot) = g(\eta, \cdot)$ for \mathbb{P}_S -a.e. $\eta \in \Omega_S$.

For our main application of sliceable numbers we introduce the function

$$\Phi : (1, \infty) \rightarrow (0, \infty), \quad \Phi(q) = \left(1 + \frac{1}{q^2} \log \left(1 + \frac{1}{2q-2}\right)\right)^{1/2} - 1, \quad (2.2)$$

so that Φ is continuous and decreasing, with $\lim_{q \rightarrow \infty} \Phi(q) = 0$ and $\lim_{q \rightarrow 1} \Phi(q) = \infty$. Furthermore, we let

$$\Psi : \{(\gamma, q) \in [0, \infty) \times (1, \infty) : 0 \leq \gamma < \Phi(q) < \infty\} \rightarrow [0, \infty),$$

$$\Psi(\gamma, q) := \left(\frac{2}{1 - \frac{2q-2}{2q-1} e^{q^2[\gamma^2+2\gamma]}}\right)^{\frac{1}{q}}.$$

The concept of sliceable numbers is motivated by Proposition 2.8 below. To formulate this statement we need the following definition:

2.7 Definition. Let $M = (M_t)_{t \in [0, T]}$ be a martingale with $M_0 \equiv 0$ such that $\mathcal{E}(M) = (\mathcal{E}(M)_t)_{t \in [0, T]} := (e^{M_t - \frac{1}{2}\langle M \rangle_t})_{t \in [0, T]}$ is a martingale as well, and let $q \in (1, \infty)$. If

$$\mathcal{RH}_q(\mathcal{E}(M))^q := \sup_{\tau} \left\| \mathbb{E}^{\mathcal{F}_{\tau}} \left[\frac{\mathcal{E}(M)_T}{\mathcal{E}(M)_{\tau}} \right]^q \right\|_{\infty} < \infty,$$

where the supremum is taken over all stopping times $\tau : \Omega \rightarrow [0, T]$, we say⁴ $\mathcal{E}(M) \in \mathcal{RH}_q$.

2.8 Proposition ([18, Theorem 5.25]). Assume that $c \in \text{BMO}(S_2)$ is d -dimensional, and that for some $N \geq 1$ it holds that $\text{sl}_N(c) < \Phi(q)$. Then, putting $(M_t)_{t \in [0, T]} := (\int_0^t c_r dW_r)_{t \in [0, T]}$, we have

$$\mathcal{RH}_q(\mathcal{E}(M)) \leq [\Psi(\text{sl}_N(c), q)]^N.$$

In particular, if M is sliceable, i.e. $\text{sl}_{\infty}(c) = 0$, then for all $q > 1$ there exists an $N \geq 1$ such that $\text{sl}_N(c) < \Phi(q)$, so that $\mathcal{E}(M) \in \bigcap_{q \in (1, \infty)} \mathcal{RH}_q$.

We end with an extension of Fefferman's inequality, which was proven in [18, Corollary 5.19] (see also [13, Lemma 1.6] and [2, Theorem 1.1(iii)]). Note that here both X and Y may be multidimensional.

2.9 Proposition. Assume that $X \in \text{BMO}(S_2)$ and that $Y = (Y_r)_{r \in [0, T]}$ is predictable and such that

$$\|Y\|_{H_p(S_2)}^p := \mathbb{E} \left(\int_0^T |Y_r|^2 dr \right)^{p/2} < \infty$$

for some $p \in [1, \infty)$. Then $\left\| \int_0^T |X_r| |Y_r| dr \right\|_{L_p} \leq \sqrt{2p} \|Y\|_{H_p(S_2)} \|X\|_{\text{BMO}(S_2)}$.

In this article we deduce conditional estimates on subintervals $[s, t] \subseteq [0, T]$, and for this we need the following conditional version of Proposition 2.9:

2.10 Proposition. Assume that $X \in \text{BMO}(S_2)$ and that $Y = (Y_r)_{r \in [0, T]}$ is predictable and such that $\|Y\|_{H_p(S_2)} < \infty$ for some $p \in [1, \infty)$, and let $c_p = (\sqrt{2p})^p$. Then we have for all $0 \leq s < t \leq T$ that

$$\mathbb{E}^{\mathcal{F}_s} \left(\int_s^t |X_r| |Y_r| dr \right)^p \leq c_p \left(\mathbb{E}^{\mathcal{F}_s} \left(\int_s^t |Y_r|^2 dr \right)^{\frac{p}{2}} \right) \sup_{r \in [s, t]} \left\| \mathbb{E}^{\mathcal{F}_r} \int_r^t |X_u|^2 du \right\|_{\infty}^{\frac{p}{2}}.$$

⁴ \mathcal{RH} stands for Reverse Hölder.

3. Weighted BMO-estimates for BSDEs

First we present our results in the general non-Markovian context in Section 3.1. Then the results are illustrated for decoupled FBSDEs in Section 3.2 where we also discuss their sharpness in Examples 3.15 and 3.16.

3.1. Non-Markovian BSDEs

We consider BSDEs of type

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s dW_s, \quad t \in [0, T], \quad (3.1)$$

where ξ is \mathcal{F}_T -measurable, and $f \in \mathcal{L}_0(\Omega_T, \mathcal{P}, \mathbb{P}_T; C(\mathbb{R}^{d+1}))^5$. Our strategy is to assume that (Y, Z) is a solution of (3.1), and assume some further conditions on Z in order to get an L_p -solution for $p \in [2, \infty)$. In Example 3.4 we present some cases when these conditions are satisfied. For $p \in [2, \infty)$ and $\theta \in [0, 1]$, we consider the conditions:

(C1) There are $L_y, L_z \geq 0$ such that for all $(t, \omega, y_0, y_1, z_0, z_1)$ one has

$$|f(t, \omega, y_0, z_0) - f(t, \omega, y_1, z_1)| \leq L_y |y_0 - y_1| + L_z [1 + |z_0| + |z_1|]^\theta |z_0 - z_1|.$$

(C2) $\int_0^T |f(s, 0, 0)| ds \in \mathcal{L}_p$.

(C3) $\left(\int_0^T |Z_s|^2 ds \right)^{\frac{1}{2}} \in \mathcal{L}_p$.

(C3') $\int_0^T |Z_s|^{1+\theta} ds \in \mathcal{L}_p$.

Assumptions (C1) and (C2) are conditions on the data of the BSDE, implicit conditions on the Z -process are (C3) and (C3').

3.1 Lemma ([18, Lemma 6.2]). *Assume that (C1)-(C3) and (C3') hold for some $p \in [2, \infty)$ and $\theta \in [0, 1]$. Then*

$$\int_0^T |f(s, Y_s, Z_s)| ds + \sup_{t \in [0, T]} |Y_t| \in \mathcal{L}_p.$$

Another implicit condition is the following "fractional BMO-assumption":

(C4) We assume that

$$\| |Z|^\theta \|_{\text{BMO}(S_2)} = \sup_{t \in [0, T]} \left\| \mathbb{E} \left(\int_t^T |Z_s|^{2\theta} ds \middle| \mathcal{F}_t \right) \right\|_\infty^{\frac{1}{2}} < \infty,$$

and fix a non-increasing sequence $s = (s_N)_{N \geq 1} \subseteq [0, \infty)$ such that

$$s_N(|Z|^\theta) \leq s_N,$$

⁵This means that $\eta \mapsto f(\eta, x)$ is \mathcal{P} -measurable for all $x \in \mathbb{R}^{d+1}$.

and put $s_\infty := \lim_{N \rightarrow \infty} s_N$. If $s_\infty = 0$, then we let $p_{(C4)} = \frac{3}{2}$, and if $s_\infty > 0$, then we let

$$p_{(C4)} := \frac{\Phi^{-1}(2\sqrt{2}L_z s_\infty)}{\Phi^{-1}(2\sqrt{2}L_z s_\infty) - 1},$$

where the function Φ is defined in (2.2).

First we show that using (C4) we may drop the assumption (C3'):

3.2 Lemma. *For all $p \in [2, \infty)$ we have the following relations:*

- (i) *If $\theta = 0$, then (C4) holds, and (C3) \Rightarrow (C3').*
- (ii) *If $\theta = 1$, then (C4) \Rightarrow (C3') \Rightarrow (C3).*
- (iii) *If $\theta \in (0, 1)$ and (C4) holds, then (C3) \Rightarrow (C3').*

Proof. (i) is obvious and (ii) follows immediately from John-Nirenberg inequality [21, Theorem 2.1]. Proposition 2.9 applied to $X = |Z|^\theta$ and $Y = |Z|$ implies (iii). \square

3.3 Remark. In addition to Lemma 3.2, the condition (C4) has an even more important role that we describe now. In our results, conditions (C4) and (C1) are assumed to hold for the same $\theta \in [0, 1]$. Then, applying Proposition 2.8, we have that a certain martingale satisfies the reverse Hölder inequality. This martingale is used to handle the quadratic or sub-quadratic nature of the generator f in the z -variable. If the number s_∞ in (C4) equals zero, then the reverse Hölder inequalities are satisfied for all indices $q \in (1, \infty)$. On the other hand, if $s_\infty > 0$, then there exists $q_0 \in (1, \infty)$ such that the reverse Hölder inequalities are satisfied for all $q \in (1, q_0)$. From this it follows that in the case $s_\infty > 0$ we need to assume more integrability than in the case $s_\infty = 0$, and this is the reason for introducing the constant $p_{(C4)}$. Note that in the uniformly Lipschitz case, i.e. $\theta = 0$, the condition (C4) is satisfied and $s_\infty = 0$. In the sub-quadratic case, i.e. $\theta \in (0, 1)$, a sufficient condition for $s_\infty = 0$ is, that there exists an $\eta \in (\theta, 1]$ such that $\| |Z|^\eta \|_{\text{BMO}(S_2)} < \infty$ (see [18, Remark 6.4]).

3.4 Example.

- (i) Assume that f satisfies (C1) and (C2) with $\theta = 0$ and $p > 1$, and that $\xi \in L_p$. Then there exists a unique solution (Y, Z) of (3.1), and (C3)-(C4) are satisfied with $\theta = 0$. This follows for example from [7, Theorem 4.2]. Note that since $\theta = 0$, we have $s_\infty = 0$.
- (ii) Assume that f satisfies (C1) and (C2) with $\theta = 1$ and $p = \infty$, and that $\xi \in L_\infty$. Then there exists a solution (Y, Z) of (3.1) such that (C3)-(C4) are satisfied with $\theta = 1$ and all $p \in [2, \infty)$. This follows for example from [23, Theorem 2.6 and Lemma 3.1] (see also [9]).
- (iii) Assume that f satisfies (C1) with $\theta_0 \in (0, 1)$, and is such that $\sup_{(r,\omega)} |f(r, \omega, 0, 0)| < \infty$. Also, assume that $\xi \in \text{cExp}$, which means that there exists some $\mu \in (0, \infty)$ such that

$$\sup_{t \in [0, T)} (T - t) \left\| \mathbb{E} \left[e^{\mu |\xi|} \middle| \mathcal{F}_t \right] \right\|_\infty < \infty.$$

Then there exists a solution (Y, Z) of (3.1) such that (C3)-(C4) are satisfied with $p = 2$ and all $\theta \in (0, 1)$, so that $s_\infty = 0$ (see [18, Theorem 6.13]).

Our final assumption is a weighted BMO-condition on ξ and f on a subinterval $[s, t] \subseteq [0, T]$. This is used in the following way: if (C1)-(C4) are satisfied and Assumption 3.5 holds on an interval $[s, t]$, then *on this interval* we have a weighted BMO-estimate and a tail estimate of (Y, Z) .

3.5 Assumption. Let $p \in [2, \infty)$ and $0 \leq s < t \leq T$. There are non-negative càdlàg processes $(w_{p,s,u,t}^\xi)_{u \in [s,t]}$ and $(w_{p,s,u,t}^f)_{u \in [s,t]}$ such that $((w_{p,s,u,t}^\xi)^p)_{u \in [s,t]}$ and $((w_{p,s,u,t}^f)^p)_{u \in [s,t]}$ are supermartingales and which satisfy, for any $u \in [s, t]$,

$$(C5) \quad \left(\mathbb{E}^{\mathcal{F}_u} |\xi - \mathbb{E}^{\mathcal{G}_u} \xi|^p \right)^{\frac{1}{p}} \leq w_{p,s,u,t}^\xi,$$

$$(C6) \quad \left(\mathbb{E}^{\mathcal{F}_u} \left(\int_u^T \sup_{y,z} |f(r, y, z) - (\mathbb{E}^{\mathcal{H}_u} f)(r, y, z)| dr \right)^p \right)^{\frac{1}{p}} \leq w_{p,s,u,t}^f,$$

where \mathcal{G}_u^t and \mathcal{H}_u^t are given in (2.1) and $(\mathbb{E}^{\mathcal{H}_u^t} f) : \Omega_T \rightarrow C(\mathbb{R} \times \mathbb{R}^d)$ is the⁶ \mathcal{H}_u^t -measurable process with

$$\mathbb{P}_T \left(\mathbb{E}^{\mathcal{H}_u^t} (f(x)) = (\mathbb{E}^{\mathcal{H}_u^t} f)(x) \right) = 1 \text{ for all } x \in \mathbb{R} \times \mathbb{R}^d.$$

To shorten the notation, we use

$$w_{p,s,u,t}^{\xi,f} := \left((w_{p,s,u,t}^\xi)^p + (w_{p,s,u,t}^f)^p \right)^{\frac{1}{p}}.$$

3.6 Remark. For a fixed $u \in (s, t]$, the weight $w_{p,s,u,t}^\xi$ is an upper bound for $\left(\mathbb{E}^{\mathcal{F}_u} |\xi - \mathbb{E}^{\mathcal{G}_u} \xi|^p \right)^{\frac{1}{p}}$, so we expect $w_{p,s,u,t}^\xi$ to depend on u and t , but not on s . We use a notation where the s is included, since we want to emphasize the fact that Assumption 3.5 is an assumption on the behaviour of (ξ, f) on the interval $[s, t]$.

We are ready to give our main result.

3.7 Theorem. Assume (C1)-(C6) for $\theta \in [0, 1]$, $p \in [2, \infty) \cap (p_{(C4)}, \infty)$, and $0 \leq s < t \leq T$. Then the following assertions hold true:

- (i) There exists $c_{(3.7)} > 0$ depending at most on $(T, d, p, L_y, L_z, (s_N)_{N \in \mathbb{N}})$ such that for any stopping time $\tau : \Omega \rightarrow [s, t]$ we have

$$\left(\mathbb{E}^{\mathcal{F}_\tau} |Y_t - Y_\tau|^p \right)^{\frac{1}{p}} \leq c_{(3.7)} w_{p,s,\tau,t}, \quad (3.2)$$

where

$$w_{p,s,u,t}^p = \left(w_{p,s,u,t}^{\xi,f} \right)^p + \mathbb{E}^{\mathcal{F}_u} \left(\int_u^t |f(r, 0, 0)| dr \right)^p + (t-u)^p \left[\mathbb{E}^{\mathcal{F}_u} \left(|\xi| + \int_t^T |f(r, 0, 0)| dr \right)^p \right].$$

- (ii) There exists $d_{(3.7)} > 0$ depending at most on $(T, d, p, L_y, L_z, (s_N)_{N \in \mathbb{N}})$ such that for any stopping time $\tau : \Omega \rightarrow [s, t]$ we have

$$\left(\mathbb{E}^{\mathcal{F}_\tau} \left(\int_\tau^t |Z_r|^2 dr \right)^{\frac{p}{2}} \right)^{\frac{1}{p}} \leq d_{(3.7)} w_{p,s,\tau,t}^{\xi,f}.$$

⁶Existence and uniqueness of such a process is proven in Lemma 7.1 below.

Theorem 3.7 is proved in Section 5.1 below. The main application of Theorem 3.7 are the following tail estimates:

3.8 Theorem. *Under the assumptions of Theorem 3.7 there exists an absolute constant $c_0 > 0$ such that for any stopping time $\tau : \Omega \rightarrow [s, t]$ we have*

$$\begin{aligned} \text{(i)} \quad & \mathbb{P}_B \left(\sup_{u \in [\tau, t]} \frac{|Y_u - Y_\tau|}{c(3.7)} > \lambda + c_0 \mu \nu \right) \leq e^{1-\mu} \mathbb{P}_B \left(\sup_{u \in [\tau, t]} \frac{|Y_u - Y_\tau|}{c(3.7)} > \lambda \right) + c_0 \mathbb{P}_B \left(\sup_{u \in [\tau, t]} w_{p,s,u,t} > \nu \right), \\ \text{(ii)} \quad & \mathbb{P}_B \left(\sup_{u \in [\tau, t]} \left| \frac{\int_\tau^u Z_r dW_r}{d(3.7) \beta_p} \right| > \lambda + c_0 \mu \nu \right) \leq e^{1-\mu} \mathbb{P}_B \left(\sup_{u \in [\tau, t]} \left| \frac{\int_\tau^u Z_r dW_r}{d(3.7) \beta_p} \right| > \lambda \right) + c_0 \mathbb{P}_B \left(\sup_{u \in [\tau, t]} w_{p,s,u,t}^{\xi, f} > \nu \right), \\ \text{(iii)} \quad & \mathbb{P}_B \left(\frac{\left(\int_\tau^t |Z_r|^2 dr \right)^{\frac{1}{2}}}{d(3.7)} > \lambda + c_0 \mu \nu \right) \leq e^{1-\mu} \mathbb{P}_B \left(\frac{\left(\int_\tau^t |Z_r|^2 dr \right)^{\frac{1}{2}}}{d(3.7)} > \lambda \right) + c_0 \mathbb{P}_B \left(\sup_{u \in [\tau, t]} w_{p,s,u,t}^{\xi, f} > \nu \right), \end{aligned}$$

for all $\lambda, \mu, \nu > 0$ and any $B \in \mathcal{F}_\tau$ of positive measure, and where β_p is the constant from Proposition 2.5.

Proof. As the tail estimates follow from Theorem 9.1 below in Appendix C, we show that the assumptions of Theorem 9.1 follow from Theorem 3.7. Assume (C1)-(C6) for $\theta \in [0, 1]$, $p \in [2, \infty) \cap (p_{(C4)}, \infty)$, and $0 \leq s < t \leq T$. Let $\epsilon > 0$, $\alpha \in (0, \frac{1}{2})$, and $R := t - s$.

(i) Define, for $r \in [0, R]$,

$$\begin{aligned} \mathcal{G}_r &:= \mathcal{F}_{r+s}, \\ A_r &:= \frac{(Y_{r+s} - Y_s) \alpha^{1/p}}{c(3.7)}, \\ \Psi_r &:= w_{p,s,r+s,t} \vee \epsilon, \end{aligned}$$

where $(w_{p,s,u,t})_{u \in [s,t]}$ is the weight process from Theorem 3.7. For $0 \leq a < b$ and a filtration $(\mathcal{H}_r)_{r \in [a,b]}$ we introduce the notation

$$\mathcal{S}_{a,b}^{\mathcal{H}} := \left\{ \tau : \Omega \rightarrow [a, b] \mid \tau \text{ is a } (\mathcal{H}_r)_{r \in [a,b]} \text{-stopping time} \right\},$$

so that in particular $\mathcal{S}_{0,R}^{\mathcal{G}} + s = \mathcal{S}_{s,t}^{\mathcal{F}}$. Then the assumptions of Theorem 9.1 are fulfilled. As the other assumptions are obvious, we will only show that equation (9.1) holds. Using Theorem 3.7 we deduce

$$\begin{aligned} \sup_{\tau \in \mathcal{S}_{0,R}^{\mathcal{G}}} \left\| \mathbb{E} \left[\frac{|A_R - A_\tau|^p}{\Psi_\tau^p} \mid \mathcal{G}_\tau \right] \right\|_\infty &= \sup_{\tau \in \mathcal{S}_{0,R}^{\mathcal{G}}} \left\| \mathbb{E} \left[\frac{|Y_t - Y_{\tau+s}|^p}{w_{p,s,\tau+s,t}^p \vee \epsilon^p} \mid \mathcal{F}_{\tau+s} \right] \right\|_\infty \frac{\alpha}{c(3.7)^p} \\ &= \sup_{\tilde{\tau} \in \mathcal{S}_{s,t}^{\mathcal{F}}} \left\| \mathbb{E} \left[\frac{|Y_t - Y_{\tilde{\tau}}|^p}{w_{p,s,\tilde{\tau},t}^p \vee \epsilon^p} \mid \mathcal{F}_{\tilde{\tau}} \right] \right\|_\infty \frac{\alpha}{c(3.7)^p} \leq \alpha. \end{aligned}$$

Hence we have by Chebyshev's inequality that for any $\nu > 0$, $\tau \in \mathcal{S}_{0,R}^{\mathcal{G}}$, and $B \in \mathcal{G}_\tau$ of positive measure:

$$\mathbb{P}_B(|A_R - A_\tau| > \nu) \leq \mathbb{P}_B(|A_R - A_\tau| > \Psi_\tau) + \mathbb{P}_B(\Psi_\tau > \nu) \leq \alpha + \mathbb{P}_B(\Psi_\tau > \nu).$$

Letting $\epsilon \rightarrow 0$ implies the claim.

(ii) The claim follows analogously as (i), when we choose

$$\begin{aligned}\mathcal{G}_r &:= \mathcal{F}_{r+s}, \\ A_r &:= \frac{\int_s^{r+s} Z_v dW_v \alpha^{1/p}}{d_{(3.7)} \beta_p}, \\ \Psi_r &:= w_{p,s,r+s,t}^{\xi,f} \vee \epsilon,\end{aligned}$$

where β_p is the constant from Proposition 2.5, as then we have by Theorem 3.7 that

$$\begin{aligned}\sup_{\tau \in \mathcal{S}_{0,R}^{\mathcal{G}}} \left\| \mathbb{E} \left[\frac{|A_R - A_\tau|^p}{\Psi_\tau^p} \mid \mathcal{G}_\tau \right] \right\|_{\infty} &= \sup_{\tau \in \mathcal{S}_{0,R}^{\mathcal{G}}} \left\| \mathbb{E} \left[\frac{\left| \int_{\tau+s}^t Z_v dW_v \right|^p}{\left(w_{p,s,\tau+s,t}^{\xi,f} \right)^p \vee \epsilon^p} \mid \mathcal{F}_{\tau+s} \right] \right\|_{\infty} \frac{\alpha}{\left(d_{(3.7)} \beta_p \right)^p} \\ &\leq \sup_{\bar{\tau} \in \mathcal{S}_{s,t}^{\mathcal{F}}} \left\| \mathbb{E} \left[\frac{\left(\int_{\bar{\tau}}^t |Z_v|^2 dv \right)^{\frac{p}{2}}}{\left(w_{p,s,\bar{\tau},t}^{\xi,f} \right)^p \vee \epsilon^p} \mid \mathcal{F}_{\bar{\tau}} \right] \right\|_{\infty} \frac{\alpha}{\left(d_{(3.7)} \right)^p} \leq \alpha.\end{aligned}$$

(iii) The claim follows analogously, when we choose

$$\begin{aligned}\mathcal{G}_r &:= \mathcal{F}_{r+s}, \\ A_r &:= \frac{\left(\int_s^{r+s} |Z_v|^2 dv \right)^{\frac{1}{2}} \alpha^{\frac{1}{p}}}{d_{(3.7)}}, \\ \Psi_r &:= w_{p,s,r+s,t}^{\xi,f} \vee \epsilon.\end{aligned}$$

□

3.2. Decoupled FBSDEs

We fix $x \in \mathbb{R}^d$ and consider the decoupled FBSDE

$$\begin{aligned}X_t &= x + \int_0^t b(r, X_r) dr + \int_0^t \sigma(r, X_r) dW_r, \quad t \in [0, T], \\ Y_t &= g(X_T) + \int_t^T h(r, X_r, Y_r, Z_r) dr - \int_t^T Z_r dW_r, \quad t \in [0, T].\end{aligned}\tag{3.3}$$

3.9 Assumption. The functions $b : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$, $\sigma : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$ and $h : [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$ are continuous, and furthermore we assume:

(A_{b,σ}) There exists $L_{b,\sigma} > 0$ such that for all $0 \leq t \leq T$ and $x, y \in \mathbb{R}^d$ one has

$$|b(t, x) - b(t, y)| + |\sigma(t, x) - \sigma(t, y)| \leq L_{b,\sigma} |x - y|.$$

(A_g) There exists $L_g > 0$ such that for all $x, y \in \mathbb{R}^d$ one has

$$|g(x) - g(y)| \leq L_g |x - y|.$$

(A_h) There exists $L_h > 0$ such that for all $0 \leq t \leq T$ and $x_i, z_i \in \mathbb{R}^d$, $y_i \in \mathbb{R}$, $i = 1, 2$, one has

$$|h(t, x_1, y_1, z_1) - h(t, x_2, y_2, z_2)| \leq L_h (|x_1 - x_2| + |y_1 - y_2| + |z_1 - z_2|).$$

3.10 Remark.

- (1) In particular it follows from Assumption 3.9, that there exist constants $L_h, K_h, K_{b,\sigma} > 0$ such that we have

$$\begin{aligned} |h(t, x, y, z)| &\leq K_h + L_h(|x| + |y| + |z|), \\ |b(t, x)| + |\sigma(t, x)| &\leq K_{b,\sigma}(1 + |x|), \end{aligned}$$

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

- (2) Under Assumption 3.9, there exists a unique solution (X, Y, Z) to FBSDE (3.3) and it holds

$$\mathbb{E} \left[\sup_{r \in [0, T]} |X_r|^p + \sup_{r \in [0, T]} |Y_r|^p + \left(\int_0^T |Z_r|^2 dr \right)^{\frac{p}{2}} \right] < \infty$$

for all $p \geq 2$ (see for example [7, Theorem 4.2]).

Assumption $(A_{b,\sigma})$ is a classical assumption for the forward equation. If $(A_{b,\sigma})$ holds, then we have a weighted BMO-estimate for the forward process X (see Lemma 5.6). Using this together with (A_g) and (A_h) we receive a weighted BMO-estimate for (Y, Z) , which gives us a tail-estimate for (Y, Z) . If we assume in addition to $(A_{b,\sigma})$ that σ is bounded, then the BMO-estimates for (X, Y, Z) are improved.

(A_σ) There exists $K_\sigma > 0$ such that for all $0 \leq t \leq T$ and $x \in \mathbb{R}^d$ one has

$$|\sigma(t, x)| \leq K_\sigma.$$

First let us give the weights from Assumption 3.5 for the FBSDE case:

3.11 Example. Assume that Assumption 3.9 holds. Then assumptions (C1)-(C6) hold true for $\theta = 0$, all $p \in [2, \infty)$, and all $0 \leq s < t \leq T$. Moreover, there exists $c_{(3.11)} > 0$ depending at most on $(T, d, p, L_g, L_h, L_{b,\sigma}, K_{b,\sigma})$ such that we may choose

$$w_{p,s,u,t}^f = w_{p,s,u,t}^\xi = c_{(3.11)}(t-u)^{1/2} \left(1 + \mathbb{E} \left[\sup_{r \in [u,t]} |X_r|^p \mid \mathcal{F}_u \right] \right)^{\frac{1}{p}}$$

for all $0 \leq s \leq u \leq t \leq T$. If additionally (A_σ) holds, then there exists $d_{(3.11)} > 0$ depending at most on $(T, d, p, L_g, L_h, L_{b,\sigma}, K_\sigma)$ such that we may choose

$$w_{p,s,u,t}^f = w_{p,s,u,t}^\xi = d_{(3.11)}(t-u)^{1/2}.$$

Example 3.11 is proved in Section 5.2 below. Now our first result is a consequence of Theorem 3.7:

3.12 Theorem. Assume that Assumption 3.9 holds and let $p \in [2, \infty)$. Then the following assertions hold true:

- (i)_Y There exists $c_{(3.12)} > 0$, depending at most on $(T, d, p, L_h, L_g, L_{b,\sigma}, K_{b,\sigma}, K_h)$, such that for any $0 \leq s < t \leq T$ and any stopping time $\tau : \Omega \rightarrow [s, t]$ we have

$$\mathbb{E} \left(|Y_t - Y_\tau|^p \mid \mathcal{F}_\tau \right) \leq c_{(3.12)}^p (t-\tau)^{p/2} [1 + |X_\tau|^p].$$

(i)_Z There exists $C_{(3.12)} > 0$ depending at most on $(T, d, p, L_h, L_g, L_{b,\sigma}, K_{b,\sigma})$ such that for any $0 \leq s < t \leq T$ and any stopping time $\tau : \Omega \rightarrow [s, t]$ we have

$$\mathbb{E} \left(\left(\int_{\tau}^t |Z_r|^2 dr \right)^{\frac{p}{2}} \middle| \mathcal{F}_{\tau} \right) \leq C_{(3.12)}^p (t - \tau)^{\frac{p}{2}} [1 + |X_{\tau}|^p].$$

If additionally (A_{σ}) holds, then we have:

(ii)_Y There exists $d_{(3.12)} > 0$, depending at most on $(T, d, p, L_h, L_g, L_{b,\sigma}, K_{b,\sigma}, K_h, K_{\sigma})$, such that for any $0 \leq s < t \leq T$ and any stopping time $\tau : \Omega \rightarrow [s, t]$ we have

$$\mathbb{E} \left(|Y_t - Y_{\tau}|^p \middle| \mathcal{F}_{\tau} \right) \leq d_{(3.12)}^p (t - \tau)^{p/2} [1 + |X_{\tau}|^p (t - \tau)^{p/2}].$$

(ii)_Z There exists $D_{(3.12)} > 0$ depending at most on $(T, d, p, L_h, L_g, L_{b,\sigma}, K_{\sigma})$ such that for any $0 \leq s < t \leq T$ and any stopping time $\tau : \Omega \rightarrow [s, t]$ we have

$$\mathbb{E}^{\mathcal{F}_{\tau}} \left(\int_{\tau}^t |Z_r|^2 dr \right)^{\frac{p}{2}} \leq D_{(3.12)}^p (t - \tau)^{\frac{p}{2}}.$$

Proof. (i)_Y Because of Example 3.11 we may use Theorem 3.7 to obtain for any $0 \leq s < t \leq T$ and any stopping time $\tau \in [s, t]$ that

$$\begin{aligned} \frac{1}{c_{(3.7)}^p} \mathbb{E}^{\mathcal{F}_{\tau}} |Y_t - Y_{\tau}|^p &\leq 2c_{(3.11)}^p (t - \tau)^{p/2} \left(1 + \mathbb{E} \left[\sup_{r \in [\tau, t]} |X_r|^p \middle| \mathcal{F}_{\tau} \right] \right) + \mathbb{E}^{\mathcal{F}_{\tau}} \left(\int_{\tau}^t |h(r, X_r, 0, 0)| dr \right)^p \\ &\quad + (t - \tau)^p \mathbb{E}^{\mathcal{F}_{\tau}} \left(g(X_T)^p + \left(\int_t^T |h(r, X_r, 0, 0)| dr \right)^p \right). \end{aligned}$$

Using (A_g) , (A_h) and the fact

$$\mathbb{E}^{\mathcal{F}_{\tau}} \sup_{\tau \leq r \leq T} |X_r|^p \leq C^p (1 + |X_{\tau}|^p), \tag{3.4}$$

where C depends at most on $(T, p, K_{b,\sigma})$, we may deduce

$$\mathbb{E}^{\mathcal{F}_{\tau}} |Y_t - Y_{\tau}|^p \leq c_{(3.12)}^p (t - \tau)^{p/2} [1 + |X_{\tau}|^p],$$

where $c_{(3.12)} > 0$ depends at most on $(T, d, p, L_g, L_h, L_{b,\sigma}, K_{b,\sigma}, K_h)$. Assertions (i)_Z and (ii)_Y follow analogously by applying Example 3.11, Theorem 3.7, and inequality (3.4). Assertion (ii)_Z, on the other hand, follows directly from Example 3.11 and Theorem 3.7. \square

One application of Theorem 3.12 are tail estimates of exponential type for (Y, Z) . In Theorem 3.13 we treat the process Y and in Theorem 3.14 the process Z . These theorems follow from Theorem 3.12 using Theorem 9.1 analogous to the proof of Theorem 3.8.

3.13 Theorem. Assume that Assumption 3.9 holds. Then there exists an absolute constant $c_0 > 0$ such that the following holds:

(i) For any $0 \leq s < t \leq T$ and any stopping time $\tau : \Omega \rightarrow [s, t]$ we have

$$\mathbb{P}_B \left(\sup_{u \in [\tau, t]} \frac{|Y_u - Y_\tau|}{c_{(3.12)} \sqrt{t-s}} > \lambda + c_0 \mu \nu \right) \leq e^{1-\mu} \mathbb{P}_B \left(\sup_{u \in [\tau, t]} \frac{|Y_u - Y_\tau|}{c_{(3.12)} \sqrt{t-s}} > \lambda \right) + c_0 \mathbb{P}_B \left(\sup_{u \in [\tau, t]} |X_u|^2 > \nu^2 - 1 \right)$$

for all $\lambda, \mu, \nu > 0$ and all $B \in \mathcal{F}_\tau$ of positive measure.

(ii) If additionally (A_σ) holds, then we have

$$\begin{aligned} \mathbb{P}_B \left(\sup_{u \in [\tau, t]} \frac{|Y_u - Y_\tau|}{d_{(3.12)} \sqrt{t-s}} > \lambda + c_0 \mu \nu \right) \\ \leq e^{1-\mu} \mathbb{P}_B \left(\sup_{u \in [\tau, t]} \frac{|Y_u - Y_\tau|}{d_{(3.12)} \sqrt{t-s}} > \lambda \right) + c_0 \mathbb{P}_B \left(\sup_{u \in [\tau, t]} |X_u|^2 (t-u) > \nu^2 - 1 \right) \end{aligned}$$

for all $\lambda, \mu, \nu > 0$ and all $B \in \mathcal{F}_\tau$ of positive measure.

3.14 Theorem. Assume that Assumption 3.9 holds. Then there exists an absolute constant $c_0 > 0$ such that the following holds:

(i) For any $0 \leq s < t \leq T$ and any stopping time $\tau : \Omega \rightarrow [s, t]$ we have

$$\begin{aligned} \mathbb{P}_B \left(\sup_{u \in [\tau, t]} \left| \frac{\int_\tau^u Z_r dW_r}{C_{(3.12)} \beta_2 \sqrt{t-s}} \right| > \lambda + c_0 \mu \nu \right) \\ \leq e^{1-\mu} \mathbb{P}_B \left(\sup_{u \in [\tau, t]} \left| \frac{\int_\tau^u Z_r dW_r}{C_{(3.12)} \beta_2 \sqrt{t-s}} \right| > \lambda \right) + c_0 \mathbb{P}_B \left(\sup_{u \in [\tau, t]} |X_u|^2 > \nu^2 - 1 \right) \end{aligned}$$

for all $\lambda, \mu, \nu > 0$ and all $B \in \mathcal{F}_\tau$ of positive measure, and where β_2 is the constant from Proposition 2.5.

(ii) If additionally (A_σ) holds, then

$$\mathbb{P}_B \left(\sup_{u \in [\tau, t]} \left| \frac{\int_\tau^u Z_r dW_r}{D_{(3.12)} \beta_2 \sqrt{t-s}} \right| > \lambda + c_0 \mu \right) \leq e^{1-\mu} \mathbb{P}_B \left(\sup_{u \in [\tau, t]} \left| \frac{\int_\tau^u Z_r dW_r}{D_{(3.12)} \beta_2 \sqrt{t-s}} \right| > \lambda \right)$$

for all $\lambda, \mu > 0$ and all $B \in \mathcal{F}_\tau$ of positive measure, and where β_2 is the constant from Proposition 2.5.

One might ask if it is necessary to use the theory of *weighted* BMO instead of non-weighted BMO. The following example shows that the weight processes of Theorem 3.12 (i)_Y and (i)_Z are sharp:

3.15 Example. Consider the FBSDE

$$X_t = \int_0^t \sqrt{4e^{-s} + X_s^2} dW_s, \quad t \in [0, T],$$

$$Y_t = X_T - \int_t^T Z_r dW_r, \quad t \in [0, T].$$

This FBSDE is of the same form as (3.3) with $d = 1$, $b \equiv 0$, $\sigma(t, x) = \sqrt{4e^{-t} + x^2}$, $h \equiv 0$ and $g(x) = x$, so that Assumption 3.9 holds. Now we have for all $p \in [2, \infty)$ and all $0 \leq s < t \leq T$ that

$$\mathbb{E} \left[|Y_t - Y_s|^p \mid \mathcal{F}_s \right] \geq \frac{(t-s)^{\frac{p}{2}} (1 + |X_s|^p)}{e^{\frac{Tp}{2}}},$$

as well as

$$\mathbb{E} \left[\left(\int_s^t |Z_r|^2 dr \right)^{\frac{p}{2}} \mid \mathcal{F}_s \right] \geq \frac{(t-s)^{\frac{p}{2}} (1 + |X_s|^p)}{\beta_p^p e^{\frac{Tp}{2}}},$$

where β_p is the constant from Proposition 2.5.

Proof. First note that $Y_t = X_t$, and that $X_t = 2 \sinh(W_t) e^{-\frac{t}{2}} = e^{W_t - \frac{t}{2}} - e^{-W_t - \frac{t}{2}}$. Furthermore, we have the equalities:

$$\begin{aligned} \mathbb{E}^{\mathcal{F}_s} |e^{W_t - \frac{t}{2}} - e^{W_s - \frac{s}{2}}|^2 &= |e^{W_s - \frac{s}{2}}|^2 (e^{t-s} - 1), \\ \mathbb{E}^{\mathcal{F}_s} |e^{-W_t - \frac{t}{2}} - e^{-W_s - \frac{s}{2}}|^2 &= |e^{-W_s - \frac{s}{2}}|^2 (e^{t-s} - 1), \\ \mathbb{E}^{\mathcal{F}_s} (e^{W_t - \frac{t}{2}} - e^{W_s - \frac{s}{2}})(e^{-W_t - \frac{t}{2}} - e^{-W_s - \frac{s}{2}}) &= e^{-t} (1 - e^{t-s}), \\ |X_s|^2 + 2e^{-s} &= |e^{W_s - \frac{s}{2}}|^2 + |e^{-W_s - \frac{s}{2}}|^2, \end{aligned}$$

so that

$$\begin{aligned} \mathbb{E}^{\mathcal{F}_s} |Y_t - Y_s|^2 &= \mathbb{E}^{\mathcal{F}_s} \left| e^{W_t - \frac{t}{2}} - e^{W_s - \frac{s}{2}} - (e^{-W_t - \frac{t}{2}} - e^{-W_s - \frac{s}{2}}) \right|^2 \\ &= |e^{W_s - \frac{s}{2}}|^2 (e^{t-s} - 1) + |e^{-W_s - \frac{s}{2}}|^2 (e^{t-s} - 1) - 2e^{-t} (1 - e^{t-s}) \\ &= (e^{t-s} - 1) (|X_s|^2 + 2(e^{-t} + e^{-s})) \\ &\geq (t-s) (|X_s|^2 + 1) e^{-T}. \end{aligned}$$

Since $\frac{p}{2} \geq 1$, we also have

$$\mathbb{E}^{\mathcal{F}_s} |Y_t - Y_s|^p \geq \left(\mathbb{E}^{\mathcal{F}_s} |Y_t - Y_s|^2 \right)^{\frac{p}{2}} \geq \left(e^{-T} (t-s) (1 + |X_s|^2) \right)^{\frac{p}{2}} \geq e^{-\frac{Tp}{2}} (t-s)^{\frac{p}{2}} (1 + |X_s|^p).$$

The result for the Z-process follows now immediately from

$$\int_s^t Z_r dW_r = Y_t - Y_s.$$

□

The following example shows that the weight processes of Theorem 3.12 (ii)_Y and (ii)_Z are sharp:

3.16 Example. Consider the FBSDE

$$X_t = \int_0^t 1 dW_s, \quad t \in [0, T],$$

$$Y_t = X_T + \int_t^T X_s ds - \int_t^T Z_s dW_s, \quad t \in [0, T],$$

This FBSDE is of the same form as (3.3) with $d = 1, b \equiv 0, \sigma \equiv 1, h(t, x, y, z) = x$, and $g(x) = x$, so that Assumptions 3.9 and (A_σ) hold. Now we have for all $p \in [2, \infty)$, and all $0 \leq s < t \leq T$ that

$$\mathbb{E} \left[|Y_t - Y_s|^p \mid \mathcal{F}_s \right] \geq (t-s)^{p/2} (1 + |X_s|^p (t-s)^{p/2})$$

as well as

$$\mathbb{E}^{\mathcal{F}_s} \left(\int_s^t |Z_r|^2 dr \right)^{\frac{p}{2}} \geq (t-s)^{\frac{p}{2}}.$$

Proof. We have for all $r \in [0, T]$ that

$$Y_r = \mathbb{E}^{\mathcal{F}_r} \left[W_T + \int_r^T W_u du \right] = W_r(1 + T - r),$$

and therefore

$$\begin{aligned} \mathbb{E}^{\mathcal{F}_s} |Y_t - Y_s|^2 &= \mathbb{E}^{\mathcal{F}_s} |W_t(1 + T - t) - W_s(1 + T - s)|^2 \\ &= (t-s)(1 + T - t)^2 + |W_s|^2 (t-s)^2 \\ &\geq (t-s)(1 + |W_s|^2 (t-s)). \end{aligned}$$

Since $\frac{p}{2} \geq 1$, we deduce

$$\mathbb{E}^{\mathcal{F}_s} |Y_t - Y_s|^p \geq \left[\mathbb{E}^{\mathcal{F}_s} |Y_t - Y_s|^2 \right]^{p/2} \geq \left[(t-s)(1 + |W_s|^2 (t-s)) \right]^{p/2} \geq (t-s)^{p/2} (1 + |W_s|^p (t-s)^{p/2}).$$

The result for the Z -process follows immediately from the fact that $Z_r = 1 + (T - r)$. \square

4. Decoupling operators

We now recall the decoupling operators introduced in [18], as well as some of their properties proven there. These operators are defined for random objects based on $\bar{\Omega}$, see Section 4.1 below, but we will use them to deduce *conditional estimates in the original probability space* $(\Omega, \mathcal{F}, \mathbb{P})$. These results are crucial in proving Theorem 3.7.

4.1. Setting

Recall the stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \in [0, T]})$ that was fixed in the beginning of Section 2. Our fundamental random object is the Brownian motion $W = (W_t)_{t \in [0, T]}$, but for our decoupling technique we also need to have a Brownian motion W' that is independent of W . Thus we proceed as follows:

Step 1. Fix another stochastic basis $(\Omega', \mathcal{F}', \mathbb{P}', (\mathcal{F}'_t)_{t \in [0, T]})$ and a standard d -dimensional Brownian motion $W' = (W'_t)_{t \in [0, T]}$ that satisfy the same assumptions as imposed on $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \in [0, T]}, (W_t)_{t \in [0, T]})$ in Section 2.

Step 2. Let

$$\bar{\Omega} := \Omega \times \Omega', \quad \bar{\mathbb{P}} := \mathbb{P} \times \mathbb{P}', \quad \bar{\mathcal{F}} := \overline{\mathcal{F} \otimes \mathcal{F}'^{\bar{\mathbb{P}}}}.$$

Step 3. Extend the Brownian motions W and W' canonically to $\bar{\Omega}$, that is,

$$\begin{aligned} W(\omega, \omega') &:= W(\omega), \\ W'(\omega, \omega') &:= W'(\omega'). \end{aligned}$$

The augmented⁷ natural filtration of the $2d$ -dimensional Brownian motion (W, W') is denoted by $\bar{\mathbb{F}} = (\bar{\mathcal{F}}_t)_{t \in [0, T]}$.

Hence, on the probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$, there are two independent d -dimensional Brownian motions W and W' . Fix a Borel-measurable function $\varphi : (0, T] \rightarrow [0, 1]$. We define another standard d -dimensional Brownian motion on $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$ by

$$W_t^\varphi := \int_0^t \sqrt{1 - \varphi(u)^2} dW_u + \int_0^t \varphi(u) dW'_u, \quad t \in [0, T],$$

and assume again continuity of all trajectories. The augmented natural filtration of W^φ is denoted by $\mathbb{F}^\varphi = (\mathcal{F}_t^\varphi)_{t \in [0, T]}$ and we obtain another stochastic basis

$$(\bar{\Omega}, \mathcal{F}_T^\varphi, \bar{\mathbb{P}}, (\mathcal{F}_t^\varphi)_{t \in [0, T]})$$

and can define, as in Notation 2.1,

$$(\bar{\Omega}_S, \Sigma_S^\varphi, \bar{\mathbb{P}}_S) := \begin{cases} (\bar{\Omega}, \mathcal{F}_T^\varphi, \bar{\mathbb{P}}) & : S = 0, \\ ([0, T] \times \bar{\Omega}, \mathcal{B}([0, T]) \otimes \mathcal{F}_T^\varphi, \frac{\lambda}{T} \otimes \bar{\mathbb{P}}) & : S = T. \end{cases} \quad (4.1)$$

Furthermore, we denote the predictable σ -algebra on the stochastic basis $(\bar{\Omega}, \mathcal{F}_T^\varphi, \bar{\mathbb{P}}, (\mathcal{F}_t^\varphi)_{t \in [0, T]})$ by \mathcal{P}^φ . Denoting the function $\varphi \equiv 0$ simply by 0, we have that W^0 and (the extension of) W are indistinguishable. Since \mathcal{F}^0 contains all $\bar{\mathbb{P}}$ -nullsets, it follows that $(\mathcal{F}_t^0)_{t \in [0, T]}$ and the augmentation of $\sigma(W_r, r \in [0, t])_{t \in [0, T]}$ coincide. Thus, we may agree to use W^0 for the extension of W , and similarly we use W^1 for the extension of W' .

4.2. Decoupling operators

Given a random variable ξ , whose randomness is given by W , we wish to define a random variable ξ^φ with the following two properties:

- (1) ξ^φ is a copy of ξ ,
- (2) The randomness of ξ^φ is given by W^φ .

We accomplish this at the level of equivalence classes. The fact that our procedure is well-defined is not proven here; all the proofs can be found in [18].

Step 1. For $\xi \in \mathcal{L}_0(\Omega, \mathcal{F}, \mathbb{P})$ take the canonical extension $\tilde{\xi} \in \mathcal{L}_0(\bar{\Omega}, \mathcal{F}^0, \bar{\mathbb{P}})$, and let $[\xi] \in L_0(\bar{\Omega}, \mathcal{F}^0, \bar{\mathbb{P}})$ be the equivalence-class that contains all \mathcal{F}^0 -measurable random variables that are $\bar{\mathbb{P}}$ -a.s. the same as $\tilde{\xi}$.

Step 2. We let $(h_k)_{k \in \mathbb{N}}$ be the $(L_2([0, T])$ -normalized) Haar-functions on $[0, T]$, and denote by $W_{s,i}^0$ the i :th component of the Brownian motion W^0 for $i = 1, \dots, d$. Now, letting $(g_n)_{n \in \mathbb{N}} : \bar{\Omega} \rightarrow \mathbb{R}$ be the family of random variables $\int_0^T h_k(s) dW_{s,i}^0$ where $i = 1, \dots, d$ and $k \in \mathbb{N}$, there exists a $\sigma(g_n, n \in \mathbb{N})$ -measurable $\xi^0 \in [\xi]$.

⁷Whenever we augment a filtration that is based on $\bar{\Omega}$, we augment it by $\bar{\mathbb{P}}$ -nullsets.

Step 3. Defining $J : \bar{\Omega} \rightarrow \mathbb{R}^{\mathbb{N}}$, $J(\eta) = (g_n(\eta))_{n \in \mathbb{N}}$, there exists a random variable $\hat{\xi} : \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}$ such that ξ^0 can be factorized through $\mathbb{R}^{\mathbb{N}}$:

$$\xi^0 : \bar{\Omega} \xrightarrow{J} \mathbb{R}^{\mathbb{N}} \xrightarrow{\hat{\xi}} \mathbb{R}.$$

Step 4. Define $(g_n^\varphi)_{n \in \mathbb{N}}$ analogously as $(g_n)_{n \in \mathbb{N}}$, using W^φ instead of W^0 , and let $J^\varphi : \bar{\Omega} \rightarrow \mathbb{R}^{\mathbb{N}}$, $J^\varphi(\eta) = (g_n^\varphi(\eta))_{n \in \mathbb{N}}$. Then it follows that $\hat{\xi}(J^\varphi)$ is a well-defined $\sigma(g_n^\varphi, n \in \mathbb{N})$ -measurable random variable.

Step 5. Finally, we let $[\xi]^\varphi \in L_0(\bar{\Omega}, \mathcal{F}^\varphi, \bar{\mathbb{P}})$ be the equivalence-class that contains all \mathcal{F}^φ -measurable random variables that are $\bar{\mathbb{P}}$ -a.s. the same as $\hat{\xi}(J^\varphi)$.

4.1 Remark.

(1) Steps 2-5 yield the *decoupling operator* $C : L_0(\bar{\Omega}, \mathcal{F}^0, \bar{\mathbb{P}}) \rightarrow L_0(\bar{\Omega}, \mathcal{F}^\varphi, \bar{\mathbb{P}})$ defined by

$$C([\xi]) = [\xi]^\varphi.$$

In the following we will identify ξ , $\tilde{\xi}$, and $[\xi]$, and denote all of them simply by ξ . Similarly, we will use the notation ξ^φ for both the equivalence class $[\xi]^\varphi$, and any representative of it.

- (2) The factorization and the approach used here is distributional, and does not require continuous paths or a gaussian distribution. As such, the approach might be useful also in other situations.
- (3) We can define X^φ for $X \in \mathcal{L}_0(\Omega_T, \Sigma_T, \mathbb{P}_T)$ analogously as above. The idea is that we change the randomness, but leave the time component unchanged. The point of defining this separately is to emphasize that $X^\varphi \in L_0(\bar{\Omega}_T, \Sigma_T^\varphi, \bar{\mathbb{P}}_T)$, i.e. that representatives of X^φ are jointly measurable.
- (4) Our approach preserves continuity: Assume that \mathbb{X} is locally σ -compact, $S \in \{0, T\}$, and $f \in \mathcal{L}_0(\Omega_S, \Sigma_S, \mathbb{P}_S; C(\mathbb{X}))$. Then we may define $f^\varphi \in L_0(\bar{\Omega}_S, \Sigma_S^\varphi, \bar{\mathbb{P}}_S; C(\mathbb{X}))$ by taking the continuous modification⁸ of $(f(x)^\varphi)_{x \in \mathbb{X}}$.

4.3. Basic properties

Predictability and adaptedness are transferred in the following sense:

4.2 Proposition ([18, Lemma 3.1 and Proposition 2.12]). *Let \mathbb{X} locally σ -compact. Then the following holds true:*

- (i) *If $\xi \in \mathcal{L}_0(\Omega, \mathcal{F}_t, \mathbb{P})$ for some $t \in [0, T]$, then all representatives of $\xi^\varphi \in L_0(\bar{\Omega}, \mathcal{F}^\varphi, \bar{\mathbb{P}})$ are \mathcal{F}_t^φ -measurable.*
- (ii) *If $f \in \mathcal{L}_0(\Omega_T, \mathcal{P}, \mathbb{P}_T; C(\mathbb{X}))^9$, then there is a \mathcal{P}^φ -measurable¹⁰ representative of*

$$f^\varphi \in L_0(\bar{\Omega}_T, \Sigma_T^\varphi, \bar{\mathbb{P}}_T; C(\mathbb{X})).$$

- (iii) *If $Y \in \mathcal{L}_0(\Omega, \mathcal{F}, \mathbb{P}; C([0, T]))$ is $(\mathcal{F}_t)_{t \in [0, T]}$ -adapted, then all representatives of*

$$Y^\varphi \in L_0(\bar{\Omega}, \mathcal{F}^\varphi, \bar{\mathbb{P}}; C([0, T]))$$

are $(\mathcal{F}_t^\varphi)_{t \in [0, T]}$ -adapted.

⁸Existence of such modification was proven in [18, Proposition A.1].

⁹This means that $\eta \mapsto f(\eta, x)$ is \mathcal{P} -measurable for all $x \in \mathbb{X}$.

¹⁰This means that $\eta \mapsto f^\varphi(\eta, x)$ is \mathcal{P}^φ -measurable for all $x \in \mathbb{X}$.

We summarize some further properties proven in [18]:

4.3 Proposition ([18, Propositions 2.5, 2.13, and Lemma 3.2]). *Let $N \geq 1$, $S \in \{0, T\}$, $X, X_1, \dots, X_N \in \mathcal{L}_0(\Omega_S, \Sigma_S, \mathbb{P}_S)$, $Y \in \mathcal{L}_1(\Omega_T, \Sigma_T, \mathbb{P}_T)$, $g : \mathbb{R}^N \rightarrow \mathbb{R}$ be a Borel function, $f \in \mathcal{L}_0(\Omega_S, \Sigma_S, \mathbb{P}_S; C(\mathbb{R}^N))$, and $Z \in \mathcal{L}_2(\Omega_T, \mathcal{P}, \mathbb{P}_T)$. Then the following holds true:*

- (i) $X \stackrel{d}{=} X^\varphi$.
- (ii) $(g(X_1, \dots, X_N))^\varphi = g(X_1^\varphi, \dots, X_N^\varphi)$.
- (iii) $(f(X_1, \dots, X_N))^\varphi = f^\varphi(X_1^\varphi, \dots, X_N^\varphi)$.
- (iv) $\left(\int_0^T Y(t) 1_{\{\int_0^t |Y(s)| ds < \infty\}} dt \right)^\varphi = \int_0^T Y^\varphi(t) 1_{\{\int_0^t |Y^\varphi(s)| ds < \infty\}} dt$.
- (v) $\left(\int_0^T Z(t) dW_t \right)^\varphi = \int_0^T Z^\varphi(t) dW_t^\varphi$ for any predictable representative of Z^φ .¹¹
- (vi) Let $X \in \mathcal{L}_0(\Omega_T, \Sigma_T, \mathbb{P}_T)$ and $Y \in \mathcal{L}_0(\overline{\Omega}_T, \overline{\Sigma}_T, \overline{\mathbb{P}}_T)$. If there is a null-set $\mathcal{N} \subseteq [0, T]$ with $Y(t) \in X(t)^\varphi$ for all $t \in [0, T] \setminus \mathcal{N}$, then $Y \in X^\varphi$.

Our next result can be interpreted as follows: if (Y, Z) is a solution of an SDE, then (Y^φ, Z^φ) is a solution of another SDE. Note that we do not assume the SDEs to have unique solutions, we only assume that (Y, Z) satisfies the equation.

4.4 Proposition ([18, Theorem 3.3]). *Assume that $f, g_i \in \mathcal{L}_0(\Omega_T, \mathcal{P}, \mathbb{P}_T; C(\mathbb{R}^{1+d}))$, $Z_i \in \mathcal{L}_0(\Omega_T, \mathcal{P}, \mathbb{P}_T)$, $i = 1, \dots, d$, that $Y \in \mathcal{L}_0(\Omega, \mathcal{F}, \mathbb{P}; C([0, T]))$ is $(\mathcal{F}_t)_{t \in [0, T]}$ -adapted, and that*

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

Furthermore, assume that $\xi \in \mathcal{L}_0(\Omega, \mathcal{F}, \mathbb{P})$, and that equation

$$Y_u = \xi + \int_u^T f(r, Y_r, Z_r) dr - \int_u^T g(r, Y_r, Z_r) dW_r, \quad u \in [0, T], \quad (4.2)$$

holds \mathbb{P} -almost surely. If we fix any predictable representatives of $f^\varphi, g_i^\varphi, Z_i^\varphi$, and an $(\mathcal{F}_t^\varphi)_{t \in [0, T]}$ -adapted (continuous) representative of Y^φ , we have

$$\mathbb{E} \left[\int_0^T |f^\varphi(r, Y_r^\varphi, Z_r^\varphi)| dr + \int_0^T |g^\varphi(r, Y_r^\varphi, Z_r^\varphi)|^2 dr \right] < \infty,$$

and we have that the equation

$$Y_u^\varphi = \xi^\varphi + \int_u^T f^\varphi(r, Y_r^\varphi, Z_r^\varphi) dr - \int_u^T g^\varphi(r, Y_r^\varphi, Z_r^\varphi) dW_r^\varphi, \quad u \in [0, T], \quad (4.3)$$

holds $\overline{\mathbb{P}}$ -almost surely.

¹¹By Proposition 4.2(ii) there exists such a representative.

4.4. Conditional results

From now on we will exclusively use functions φ of the form

$$\chi_{(s,t]} : (0, T] \rightarrow [0, 1], \quad \chi_{(s,t]}(r) = \begin{cases} 1, & \text{if } r \in (s, t], \\ 0, & \text{if } r \notin (s, t], \end{cases}$$

where $0 \leq s < t \leq T$.¹² To keep the notation light, we let

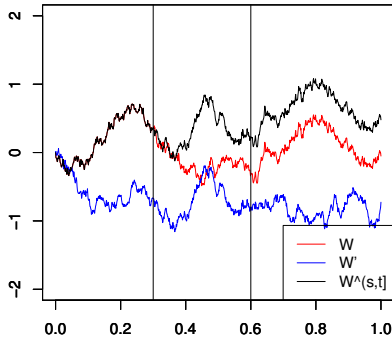


Figure 1: Brownian motions W, W' and $W^{(s,t]}$. Here $s = 0.3, t = 0.6$ and $T = 1$.

$$X^{(s,t]} := X^{\chi_{(s,t]}}. \quad (4.4)$$

Recall that the random object $X^{(s,t]}$ is obtained by changing the underlying Brownian motion W^0 to an independent one on the interval $(s, t]$. If X is independent of $\sigma(W_r^0 - W_s^0, r \in (s, t])$, we ought to have $X^{(s,t]} = X$. Precisely in what sense this holds, is answered by the following proposition:

4.5 Proposition. Let $0 \leq s < t \leq T$, and define the sigma-algebras

$$\overline{\mathcal{G}}_s^t := \sigma(W_r^0, r \in [0, s]) \vee \sigma(W_r^0 - W_t^0, r \in [t, T]) \vee \overline{N} \quad \text{and} \quad \overline{\mathcal{H}}_s^t := \mathcal{B}([0, T]) \otimes \overline{\mathcal{G}}_s^t, \quad (4.5)$$

where \overline{N} are the $\overline{\mathbb{P}}$ -nullsets. Then

- (i) $\mathbb{E}^{\overline{\mathcal{H}}_s^t} X = \mathbb{E}^{\overline{\mathcal{H}}_s^t} X^{(s,t]}$ for any $X \in L_1(\overline{\Omega}_T, \Sigma_T^0, \overline{\mathbb{P}}_T)$,
- (ii) $\mathbb{E}^{\overline{\mathcal{G}}_s^t} \alpha = \mathbb{E}^{\overline{\mathcal{G}}_s^t} \alpha^{(s,t]}$ for any $\alpha \in L_1(\overline{\Omega}, \mathcal{F}^0, \overline{\mathbb{P}})$,
- (iii) $\alpha = \alpha^{(s,t]}$ for any $\alpha \in L_0(\overline{\Omega}, \overline{\mathcal{G}}_s^t, \overline{\mathbb{P}})$,
- (iv) $X \in X^{(s,t]}$ for any $X \in \mathcal{L}_0(\overline{\Omega}_T, \overline{\mathcal{H}}_s^t, \overline{\mathbb{P}}_T)$,
- (v) $f \in f^{(s,t]}$ for any $f \in \mathcal{L}_0(\overline{\Omega}_T, \overline{\mathcal{H}}_s^t, \overline{\mathbb{P}}_T; C(\mathbb{X}))$.

¹²For a picture of the different Brownian motions W, W' , and $W^{(s,t]}$ see Figure 1.

Proof. First of all we note that

$$\overline{\mathcal{G}}_s^t = \sigma(W_r^{(s,t)}, r \in [0, s]) \vee \sigma(W_r^{(s,t)} - W_t^{(s,t)}, r \in [t, T]) \vee \overline{N}.$$

Hence, similarly as in Proposition 4.2(i) (i.e. [18, Lemma 3.1]), we have that if $\alpha \in \mathcal{L}_0(\overline{\Omega}, \overline{\mathcal{G}}_s^t, \overline{\mathbb{P}})$, then all representatives of $\alpha^{(s,t)}$ are $\overline{\mathcal{G}}_s^t$ -measurable.

(i) We need to prove that

$$\int_A X d\overline{\mathbb{P}}_T = \int_A X^{(s,t)} d\overline{\mathbb{P}}_T$$

for all $A \in \overline{\mathcal{H}}_s^t$. By linearity of the decoupling operator we may assume that $X \geq 0$, and it is enough to consider a generating π -system, so that we assume A to be of the form

$$(r, \widehat{W}) := (r, W_{s_1}^0, \dots, W_{s_n}^0, W_{t_1}^0 - W_t^0, \dots, W_{t_m}^0 - W_t^0) \in B_1 \times B_2,$$

where $n, m \in \mathbb{N}$, $0 \leq s_i \leq s < t \leq t_j \leq T$, $0 \leq r \leq T$, and $B_1 \in \mathcal{B}([0, T])$, $B_2 \in \mathcal{B}(\mathbb{R}^{(n+m)d})$. Letting $Y(r, \omega) := \chi_{B_1 \times B_2}(r, \widehat{W}(\omega))$ we have $Y \in \mathcal{L}_\infty(\overline{\Omega}_T, \overline{\mathcal{H}}_s^t, \overline{\mathbb{P}}_T)$, and $Y \in Y^{(s,t)}$ because of Proposition 4.3(ii) and (vi). Thus, again using Proposition 4.3,

$$\begin{aligned} \int_A X d\overline{\mathbb{P}}_T &= \int_{\overline{\Omega}_T} XY d\overline{\mathbb{P}}_T = \int_{\overline{\Omega}_T} (XY)^{(s,t)} d\overline{\mathbb{P}}_T = \int_{\overline{\Omega}_T} X^{(s,t)} Y^{(s,t)} d\overline{\mathbb{P}}_T \\ &= \int_{\overline{\Omega}_T} X^{(s,t)} Y d\overline{\mathbb{P}}_T = \int_A X^{(s,t)} d\overline{\mathbb{P}}_T. \end{aligned}$$

(ii) Can be shown similarly as (i).

(iii) First let $\alpha \in L_1(\overline{\Omega}, \overline{\mathcal{G}}_s^t, \overline{\mathbb{P}})$. Then we have that $\mathbb{E}^{\overline{\mathcal{G}}_s^t} \alpha^{(s,t)} = \alpha^{(s,t)}$, but from (ii) we have that $\mathbb{E}^{\overline{\mathcal{G}}_s^t} \alpha^{(s,t)} = \alpha$ as well. For $\alpha \in L_0(\overline{\Omega}, \overline{\mathcal{G}}_s^t, \overline{\mathbb{P}})$ the claim follows from the fact that for all $N \in \mathbb{N}$

$$(N \wedge \alpha \vee (-N))^{(s,t)} = N \wedge \alpha^{(s,t)} \vee (-N).$$

(iv) If $X \in \mathcal{L}_0(\overline{\Omega}_T, \overline{\mathcal{H}}_s^t, \overline{\mathbb{P}}_T)$, then by Fubini's theorem $X(r) \in \mathcal{L}_0(\overline{\Omega}, \overline{\mathcal{G}}_s^t, \overline{\mathbb{P}})$ for all $r \in [0, T]$, so that (iii) implies that $X(r) \in X(r)^{(s,t)}$ for all $r \in [0, T]$. Since $\overline{\mathcal{H}}_s^t \subseteq \Sigma_T^{(s,t)}$, we have that $X \in \mathcal{L}_0(\overline{\Omega}_T, \Sigma_T^{(s,t)}, \overline{\mathbb{P}}_T)$ so that the claim follows from Proposition 4.3(vi).

(v) Follows directly from (iv) and the definition of $f^{(s,t)}$. \square

We want to deduce conditional estimates for random variables based on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ from estimates obtained using the decoupling operators. Recall, that \mathcal{F}^0 and \mathcal{F}_s^0 were defined in Section 4.1, $\overline{\mathcal{G}}_s^t$ and $\overline{\mathcal{G}}_s^t$ by equations (4.5), \mathcal{G}_s^t and $\overline{\mathcal{G}}_s^t$ by (2.1), and $\xi^{(a,b)}$ by (4.4). The following result is vital:

4.6 Lemma. *Let $p \in [1, \infty)$, $0 \leq s < t \leq T$, and $\xi \in \mathcal{L}_p(\overline{\Omega}, \mathcal{F}^0, \overline{\mathbb{P}})$. Then*

$$\frac{1}{2^p} \mathbb{E}^{\overline{\mathcal{G}}_s^t} |\xi - \xi^{(s,t)}|^p \leq \mathbb{E}^{\overline{\mathcal{G}}_s^t} |\xi - \mathbb{E}^{\overline{\mathcal{G}}_s^t} \xi|^p \leq \mathbb{E}^{\overline{\mathcal{G}}_s^t} |\xi - \xi^{(s,t)}|^p. \quad (4.6)$$

Proof. We know from [18, Lemma 4.23] that for any $X \in \mathcal{L}_p(\overline{\Omega}, \mathcal{F}^0, \overline{\mathbb{P}})$

$$\frac{1}{2} \|X - X^{(s,t)}\|_p \leq \|X - \mathbb{E}^{\overline{\mathcal{G}}_s^t} X\|_p \leq \|X - X^{(s,t)}\|_p. \quad (4.7)$$

Let $A \in \overline{\mathcal{G}}_s^t$ such that $\overline{\mathbb{P}}(A) > 0$. Using Propositions 4.3(ii) and 4.5(iii) we have $(\xi \chi_A)^{(s,t)} = \chi_A \xi^{(s,t)}$ and $\mathbb{E}^{\overline{\mathcal{G}}_s^t}(\xi \chi_A) = \chi_A \mathbb{E}^{\overline{\mathcal{G}}_s^t} \xi$, so that applying equation (4.7) with $X = \xi \chi_A$ implies the claim. \square

4.7 Corollary. Let $p \in [1, \infty)$, $0 \leq s < t \leq T$, $\xi \in \mathcal{L}_p(\Omega, \mathcal{F}, \mathbb{P})$ and $\Psi \in \mathcal{L}_0(\Omega, \mathcal{F}, \mathbb{P})$, and denote their canonical extensions by $\tilde{\xi}, \tilde{\Psi}$, respectively. Then

$$\mathbb{E}^{\mathcal{F}_s^0} |\tilde{\xi} - \xi^{(s,t)}|^p \leq \tilde{\Psi} \quad \Rightarrow \quad \mathbb{E}^{\mathcal{F}_s} |\xi - \mathbb{E}^{\mathcal{G}_s} \xi|^p \leq \Psi \quad \Rightarrow \quad \mathbb{E}^{\mathcal{F}_s^0} |\tilde{\xi} - \xi^{(s,t)}|^p \leq 2^p \tilde{\Psi}.$$

Proof. We have that $\tilde{\xi} \in \mathcal{L}_p(\overline{\Omega}, \mathcal{F}^0, \overline{\mathbb{P}})$, and that the canonical extension of $\mathbb{E}^{\mathcal{F}_s} |\xi - \mathbb{E}^{\mathcal{G}_s} \xi|^p$ is $\overline{\mathbb{P}}$ -a.s. equal to $\mathbb{E}^{\mathcal{F}_s^0} |\tilde{\xi} - \mathbb{E}^{\overline{\mathcal{G}_s}} \tilde{\xi}|^p$. Applying $\mathbb{E}^{\mathcal{F}_s^0}$ on equation (4.6), we have $\overline{\mathbb{P}}$ -a.s.

$$\frac{1}{2^p} \mathbb{E}^{\mathcal{F}_s^0} |\tilde{\xi} - \xi^{(s,t)}|^p \leq \mathbb{E}^{\mathcal{F}_s^0} |\tilde{\xi} - \mathbb{E}^{\overline{\mathcal{G}_s}} \tilde{\xi}|^p \leq \mathbb{E}^{\mathcal{F}_s^0} |\tilde{\xi} - \xi^{(s,t)}|^p, \quad (4.8)$$

and the claim follows. \square

The same idea applies also for the generator of a BSDE. However, the result corresponding to Lemma 4.6 being technically involved, is proven in the appendix.

4.8 Corollary. Assume that $f \in \mathcal{L}_0(\overline{\Omega}_T, \Sigma_T^0, \overline{\mathbb{P}}_T; C(\mathbb{R}^{1+d}))$ satisfies (C1) with $\theta = 1$ and (C2) with $p = 1$, let $q \in [1, \infty)$ and $0 \leq s < t \leq T$. Moreover, let $\Psi \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$, and denote its canonical extension by $\tilde{\Psi}$. Then

$$\mathbb{E}^{\mathcal{F}_s} \left(\int_s^T \sup_{x \in \mathbb{R}^{d+1}} |f(r, x) - (\mathbb{E}^{\mathcal{H}_s^t} f)(r, x)| dr \right)^q \leq \Psi \quad \Rightarrow \quad \mathbb{E}^{\mathcal{F}_s^0} \left(\int_s^T \sup_{x \in \mathbb{R}^{d+1}} |f(r, x) - f^{(s,t)}(r, x)| dr \right)^q \leq 2^q \tilde{\Psi},$$

and conversely,

$$\mathbb{E}^{\mathcal{F}_s^0} \left(\int_s^T \sup_{x \in \mathbb{R}^{d+1}} |f(r, x) - f^{(s,t)}(r, x)| dr \right)^q \leq \tilde{\Psi} \quad \Rightarrow \quad \mathbb{E}^{\mathcal{F}_s} \left(\int_s^T \sup_{x \in \mathbb{R}^{d+1}} |f(r, x) - (\mathbb{E}^{\mathcal{H}_s^t} f)(r, x)| dr \right)^q \leq \Psi.$$

Proof. The canonical extension of $\mathbb{E}^{\mathcal{F}_s} \left(\int_s^T \sup_{x \in \mathbb{R}^{d+1}} |f(r, x) - (\mathbb{E}^{\mathcal{H}_s^t} f)(r, x)| dr \right)^q$ is $\overline{\mathbb{P}}$ -a.s. equal to $\mathbb{E}^{\mathcal{F}_s^0} \left(\int_s^T \sup_{x \in \mathbb{R}^{d+1}} |f(r, x) - (\mathbb{E}^{\overline{\mathcal{H}_s^t}} f)(r, x)| dr \right)^q$, so that the result follows by applying $\mathbb{E}^{\mathcal{F}_s^0}$ to the conclusion of Proposition 8.6 with $u_1 = s$ and $u_2 = T$. To apply Proposition 8.6, we show that for all $R > 0$ it holds $\int_{\Omega_T} \sup_{x \in \overline{B}(0, R)} |f(x)| d\mathbb{P}_T < \infty$, where $\overline{B}(0, R) \subseteq \mathbb{R}^{1+d}$ is the closed ball of radius R . Indeed, it follows from (C2) and (C3) that

$$\begin{aligned} \int_{\Omega_T} \sup_{(y,z) \in \overline{B}(0, R)} |f(r, \omega, y, z)| d\mathbb{P}_T(r, \omega) &\leq \int_{\Omega_T} \sup_{(y,z) \in \overline{B}(0, R)} |f(r, \omega, 0, 0)| + L_y |y| + L_z (1 + |z|) |z| d\mathbb{P}_T(r, \omega) \\ &\leq \mathbb{E} \left(\int_0^T |f(r, 0, 0)| dr \right) + L_y R + L_z (1 + R) R < \infty. \end{aligned}$$

\square

5. Proof of Theorem 3.7 and Example 3.11

Again we use \mathcal{F}_u^0 from Section 4.1 and $\xi^{(a,b)}$ from (4.4). The following is the counterpart to Assumption 3.5:

5.1 Assumption. Let $p \in [2, \infty)$ and $0 \leq s < t \leq T$. There are non-negative càdlàg processes $(w_{p,s,u,t}^\xi)_{u \in [s,t]}$ and $(w_{p,s,u,t}^f)_{u \in [s,t]}$, such that $((w_{p,s,u,t}^\xi)^p)_{u \in [s,t]}$ and $((w_{p,s,u,t}^f)^p)_{u \in [s,t]}$ are $(\mathcal{F}_r)_{r \in [0,T]}$ -supermartingales, whose canonical extensions $(\tilde{w}_{p,s,u,t}^\xi)_{u \in [s,t]}$ and $(\tilde{w}_{p,s,u,t}^f)_{u \in [s,t]}$ satisfy, for any $u \in [s, t]$,

$$(\overline{C5}) \left(\mathbb{E}^{\mathcal{F}_u^0} |\xi - \xi^{(u,t)}|^p \right)^{\frac{1}{p}} \leq \tilde{w}_{p,s,u,t}^{\xi}$$

$$(\overline{C6}) \left(\mathbb{E}^{\mathcal{F}_u^0} \left(\int_u^T \sup_{y,z} |f(r,y,z) - f^{(u,t)}(r,y,z)| dr \right)^p \right)^{\frac{1}{p}} \leq \tilde{w}_{p,s,u,t}^f$$

5.2 Remark. It is immediate from Corollaries 4.7 and 4.8 that Assumptions 3.5 and 5.1 are equivalent in the sense, that when passing from one assumption to the other one can use the same weights multiplied by the factor 2.

5.1. Proof of Theorem 3.7

In this Section we deduce upper bounds for $\mathbb{E}^{\mathcal{F}_\tau} |Y_t - Y_\tau|^p$ and $\mathbb{E}^{\mathcal{F}_\tau} \left(\int_\tau^t |Z_r|^2 dr \right)^{\frac{p}{2}}$, where $\tau : \Omega \rightarrow [s, t]$ is any stopping time, and $0 \leq s < t \leq T$ are such that Assumption 3.5 is satisfied.

Our procedure consists of the following steps:

Step 1: Let $0 \leq s < t \leq T$, $u \in [s, t]$, and consider the decomposition

$$\left(\mathbb{E}^{\mathcal{F}_u} |Y_t - Y_u|^p \right)^{1/p} \leq \left(\mathbb{E}^{\mathcal{F}_u} |Y_t - \mathbb{E}^{\mathcal{F}_u} Y_t|^p \right)^{1/p} + \left(\mathbb{E}^{\mathcal{F}_u} |Y_u - \mathbb{E}^{\mathcal{F}_u} Y_t|^p \right)^{1/p} =: I_1^{1/p} + I_2^{1/p}. \quad (5.1)$$

Step 2: With the assumptions of Theorem 3.7, Proposition 5.3 together with Corollary 4.7 implies

$$I_1 + \mathbb{E}^{\mathcal{F}_u} \left(\int_u^t |Z_r|^2 dr \right)^{\frac{p}{2}} \leq c_{(5.3)}^p 2^p \left(w_{p,s,u,t}^{\xi,f} \right)^p,$$

where $c_{(5.3)} > 0$ depends at most on $(T, d, p, L_y, L_z, (s_N)_N)$.

Step 3: With the assumptions of Theorem 3.7, Proposition 5.5 implies that

$$I_2 \leq c_{(5.5)}^p w_{p,s,u,t}^p,$$

where $c_{(5.5)} > 0$ depends at most on $(T, d, p, L_y, L_z, (s_N)_N)$.

Step 4: In the end we extend the result from all deterministic times $u \in [s, t]$ to all stopping times $\tau : \Omega \rightarrow [s, t]$.

The next Proposition is a conditional version of [18, Theorem 6.3]. Note that Assumption 3.5 is not needed for this result.

5.3 Proposition. Assume (C1)-(C4) for $\theta \in [0, 1]$ and $p \in [2, \infty) \cap (p_{(C4)}, \infty)$, and fix $0 \leq u < t \leq T$. Then there exists $c_{(5.3)} > 0$ depending at most on $(T, d, p, L_y, L_z, (s_N)_N)$ such that

$$\begin{aligned} & \mathbb{E}^{\mathcal{F}_u^0} \sup_{r \in [u, T]} |Y_r^{(u,t)} - Y_r|^p + \mathbb{E}^{\mathcal{F}_u^0} \left[\left(\int_u^t |Z_r|^2 dr \right)^{\frac{p}{2}} + \left(\int_u^T |Z_r^{(u,t)} - Z_r|^2 dr \right)^{\frac{p}{2}} \right] \\ & \leq c_{(5.3)}^p \mathbb{E}^{\mathcal{F}_u^0} \left(|\xi^{(u,t)} - \xi| + \int_u^T |f(r, Y_r, Z_r) - f^{(u,t)}(r, Y_r, Z_r)| \right)^p. \end{aligned}$$

Proof. Let $A^0 \in \mathcal{F}_u^0$ such that $\bar{\mathbb{P}}(A^0) > 0$. Since the σ -algebras \mathcal{F}_u^0 and $\mathcal{F}_u \otimes \{\emptyset, \Omega'\}$ differ only by null-sets, it follows that there exists $A \in \mathcal{F}_u$ with $\mathbb{P}(A) > 0$ such that $\bar{\mathbb{P}}(1_{(A \times \Omega')} = 1_{A^0}) = 1$. Now we define

$$\begin{aligned}\bar{\xi} &:= (\xi - Y_u)1_A, \\ \bar{f}(r, y, z) &:= f(r, y + Y_u, z)1_A 1_{(u, T]}(r), \\ \bar{Y}_r &:= (Y_r - Y_u)1_A 1_{(u, T]}(r), \\ \bar{Z}_r &:= Z_r 1_A 1_{(u, T]}(r).\end{aligned}$$

Note that \bar{f} is designed to satisfy for all $r \in [0, T]$ the equation

$$\bar{f}(r, \bar{Y}_r, \bar{Z}_r) = f(r, Y_r, Z_r)1_A 1_{(u, T]}(r).$$

It is straight-forward to check that since (f, Y, Z) satisfy (C1)-(C4), also $(\bar{f}, \bar{Y}, \bar{Z})$ satisfy (C1)-(C4). Moreover, $(t, \omega) \mapsto \bar{f}(t, \omega, y, z)$ is predictable for all $(y, z) \in \mathbb{R}^{1+d}$. Now we have that (\bar{Y}, \bar{Z}) is a solution to

$$\bar{Y}_t = \bar{\xi} + \int_t^T \bar{f}(r, \bar{Y}_r, \bar{Z}_r) dr - \int_t^T \bar{Z}_r dW_r.$$

Since $(\bar{f}, \bar{Y}, \bar{Z})$ satisfy conditions (C1)-(C4), and because of Lemma 3.1, it follows that they also satisfy the assumptions of [18, Theorem 6.3]. Applying [18, Theorem 6.3] with $\psi := 0$, and $\varphi := 1_{(u, t]}$ implies that there exists $c_{(6.3)} > 0$ depending at most on $(T, d, p, L_y, L_z, (s_N)_N)$ such that

$$\begin{aligned}& \left\| \sup_{r \in [u, T]} |\bar{Y}_r^{(u, t]} - \bar{Y}_r| \right\|_p + \left\| \left(\int_u^t |\bar{Z}_r|^2 dr \right)^{\frac{1}{2}} \right\|_p + \left\| \left(\int_u^T |\bar{Z}_r^{(u, t]} - \bar{Z}_r|^2 dr \right)^{\frac{1}{2}} \right\|_p \\ & \leq c_{(6.3)} \left[\left\| \bar{\xi}^{(u, t]} - \bar{\xi} \right\|_p + \left\| \int_u^T |\bar{f}^{(u, t]}(r, \bar{Y}_r, \bar{Z}_r) - \bar{f}(r, \bar{Y}_r, \bar{Z}_r)| dr \right\|_p \right].\end{aligned}$$

By definitions of $(\bar{\xi}, \bar{Y}, \bar{Z}, \bar{f})$ and using properties of the decoupling operators, in particular note that $1_{A^0}^{(u, t]} = 1_{A^0}$ since $A^0 \in \mathcal{F}_u^0$, this reads as

$$\begin{aligned}& \left\| \sup_{r \in [u, T]} |Y_r^{(u, t]} - Y_r| 1_{A^0} \right\|_p + \left\| \left(\int_u^t |Z_r|^2 dr \right)^{\frac{1}{2}} 1_{A^0} \right\|_p + \left\| \left(\int_u^T |Z_r^{(u, t]} - Z_r|^2 dr \right)^{\frac{1}{2}} 1_{A^0} \right\|_p \\ & \leq c_{(6.3)} \left[\left\| (\xi^{(u, t]} - \xi) 1_{A^0} \right\|_p + \left\| \int_u^T |f^{(u, t]}(r, Y_r, Z_r) - f(r, Y_r, Z_r)| dr 1_{A^0} \right\|_p \right],\end{aligned}$$

which immediately implies the claim. \square

Next we try to find an upper bound for $I_2 = \mathbb{E}^{\mathcal{F}_u} |Y_u - \mathbb{E}^{\mathcal{F}_u} Y_t|^p$. We accomplish this by upper bounding $\mathbb{E}^{\mathcal{F}_u} |\int_u^t f(r, Y_r, Z_r) dr|^p$. First we have a simple upper bound for the Y -term, given in terms of the data (ξ, f) .

5.4 Lemma. *Assume (C1)-(C4) for $\theta \in [0, 1]$ and $p \in [2, \infty) \cap (p_{(C4)}, \infty)$. Then we have for any $u \in [0, T]$ that*

$$\mathbb{E}^{\mathcal{F}_u} \left(\sup_{r \in [u, T]} |Y_r| + \left(\int_u^T |Z_r|^2 dr \right)^{\frac{1}{2}} \right)^p \leq c_{(5.4)}^p \mathbb{E}^{\mathcal{F}_u} \left(|\xi| + \int_u^T |f(r, 0, 0)| dr \right)^p,$$

where $c_{(5.4)} > 0$ depends at most on $(T, d, p, L_y, L_z, (s_N)_N)$.

Proof. Let $A \in \mathcal{F}_u$, and put

$$\begin{aligned}\xi^0 &= (\xi - Y_u)1_A, \\ f^0(r, y, z) &= f(r, y + Y_u, z)1_A 1_{(u, T]}(r), \\ Y_r^0 &= (Y_r - Y_u)1_A 1_{(u, T]}(r), \\ Z_r^0 &= Z_r 1_A 1_{(u, T]}(r),\end{aligned}$$

as well as $\xi^1 = 0$, $f^1(r, y, z) = 0$, $Y_r^1 = 0$, $Z_r^1 = 0$. As in the proof of Proposition 5.3, we have that (f^0, Y^0, Z^0) satisfy (C1)-(C4). This yields the assumptions of [18, Lemma 5.26], which immediately implies the claim. \square

Next we deduce the desired upper bound for I_2 .

5.5 Proposition. *Assume (C1)-(C4) for $\theta \in [0, 1]$ and $p \in [2, \infty) \cap (p_{(C4)}, \infty)$, and let $0 \leq s < t \leq T$ such that (C5) and (C6) are satisfied. Then we have for any $u \in [s, t]$ that*

$$|Y_u - \mathbb{E}^{\mathcal{F}_u} Y_t|^p \leq c_{(5.5)}^p \left[(w_{p,s,u,t}^{\xi,f})^p + \mathbb{E}^{\mathcal{F}_u} \left(\int_u^t |f(r, 0, 0)| dr \right)^p + (t-u)^p \mathbb{E}^{\mathcal{F}_u} \left(|\xi| + \int_u^T |f(r, 0, 0)| dr \right)^p \right],$$

where $c_{(5.5)} > 0$ depends at most on $(T, d, p, L_y, L_z, (s_N)_N)$.

Proof. We have directly

$$\begin{aligned}|Y_u - \mathbb{E}^{\mathcal{F}_u} Y_t|^p &= \left| \mathbb{E}^{\mathcal{F}_u} \int_u^t f(r, Y_r, Z_r) dr \right|^p \\ &\leq \mathbb{E}^{\mathcal{F}_u} \left| \int_u^t |f(r, 0, 0)| + L_y |Y_r| + L_z [1 + |Z_r|]^\theta |Z_r| dr \right|^p \\ &\leq C_p \mathbb{E}^{\mathcal{F}_u} \left[\left(\int_u^t |f(r, 0, 0)| dr \right)^p + L_y^p (t-u)^p \sup_{r \in [u, t]} |Y_r|^p \right. \\ &\quad \left. + L_z^p \left(\int_u^t |Z_r| dr \right)^p + L_z^p \left(\int_u^t |Z_r|^{1+\theta} dr \right)^p \right].\end{aligned}$$

Lemma 5.4 gives us an upper bound for the second term. For the third term we may apply Proposition 5.3 and Assumption 5.1 with Remark 5.2 to deduce

$$\mathbb{E}^{\mathcal{F}_u} \left(\int_u^t |Z_r| dr \right)^p \leq (t-u)^{p/2} \mathbb{E}^{\mathcal{F}_u} \left(\int_u^t |Z_r|^2 dr \right)^{p/2} \leq (2c_{(5.3)})^p (t-u)^{p/2} (w_{p,s,u,t}^{\xi,f})^p.$$

For the last term we use also Proposition 2.10 and Assumption (C4) to deduce

$$\begin{aligned}\mathbb{E}^{\mathcal{F}_u} \left(\int_u^t |Z_r|^{1+\theta} dr \right)^p &\leq c_p \sup_{r \in [u, t]} \left\| \mathbb{E}^{\mathcal{F}_r} \int_r^t |Z_v|^{2\theta} dv \right\|_\infty^{\frac{p}{2}} \mathbb{E}^{\mathcal{F}_u} \left(\int_u^t |Z_r|^2 dr \right)^{\frac{p}{2}} \\ &\leq c_p \|\chi_{(u,t]} |Z|^\theta\|_{\text{BMO}(\mathcal{S}_2)}^p (2c_{(5.3)})^p (w_{p,s,u,t}^{\xi,f})^p.\end{aligned}$$

\square

Proof of Theorem 3.7:

Assume that (C1)-(C6) hold for $\theta \in [0, 1]$, $p \in [2, \infty) \cap (p_{(C4)}, \infty)$, and $0 \leq s < t \leq T$.

(i) It follows from Propositions 5.3 and 5.5 that there exists a constant $C > 0$ depending at most on $(T, d, p, L_y, L_z, (s_N)_N)$, such that for all $u \in [s, t]$

$$\mathbb{E}^{\mathcal{F}_u} |Y_t - Y_u|^p \leq C w_{p,s,u,t}^p.$$

Since $((w_{p,s,u,t}^\xi)^p)_{u \in [s,t]}$ and $((w_{p,s,u,t}^f)^p)_{u \in [s,t]}$ are supermartingales, it follows that $(w_{p,s,u,t}^p)_{u \in [s,t]}$ as well is a supermartingale. Applying [20, Theorem 3.13, page 16] on $\mathbb{E}^{\mathcal{F}_u} \int_u^t |f(r, 0, 0)| dr$, we deduce that $(w_{p,s,u,t}^p)_{u \in [s,t]}$ has a càdlàg modification, to which we will switch without changing the notation. Applying Lemma 7.2 with $\alpha_u := |Y_t - Y_u|^p$ and $w_u := C w_{p,s,u,t}^p$ implies the claim.

(ii) It follows from Proposition 5.3 that there exists $C > 0$ depending at most on $(T, d, p, L_y, L_z, (s_N)_N)$, such that for all $u \in [s, t]$ we have

$$\mathbb{E}^{\mathcal{F}_u} \left(\int_u^t |Z_r|^2 dr \right)^{\frac{p}{2}} \leq C^p (w_{p,s,u,t}^{\xi,f})^p.$$

Hence, the claim follows by applying Lemma 7.2 with $\alpha_u := \left(\int_u^t |Z_r|^2 dr \right)^{\frac{p}{2}}$ and $w_u := C^p (w_{p,s,u,t}^{\xi,f})^p$. \square

5.2. Proof of Example 3.11

We start with an inequality, which proof is the same as that of [16, Theorem 2.5]. To do so, recall that $(\overline{\mathcal{F}}_t)_{t \in [0, T]}$ is the natural augmented filtration of (W, W') .

5.6 Lemma. Assume $(A_{b,\sigma})$, let $0 \leq s < t \leq T$ and $p \in [2, \infty)$. Then there exists $C_{(5.6)} > 0$ depending at most on $(T, d, p, L_{b,\sigma}, K_{b,\sigma})$ such that

$$\mathbb{E} \left[\sup_{r \in [s, T]} |X_r^{(s,t)} - X_r|^p \mid \overline{\mathcal{F}}_s \right] \leq C_{(5.6)}^p (t-s)^{p/2} \left(1 + \mathbb{E} \left[\sup_{r \in [s, t]} |X_r|^p \mid \overline{\mathcal{F}}_s \right] \right).$$

If additionally (A_σ) holds, then there exists $D_{(5.6)} > 0$ depending at most on $(T, d, p, L_{b,\sigma}, K_\sigma)$ such that

$$\mathbb{E} \left[\sup_{r \in [s, T]} |X_r^{(s,t)} - X_r|^p \mid \overline{\mathcal{F}}_s \right] \leq D_{(5.6)}^p (t-s)^{p/2}.$$

Proof. Using Proposition 4.4 we have

$$\begin{aligned} X_r^{(s,t)} - X_r &= \int_s^r (b(u, X_u^{(s,t)}) - b(u, X_u)) du + \int_s^r \sigma(u, X_u^{(s,t)}) 1_{(s,t]}(u) dW_u^1 - \int_s^r \sigma(u, X_u) 1_{(s,t]}(u) dW_u^0 \\ &\quad + \int_s^r (\sigma(u, X_u^{(s,t)}) - \sigma(u, X_u)) (1 - 1_{(s,t]}(u)) dW_u^0 \end{aligned}$$

for all $r \in [s, T]$, $\overline{\mathbb{P}}$ -a.s. Next we let $A \in \overline{\mathcal{F}}_s$ with $\overline{\mathbb{P}}(A) > 0$, and define $g : [s, T] \rightarrow [0, \infty)$ by

$$g(v) := \mathbb{E} \left(\sup_{s \leq r \leq v} |X_r^{(s,t)} - X_r|^p 1_A \right) = \int_A \sup_{s \leq r \leq v} |X_r^{(s,t)} - X_r|^p d\overline{\mathbb{P}}.$$

Using basic inequalities, we have for all $v \in [s, T]$ that

$$g(v) = \mathbb{E} \sup_{s \leq r \leq v} \left| \int_s^r (b(u, X_u^{(s,t)}) - b(u, X_u)) 1_A du + \int_s^r \sigma(u, X_u^{(s,t)}) 1_{(s,t]}(u) dW_u^1 \right|$$

$$\begin{aligned}
 & - \int_s^r \sigma(u, X_u) 1_{(s,t]}(u) 1_A dW_u^0 + \int_s^r (\sigma(u, X_u^{(s,t]}) - \sigma(u, X_u)) 1_{(s,t]}(u) 1_A dW_u^0 \Big|^p \\
 & \leq C(t-s)^{p/2} \int_A 1 + \sup_{s \leq r \leq t} |X_r|^p d\bar{\mathbb{P}} + C \int_s^v g(u) du,
 \end{aligned} \tag{5.2}$$

where the constant C depends at most on $(T, d, p, L_{b,\sigma}, K_{b,\sigma})$. Then it follows from Gronwall's lemma that

$$\int_A \sup_{s \leq r \leq T} |X_r^{(s,t]} - X_r|^p d\bar{\mathbb{P}} = g(T) \leq C_{(5.6)}^p (t-s)^{p/2} \int_A 1 + \sup_{s \leq r \leq t} |X_r|^p d\bar{\mathbb{P}},$$

where the constant $C_{(5.6)}$ depends at most on $(T, d, p, L_{b,\sigma}, K_{b,\sigma})$. If (A_σ) holds, then we can deduce from Equation (5.2) that

$$g(v) \leq C(t-s)^{p/2} \bar{\mathbb{P}}(A) + C \int_s^v g(u) du,$$

where the constant C now depends at most on $(T, d, p, L_{b,\sigma}, K_{b,\sigma})$. The result again follows from Gronwall's lemma. \square

Proof of Example 3.11:

(C1)-(C4): Follow from [7, Theorem 4.2], since (A_h) implies that (C1) holds with $\theta = 0$, $(A_{b,\sigma})$ together with (A_g) implies $\mathbb{E}|g(X_T)|^p < \infty$, and $(A_{b,\sigma})$ together with (A_h) implies $\mathbb{E} \left(\int_0^T |h(r, X_r, 0, 0)| dr \right)^p < \infty$.

(C5): Let $0 \leq s \leq u \leq t \leq T$. Using Proposition 4.3(ii) we have that

$$(g(X_T))^{(u,t]} = g(X_T^{(u,t]}),$$

and Proposition 4.4 implies that $X^{(u,t]}$ is the solution of

$$X_r^{(u,t]} = x + \int_0^r b(v, X_v^{(u,t]}) dv + \int_0^r \sigma(v, X_v^{(u,t]}) dW_v^{(u,t]}, \quad r \in [0, T].$$

It follows from (A_g) that

$$\mathbb{E}^{\mathcal{F}_u^0} |g(X_T) - g(X_T^{(u,t]})|^p \leq L_g^p \mathbb{E}^{\mathcal{F}_u^0} |X_T - X_T^{(u,t]}|^p.$$

Finally, Lemma 5.6 implies

$$\mathbb{E}^{\mathcal{F}_u^0} |X_T - X_T^{(u,t]}|^p \leq C_{(5.6)}^p (t-u)^{p/2} \left(1 + \mathbb{E} \left[\sup_{r \in [u,t]} |X_r|^p \mid \mathcal{F}_u^0 \right] \right),$$

and if (A_σ) holds, then Lemma 5.6 implies

$$\mathbb{E}^{\mathcal{F}_u^0} |X_T - X_T^{(u,t]}|^p \leq D_{(5.6)}^p (t-u)^{p/2}.$$

(C6): Let $0 \leq s \leq u \leq t \leq T$. We notice that Proposition 4.3 implies

$$(h(r, X_r, y, z))^{(u,t]} = h(r, X_r^{(u,t]}, y, z).$$

The result again follows from Lemma 5.6, since (A_h) implies

$$\mathbb{E}^{\mathcal{F}_u^0} \left(\int_0^T \sup_{y,z} |h(r, X_r, y, z) - h(r, X_r^{(u,t]}, y, z)| dr \right)^p \leq \mathbb{E}^{\mathcal{F}_u^0} \left(\int_0^T L_h |X_r^{(u,t]} - X_r| dr \right)^p.$$

For all $u \in [s, t]$ we let

$$w_u := w_{p,s,u,t} := (t-u)^{p/2} \left(1 + \mathbb{E} \left[\sup_{r \in [u,t]} |X_r|^p \mid \mathcal{F}_u \right] \right),$$

and get that the process $(w_u)_{u \in [s,t]}$ is a supermartingale. Since $u \mapsto \mathbb{E}w_u$ is continuous, there exists a càdlàg modification of the process $(w_r)_{r \in [s,t]}$. This modification is a càdlàg supermartingale, and for any fixed $u \in [s, t]$ its canonical extension coincides \mathbb{P} -a.s. with

$$(t-u)^{p/2} \left(1 + \mathbb{E} \left[\sup_{r \in [u,t]} |X_r|^p \mid \mathcal{F}_u^0 \right] \right).$$

Hence, there exists $C > 0$ depending at most on $(T, d, p, L_g, L_h, L_{b,\sigma}, K_{b,\sigma})$ such that Assumption 5.1 holds for all $0 \leq s < t \leq T$ with

$$\left(\tilde{w}_{p,s,u,t}^f \right)^p = \left(\tilde{w}_{p,s,u,t}^\xi \right)^p := C^p (t-u)^{p/2} \left(1 + \mathbb{E} \left[\sup_{r \in [u,t]} |X_r|^p \mid \mathcal{F}_u^0 \right] \right).$$

If additionally (A_σ) holds, then there exists $D > 0$ depending at most on $(T, d, p, L_g, L_h, L_{b,\sigma}, K_\sigma)$ such that we may choose

$$\tilde{w}_{p,s,u,t}^f = \tilde{w}_{p,s,u,t}^\xi = D(t-u)^{1/2}.$$

□

6. Some Applications

In this section we discuss some applications of the tail estimates obtained in Theorem 3.13. We can use them in two different ways: Firstly, we can exploit the tail estimates (Sections 6.1 and 6.2), secondly we may exploit the fact that we can control *all conditional moments* which might allow us for a change of the underlying measure (Section 6.3).

6.1. Uniform spline approximation of the process Y

To get a path-dependent approximation of the process $Y = (Y_t)_{t \in [0,T]}$ based on a method that provides approximations \widehat{Y}_{t_i} of Y_{t_i} for some deterministic time-net $\pi = (t_i)_{i=0}^n$, $0 = t_0 < \dots < t_n = T$, one can consider a linear spline

$$\widehat{Y}_t^\pi := (1-\theta)\widehat{Y}_{t_{i-1}} + \theta\widehat{Y}_{t_i} \quad \text{for } t \in I_i^\pi := [t_{i-1}, t_i] \text{ with } t = (1-\theta)t_{i-1} + \theta t_i.$$

We get that

$$\sup_{i=0, \dots, n} |\widehat{Y}_{t_i}^\pi - Y_{t_i}| \leq \|\widehat{Y}^\pi - Y\|_{C([0,T])} \leq \sup_{i=0, \dots, n} |\widehat{Y}_{t_i}^\pi - Y_{t_i}^\pi| + \|Y^\pi - Y\|_{C([0,T])},$$

where, as above,

$$Y_t^\pi := (1-\theta)Y_{t_{i-1}} + \theta Y_{t_i} \quad \text{for } t \in I_i^\pi \text{ with } t = (1-\theta)t_{i-1} + \theta t_i.$$

The process Y^π is a piece-wise linear and continuous process, but fails to be adapted in general. In this section we provide in Propositions 6.4 and 6.5 below large deviation type estimates for $\|Y^\pi - Y\|_{C([0,T])}$. We start with the following simple observation that links the distribution of the spline to our results:

6.1 Lemma. Assume that there is a $\lambda_0 \geq 0$ and a function $G : [\lambda_0, \infty) \times [0, T] \rightarrow [0, \infty)$, non-decreasing in its second component, such that

$$\mathbb{P}\left(\sup_{u \in [s, t]} \frac{|Y_u - Y_s|}{\sqrt{t-s}} > \lambda\right) \leq G(\lambda, t-s) \quad \text{for } \lambda \geq \lambda_0.$$

Then one has that

$$\mathbb{P}\left(\sup_{t \in [0, T]} |Y_t - Y_t^\pi| > \lambda\right) \leq nG\left(\frac{\lambda}{2\sqrt{|\pi|}}, |\pi|\right) \quad \text{for } \lambda \geq 2\sqrt{|\pi|}\lambda_0.$$

Proof. We have that

$$\begin{aligned} \sup_{t \in [0, T]} |Y_t - Y_t^\pi| &= \sup_{i=1, \dots, n} \sup_{\theta \in [0, 1]} |Y_{(1-\theta)t_{i-1} + \theta t_i} - ((1-\theta)Y_{t_{i-1}} + \theta Y_{t_i})| \\ &\leq \sup_{i=1, \dots, n} \max \left\{ \sup_{t \in I_i^\pi} |Y_t - Y_{t_{i-1}}|, \sup_{t \in I_i^\pi} |Y_t - Y_{t_i}| \right\} \\ &\leq \sup_{i=1, \dots, n} \left(\sup_{t \in I_i^\pi} |Y_t - Y_{t_{i-1}}| + |Y_{t_i} - Y_{t_{i-1}}| \right) \\ &\leq 2 \sup_{i=1, \dots, n} \sup_{t \in I_i^\pi} |Y_t - Y_{t_{i-1}}|. \end{aligned}$$

For $\lambda \geq 2\lambda_0$ this implies our statement because

$$\begin{aligned} \mathbb{P}\left(\sup_{t \in [0, T]} |Y_t - Y_t^\pi| > \sqrt{|\pi|}\lambda\right) &\leq \sum_{i=1}^n \mathbb{P}\left(2 \sup_{t \in I_i^\pi} |Y_t - Y_{t_{i-1}}| > \sqrt{|\pi|}\lambda\right) \\ &\leq \sum_{i=1}^n \mathbb{P}\left(\sup_{t \in I_i^\pi} |Y_t - Y_{t_{i-1}}| > \frac{\sqrt{|\pi|}\lambda}{2}\right) \\ &\leq \sum_{i=1}^n G\left(\frac{\lambda}{2}, |I_i^\pi|\right) \\ &\leq nG\left(\frac{\lambda}{2}, |\pi|\right). \quad \square \end{aligned}$$

In order to apply Theorem 3.13 we let, for $\lambda > 0$ and $0 \leq s < t \leq T$,

$$\begin{aligned} F(\lambda) &:= \mathbb{P}\left(\sup_{u \in [0, T]} |X_u| > \lambda\right), \\ G_\ell(\lambda) &:= \inf \left\{ e^{-\mu} + F\left(\sqrt{v^2 - 1}\right) : \lambda = \mu v \text{ with } \mu > 0, v > 1 \right\}, \\ G_b(\lambda, t-s) &:= \inf \left\{ e^{-\mu} + F\left(\sqrt{\frac{v^2 - 1}{t-s}}\right) : \lambda = \mu v \text{ with } \mu > 0, v > 1 \right\}. \end{aligned}$$

The subscript ℓ stands for a linear growth of σ , the subscript b for a bounded σ . For the function F we get the following upper bounds:

6.2 Lemma.

- (i) Under the condition $(A_{b,\sigma})$ there exist $\alpha > 0$ and $\lambda_0 \geq 1$ depending at most on (x, b, σ, T) such that, for $\lambda \geq \lambda_0$,

$$F(\lambda) \leq e^{-\alpha(\log \lambda)^2}.$$

- (ii) Under the conditions $(A_{b,\sigma})$ and (A_σ) there is a $c > 0$ depending at most on (x, b, σ, T) such that, for $\lambda \geq 0$,

$$F(\lambda) \leq ce^{-(\lambda/c)^2}.$$

Proof. For $p \in [2, \infty)$ one has the estimates

$$\left\| \sup_{t \in [0, T]} |X_t| \right\|_p \leq e^{cp} \quad \text{and} \quad \left\| \sup_{t \in [0, T]} |X_t| \right\|_p \leq c\sqrt{p}$$

under $(A_{b,\sigma})$ and $(A_{b,\sigma}, A_\sigma)$, respectively, for constants $c > 0$ depending at most on (x, b, σ, T) . Both estimates are known. They can be proved by the standard Gronwall argument (cf. [1, Lemma A.2]) but one has to use the estimate $\beta_p \leq c\sqrt{p}$ for $p \in [2, \infty)$ from Proposition 2.5.

- (i) For all $\lambda > 0$,

$$\mathbb{P}(X_T^* > \lambda) \leq \frac{1}{\lambda^p} \mathbb{E}|X_T^*|^p \leq \frac{1}{\lambda^p} e^{cp^2}.$$

We set $\lambda_0 := e^{4c}$ and get for $\lambda \geq \lambda_0$ a $p \in [2, \infty)$ with $p = \frac{\log \lambda}{2c}$, so that $\frac{1}{\lambda^p} e^{cp^2} = e^{-\frac{(\log \lambda)^2}{4c}}$.

- (ii) Again, for all $\lambda > 0$,

$$\mathbb{P}(X_T^* > \lambda) \leq \frac{1}{\lambda^p} \mathbb{E}|X_T^*|^p \leq \frac{1}{\lambda^p} c^p p^{\frac{p}{2}}.$$

Assume $\lambda \geq \sqrt{2}ce$ and set $p := (\lambda/(ce))^2 \in [2, \infty)$. Then

$$\mathbb{P}(X_T^* > \lambda) \leq e^{-\lambda^2/(ce)^2}.$$

Consequently, $\mathbb{P}(X_T^* > \lambda) \leq e^{-\lambda^2/(ce)^2}$ for all $\lambda \geq 0$. □

We derive the following bounds for G_ℓ and G_b :

6.3 Lemma.

- (i) Under the condition $(A_{b,\sigma})$ there exist $\alpha > 0$ and $\lambda_0 \geq 1$ depending at most on (x, b, σ, T) such that, for $\lambda \geq \lambda_0$,

$$G_\ell(\lambda) \leq e^{-\alpha(\log \lambda)^2}.$$

- (ii) Under the conditions $(A_{b,\sigma})$ and (A_σ) there is a $c > 0$ depending at most on (x, b, σ, T) such that, for $0 \leq s < t \leq T$,

$$G_b(\lambda, t-s) \leq \begin{cases} ce^{-\frac{1}{c}\lambda} & 0 < \lambda \leq \frac{1}{t-s} \\ ce^{-\frac{1}{c}\lambda^{\frac{2}{3}}(t-s)^{-\frac{1}{3}}} & \lambda > \frac{1}{t-s} \end{cases}.$$

Under the conditions $(A_{b,\sigma}, A_\sigma)$ we let $\lambda_0 := 0$ and $G_b(0, t-s) := \lim_{\lambda \downarrow 0} G_b(\lambda, t-s)$ so that $G_b(0, t-s) \leq c$.

Proof of Lemma 6.3. For both cases we can replace $\nu > 1$ in the definitions of G_ℓ and G_b by $\nu \geq \sqrt{4/3}$ to replace the term $\sqrt{\nu^2 - 1}$ by $\nu/2$ to simplify the computation.

(i) We use the decomposition $\lambda = \mu\nu = \sqrt{\lambda}\sqrt{\lambda}$ and Lemma 6.2(i) (where $\alpha, \lambda_0 > 0$ might change).

(ii) In the case $\lambda \leq \frac{1}{t-s}$ we use the decomposition $\mu = \sqrt{3/4}\lambda$ and $\nu = \sqrt{4/3}$, and in the case $\lambda > \frac{1}{t-s}$ we use $\mu = \sqrt{3/4}\lambda^{2/3}(t-s)^{-1/3}$ and $\nu = \sqrt{4/3}\lambda^{1/3}(t-s)^{1/3}$. Then we apply Lemma 6.2(ii). \square

From Theorem 3.13 we know that

$$\mathbb{P}\left(\sup_{u \in [s,t]} \frac{|Y_u - Y_s|}{\sqrt{t-s}} > A_{(3.13)}\lambda\right) \leq A \begin{cases} G_\ell(\lambda) & : (A_{b,\sigma}) \\ G_b(\lambda, t-s) & : (A_{b,\sigma}, A_\sigma) \end{cases} \quad (6.1)$$

for $\lambda \geq \lambda_0$ and $0 \leq s < t \leq T$. Here $A := c_0 \vee e$ with $c_0 > 0$ taken from Theorem 3.13, $A_{(3.13)} := c_0 c_{(3.12)}$ in the case $(A_{b,\sigma})$, and $A_{(3.13)} := c_0 d_{(3.12)}$ in the case $(A_{b,\sigma}, A_\sigma)$. To provide the large deviation type inequalities, we let $\pi_n = (iT/n)_{i=0}^n$ be the equidistant net with $n+1$ knots and denote $Y^n := Y^{\pi_n}$.

6.4 Proposition. *Under the condition $(A_{b,\sigma})$ one has for $n \geq 2$ and $\lambda \geq \lambda_0$ that*

$$\mathbb{P}(\|Y - Y^n\|_{C([0,T])} > \alpha_n \lambda) \leq 2Ae^{-\alpha(\log \lambda)^2}.$$

where $\alpha_n := 2\sqrt{T}A_{(3.13)}n^{-1/2}e^{\sqrt{\frac{1}{\alpha}\log \frac{n}{2}}}$ and α, λ_0 are taken from Lemma 6.3(i).

Proof. For $n \geq 2$ and $\lambda \geq \lambda_0$ we get from Lemmas 6.1 and 6.3(i) that

$$\begin{aligned} \mathbb{P}(\|Y - Y^n\|_{C([0,T])} > \alpha_n \lambda) &\leq nAG_\ell\left(\frac{\alpha_n \lambda \sqrt{n}}{2\sqrt{T}A_{(3.13)}}\right) \\ &= nAG_\ell\left(\lambda e^{\sqrt{\frac{1}{\alpha}\log \frac{n}{2}}}\right) \\ &\leq nAe^{-\alpha\left(\log\left(\lambda e^{\sqrt{\frac{1}{\alpha}\log \frac{n}{2}}}\right)\right)^2} \\ &= nAe^{-\alpha\left(\log \lambda + \sqrt{\frac{1}{\alpha}\log \frac{n}{2}}\right)^2} \\ &\leq nAe^{-\alpha(\log \lambda)^2 - \log \frac{n}{2}} \\ &= 2Ae^{-\alpha(\log \lambda)^2}. \end{aligned}$$

\square

Using (6.1) for $(A_{b,\sigma}, A_\sigma)$ gives the following large deviation estimate:

6.5 Proposition. *Under the conditions $(A_{b,\sigma})$ and (A_σ) there is a constant $c > 0$ such that*

$$\begin{aligned} \limsup_{\lambda \rightarrow \infty} \frac{\log \mathbb{P}(\|Y - Y^n\|_{C([0,T])} > \lambda)}{\lambda^{\frac{2}{3}}} &\leq -cn^{\frac{2}{3}} \quad \text{for } n \geq 1, \\ \limsup_{n \rightarrow \infty} \frac{\log \mathbb{P}(\|Y - Y^n\|_{C([0,T])} > \lambda)}{n^{\frac{1}{2}}} &\leq -c\lambda \quad \text{for } \lambda > 0. \end{aligned}$$

Proof. For the first inequality we use the case $\lambda > \frac{1}{t-s} = \frac{n}{T}$ in inequality (6.1), for the second inequality the case $0 < \lambda \leq \frac{1}{t-s} = \frac{n}{T}$ in inequality (6.1). \square

6.2. Confidence interval for direct simulation

Assume that we are interested in the computation of $E(Y_t - Y_s)$ for fixed $0 \leq s < t \leq T$ and that we can simulate independent copies D_1, \dots, D_n of $Y_t - Y_s$. Below we give an estimate on the confidence interval that is based on our tail-estimates. We start with a general lemma, that should be folklore. To this end, assume an i.i.d. sequence of random variables $D_1, D_2, \dots : \widehat{\Omega} \rightarrow \mathbb{R}$ such that $D_1 \in \bigcap_{p \in (0, \infty)} L_p(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}})$, and let

$$S_n := \frac{1}{n}(D_1 + \dots + D_n) \quad \text{and} \quad \mu := \mathbb{E}D_1.$$

6.6 Lemma. *For $\varepsilon > 0$ one has*

$$\widehat{\mathbb{P}}(|S_n - \mu| > \varepsilon) \leq \inf_{p \in [2, \infty)} \left(\frac{2(p-1)}{\sqrt{n}\varepsilon} \|D_1\|_p \right)^p.$$

Proof. For $p \in [2, \infty)$ we have

$$\|S_n - \mu\|_p = \frac{1}{n} \left\| \sum_{i=1}^n [D_i - \mu] \right\|_p \leq \frac{\beta_p}{n} \left\| \left(\sum_{i=1}^n |D_i - \mu|^2 \right)^{\frac{1}{2}} \right\|_p \leq \frac{\beta_p}{n} \left(\sum_{i=1}^n \|D_i - \mu\|_p^2 \right)^{\frac{1}{2}} = \frac{\beta_p}{\sqrt{n}} \|D_1 - \mu\|_p$$

where from [10, Theorem 3.3] we know that we can take $\beta_p = p - 1$. Therefore, for $\varepsilon > 0$,

$$\widehat{\mathbb{P}}(|S_n - \mu| > \varepsilon) \leq \frac{1}{\varepsilon^p} \|S_n - \mu\|_p^p \leq \frac{1}{\varepsilon^p} \left(\frac{p-1}{\sqrt{n}} \right)^p \|D_1 - \mu\|_p^p \leq \left(\frac{2(p-1)}{\sqrt{n}\varepsilon} \right)^p \|D_1\|_p^p.$$

□

Now let us assume that condition $(A_{b,\sigma})$ is satisfied and fix $0 \leq s < t \leq T$. Let S_n be a direct simulation of $Y_t - Y_s$. From Lemma 6.3(i) we can deduce

$$\left\| \frac{Y_t - Y_s}{\sqrt{t-s}} \right\|_p \leq e^{cp}$$

for some $c > 0$ and all $p \in [2, \infty)$. By Lemma 6.6,

$$\widehat{\mathbb{P}}(|S_n - \mu| > \varepsilon) \leq \left(\frac{2(p-1)}{\sqrt{n}\varepsilon} \sqrt{t-s} e^{cp} \right)^p \leq \left(\alpha \frac{\sqrt{t-s}}{\sqrt{n}\varepsilon} e^{\alpha p} \right)^p$$

for some $\alpha = \alpha(c) > 0$ and all $p \in [2, \infty)$. Define $\psi(\delta) := \inf_{p \in [2, \infty)} (\alpha \delta e^{\alpha p})^p$ for $\delta > 0$. Then

$$\widehat{\mathbb{P}}(|S_n - \mu| > \varepsilon) \leq \psi \left(\frac{\sqrt{t-s}}{\sqrt{n}\varepsilon} \right).$$

It is not difficult to check that

$$\lim_{\delta \downarrow 0} \frac{\psi(\delta)}{\delta^M} = 0 \quad \text{for all } M > 0$$

(consider $\delta \in (0, e^{-4\alpha})$ and choose $p \in [2, \infty)$ with $\delta = e^{-2\alpha p}$ so that $\psi(\delta) \leq (\alpha^2 \delta)^{\frac{p}{2}}$). For example, this implies

$$\lim_{n \rightarrow \infty} n^M \psi \left(\frac{\sqrt{t-s}}{\sqrt{n}\varepsilon} \right) = 0 \quad \text{for all } M > 0.$$

6.3. Change of measure

We describe a consequence of the BMO-estimates with respect to a change of the underlying measure. Let \mathbb{Q} be a probability measure that is absolutely continuous with respect to \mathbb{P} and such that for $L := d\mathbb{Q}/d\mathbb{P} > 0$ there are $c > 0$ and $\nu \in (1, \infty)$ such that

$$\mathbb{E} \left[L^\nu \mid \mathcal{F}_\tau \right] \leq c^\nu \left[\mathbb{E} \left[L \mid \mathcal{F}_\tau \right] \right]^\nu,$$

for all stopping times $\tau : \Omega \rightarrow [0, T]$ (i.e. \mathbb{Q} satisfies a reverse Hölder inequality with exponent ν , cf. Definition 2.7). Assume a positive càdlàg and adapted process $\Phi = (\Phi_t)_{t \in [0, T]}$, $p \in (0, \infty)$, and a continuous and adapted process $A = (A_t)_{t \in [0, T]}$ with $A_0 \equiv 0$ with $\|A\|_{\text{BMO}_p^\Phi} < \infty$ (see equation (1.2)). Let $\tau : \Omega \rightarrow [0, T]$ be a stopping time, $B \in \mathcal{F}_\tau$ be of positive measure, and $\varepsilon, \nu > 0$. If \mathbb{Q}_B is the normalized restriction of \mathbb{Q} to B and $1 = \frac{1}{u} + \frac{1}{v}$, then

$$\begin{aligned} & \mathbb{Q}_B(|A_T - A_\tau| > (1 + \varepsilon)\nu) \\ & \leq \mathbb{Q}_B(|A_T - A_\tau| > (1 + \varepsilon)\Phi_\tau) + \mathbb{Q}_B(\Phi_\tau > \nu) \\ & = \frac{1}{\mathbb{Q}(B)} \int_B \mathbb{1}_{\{|A_T - A_\tau| > (1 + \varepsilon)\Phi_\tau\}} L d\mathbb{P} + \mathbb{Q}_B(\Phi_\tau > \nu) \\ & = \frac{1}{\mathbb{Q}(B)} \int_B \mathbb{E} \left[\mathbb{1}_{\{|A_T - A_\tau| > (1 + \varepsilon)\Phi_\tau\}} L \mid \mathcal{F}_\tau \right] d\mathbb{P} + \mathbb{Q}_B(\Phi_\tau > \nu) \\ & \leq \frac{1}{\mathbb{Q}(B)} \int_B \mathbb{E} \left[\mathbb{1}_{\{|A_T - A_\tau| > (1 + \varepsilon)\Phi_\tau\}} \mid \mathcal{F}_\tau \right]^{\frac{1}{u}} \mathbb{E} \left[L^\nu \mid \mathcal{F}_\tau \right]^{\frac{1}{v}} d\mathbb{P} + \mathbb{Q}_B(\Phi_\tau > \nu) \\ & \leq \frac{c}{\mathbb{Q}(B)} \int_B \mathbb{E} \left[\mathbb{1}_{\{|A_T - A_\tau| > (1 + \varepsilon)\Phi_\tau\}} \mid \mathcal{F}_\tau \right]^{\frac{1}{u}} \mathbb{E}[L \mid \mathcal{F}_\tau] d\mathbb{P} + \mathbb{Q}_B(\Phi_\tau > \nu) \\ & = c \int_B \mathbb{E} \left[\mathbb{1}_{\{|A_T - A_\tau| > (1 + \varepsilon)\Phi_\tau\}} \mid \mathcal{F}_\tau \right]^{\frac{1}{u}} d\mathbb{Q}_B + \mathbb{Q}_B(\Phi_\tau > \nu) \\ & \leq c \int_B \left[\frac{1}{(1 + \varepsilon)^p} \|A\|_{\text{BMO}_p^\Phi}^p \right]^{\frac{1}{u}} d\mathbb{Q}_B + \mathbb{Q}_B(\Phi_\tau > \nu) \\ & = \frac{c}{(1 + \varepsilon)^{\frac{p}{u}}} \|A\|_{\text{BMO}_p^\Phi}^{\frac{p}{u}} + \mathbb{Q}_B(\Phi_\tau > \nu) \\ & \leq \frac{c}{(1 + \varepsilon)^{\frac{p}{u}}} \|A\|_{\text{BMO}_p^\Phi}^{\frac{p}{u}} + \mathbb{Q}_B \left(\sup_{u \in [\tau, T]} \Phi_u > \nu \right). \end{aligned}$$

As a consequence we can apply Theorem 9.1, but now for the measure \mathbb{Q} instead of \mathbb{P} . Let us come back to our setting and recall the inequality (1.3), i.e.

$$\|(Y_t - Y_0)_{t \in [0, T]}\|_{\text{BMO}_p^\Phi} \leq C(3.7).$$

So we can apply this change of measure technique in our context. A careful investigation of local settings (i.e. the consideration of fixed general sub-intervals $[s, t] \subset [0, T]$) is not yet done.

6.4. Outlook

The methodology to use weighted BMO spaces in stochastic problems, in order to replace L_p spaces, is exploited in the context of approximation problems for stochastic integrals in [17] and in the context of variational problems for BSDEs in this article. The natural question is, to which other problems this general

technique might be applied. A natural candidate for such a problem would be the investigation of existing approximation schemes for BSDEs from the literature (for example, [6],[5], [30], [29]). It might be that the partial backward structure of these schemes helps to apply weighted BMO techniques where one could use existing L_2 results.

7. Appendix A: General tools

The following lemmas have been used before, Lemma 7.1 to justify assumption (C6), and Lemma 7.2 in the proof of Theorem 3.7.

7.1 Lemma. *Assume that \mathbb{X} is locally σ -compact. Let $(f(x))_{x \in \mathbb{X}}$ be a continuous stochastic process defined on a probability space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$, such that $\mathbb{E} \sup_{x \in K_n} |f(x)| < \infty$ for all $n \in \mathbb{N}$. If $\mathcal{G} \subseteq \hat{\mathcal{F}}$ is a σ -algebra, then there exists a unique¹³ continuous stochastic process $((\mathbb{E}^{\mathcal{G}} f)(x))_{x \in \mathbb{X}} := (g(x))_{x \in \mathbb{X}}$ such that $\hat{\mathbb{P}}(\mathbb{E}^{\mathcal{G}}(f(x)) = g(x)) = 1$ for all $x \in \mathbb{X}$, and such that $g(x)$ is \mathcal{G} -measurable for every $x \in \mathbb{X}$.*

Proof. (i) Let K be one of the sets K_n as in Definition 2.2, and consider f as the Banach-space valued random variable $f : \hat{\Omega} \rightarrow C(K)$, where $C(K)$ is the space of continuous functions on K equipped with the sup-norm. This space is separable, so that applying [14, Theorem V.1.4] and properties of the Bochner integral we find a $g : \hat{\Omega} \rightarrow C(K)$ with the required properties.

(ii) Defining $(g^{K_n}(x))_{x \in K_n}$ and $(g^{K_{n+1}}(x))_{x \in K_{n+1}}$ as in step (i), we have that g^{K_n} and $g^{K_{n+1}}$ are indistinguishable in K_n . Hence, we can consistently define one process in $\bigcup_{n=1}^{\infty} K_n = \mathbb{X}$. \square

7.2 Lemma. *Let $0 \leq s < t \leq T$, and assume that $(\alpha_u)_{u \in [s,t]}$ is a process with càdlàg paths, and such that $\mathbb{E} \sup_{r \in [s,t]} |\alpha_r| < \infty$. If for all $u \in [s,t]$ we have $\mathbb{E}^{\mathcal{F}_u} |\alpha_u| \leq w_u$, where $(w_u)_{u \in [s,t]}$ is a supermartingale with càdlàg paths, then $\mathbb{E}^{\mathcal{F}_\tau} |\alpha_\tau| \leq w_\tau$ holds for all stopping times $\tau : \Omega \rightarrow [s,t]$.*

Proof. (i) Assume that $\tau : \Omega \rightarrow \{s_1, \dots, s_n\}$ is a stopping time for some $n \in \mathbb{N}$, $s \leq s_1 \leq \dots \leq s_n \leq t$. We have for all $i = 1, \dots, n$ that $\mathbb{E}^{\mathcal{F}_{s_i}} |\alpha_{s_i}| \leq w_{s_i}$. Now we have for any $A \in \mathcal{F}_\tau$ that

$$\int_A |\alpha_\tau| d\mathbb{P} = \sum_{i=1}^n \int_{A \cap \{\tau=s_i\}} |\alpha_{s_i}| d\mathbb{P} \leq \sum_{i=1}^n \int_{A \cap \{\tau=s_i\}} w_{s_i} d\mathbb{P} = \int_A w_\tau d\mathbb{P}.$$

(ii) Let $\tau : \Omega \rightarrow [s,t]$ be a stopping time, and let $(\tau_n)_{n \in \mathbb{N}}$ be a sequence of stopping times such that $\tau_n(\omega) \downarrow \tau(\omega)$ for all $\omega \in \Omega$, and $\tau_n : \Omega \rightarrow [s,t]$ has a finite range. By step (i) we know that for all $n \geq 1$ we have

$$\mathbb{E}^{\mathcal{F}_{\tau_n}} |\alpha_{\tau_n}| \leq w_{\tau_n}. \tag{7.1}$$

Consider now the martingale

$$(M_r^n)_{r \in [s,t]} := \left(\mathbb{E} \left[|\alpha_{\tau_n}| \mid \mathcal{F}_r \right] \right)_{r \in [s,t]}.$$

By optional stopping, and the fact that $\tau \leq \tau_n \leq t$ for all $n \geq 1$, we have

$$\mathbb{E} \left[M_{\tau_n}^n \mid \mathcal{F}_\tau \right] = M_\tau^n.$$

¹³Unique up to indistinguishability.

Moreover, using optional stopping and the fact that w is a right-continuous supermartingale, we deduce

$$\mathbb{E}^{\mathcal{F}_\tau} w_{\tau_n} \leq w_\tau.$$

Now, applying $\mathbb{E}^{\mathcal{F}_\tau}$ on both sides of equation (7.1), we have that

$$\mathbb{E}^{\mathcal{F}_\tau} |\alpha_{\tau_n}| \leq w_\tau.$$

Since α is right-continuous, we may apply dominated convergence to deduce that we have for any $A \in \mathcal{F}_\tau$

$$\int_A \mathbb{E} \left[|\alpha_\tau| \mid \mathcal{F}_\tau \right] d\mathbb{P} = \lim_n \int_A |\alpha_{\tau_n}| d\mathbb{P} \leq \lim_n \int_A w_\tau d\mathbb{P}.$$

□

8. Appendix B: Tools related to decoupling

The aim of the section is the proof of Proposition 8.6 below that was used in the proof of Corollary 4.8. We start with some preparations before we turn to Proposition 8.6.

Given a probability space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$, the space of equivalence classes $L_0(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$ can be equipped with the metric

$$d(X, X') := \int_{\hat{\Omega}} \frac{|X - X'|}{1 + |X - X'|} d\hat{\mathbb{P}}.$$

It is proven in [18, Proposition 2.5] that the decoupling operators defined in Section 4 are isometries. In particular, given a Borel-measurable function $\varphi : (0, T] \rightarrow [0, 1]$ and $S \in \{0, T\}$, it follows for any $X, Y \in L_0(\bar{\Omega}_S, \Sigma_S^0, \bar{\mathbb{P}}_S)$ that $d(X, Y) = d(X^\varphi, Y^\varphi)$.

8.1 Lemma. *Assume that \mathbb{X} is locally σ -compact, and let $A \subseteq \mathbb{X}$ dense and countable. If $h : \mathbb{X} \rightarrow \mathbb{R}$ is continuous, then $\sup_{x \in \mathbb{X}} h(x) = \sup_{x \in A} h(x)$. Furthermore, if $f_1, f_2 \in [f] \in L_0(\bar{\Omega}_T, \mathcal{B}([0, T]) \otimes \bar{\mathcal{F}}, \bar{\mathbb{P}}; C(\mathbb{X}))$, then*

$$\mathbb{E} \int_0^T \sup_{x \in \mathbb{X}} |f_2(r, x)| dr = \mathbb{E} \int_0^T \sup_{x \in \mathbb{X}} |f_1(r, x)| dr.$$

For the following recall that $(\bar{\Omega}_S, \Sigma_S^\varphi, \bar{\mathbb{P}}_S)$ was introduced in equation (4.1).

8.2 Lemma. *Assume that \mathbb{X} is locally σ -compact. Let $S \in \{0, T\}$, $f \in \mathcal{L}_0(\bar{\Omega}_S, \Sigma_S^0, \bar{\mathbb{P}}_S; C(\mathbb{X}))$, and put for all $\eta \in \bar{\Omega}_S$ and all $x \in \mathbb{X}$*

$$g(\eta, x) := f(\eta, x) 1_{\{|\bar{\eta} \in \bar{\Omega}_S| \sup_{y \in \mathbb{X}} f(\bar{\eta}, y) \in \mathbb{R}\}}.$$

Then it holds that $g \in \mathcal{L}_0(\bar{\Omega}_S, \Sigma_S^\varphi, \bar{\mathbb{P}}_S; C(\mathbb{X}))$, and any representative $g^\varphi \in \mathcal{L}_0(\bar{\Omega}_S, \Sigma_S^\varphi, \bar{\mathbb{P}}_S; C(\mathbb{X}))$ satisfies $\bar{\mathbb{P}}_S(\sup_{x \in \mathbb{X}} g^\varphi(x) \in \mathbb{R}) = 1$, and

$$\sup_{x \in \mathbb{X}} g^\varphi(x) 1_{\{\sup_{y \in \mathbb{X}} g^\varphi(y) \in \mathbb{R}\}} \in \left(\sup_{x \in \mathbb{X}} g(x) \right)^\varphi.$$

Consequently, there exists a representative h^φ of $g^\varphi \in L_0(\bar{\Omega}_S, \Sigma_S^\varphi, \bar{\mathbb{P}}_S; C(\mathbb{X}))$ such that

$$\sup_{x \in \mathbb{X}} h^\varphi(x) \in \left(\sup_{x \in \mathbb{X}} g(x) \right)^\varphi.$$

Proof. The claim $g \in \mathcal{L}_0(\overline{\Omega}_S, \Sigma_S^0, \overline{\mathbb{P}}_S; C(\mathbb{X}))$ follows from Lemma 8.1. Since $\sup_{x \in \mathbb{A}} g(\eta, x) \in \mathbb{R}$ for all $\eta \in \overline{\Omega}_S$, we have that $\sup_{x \in \mathbb{A}} g(x) \in \mathcal{L}_0(\overline{\Omega}_S, \Sigma_S^0, \overline{\mathbb{P}}_S)$. Since \mathbb{A} is countable, we can fix finite sets $\mathbb{A}_1 \subseteq \mathbb{A}_2 \subseteq \dots \subseteq \mathbb{A}$ such that $\bigcup_{n \in \mathbb{N}} \mathbb{A}_n = \mathbb{A}$. Using Proposition 4.3(ii) and the isometry-property, we have

$$d\left(\left(\sup_{x \in \mathbb{A}} g(x)\right)^\varphi, \sup_{x \in \mathbb{A}_k} g^\varphi(x)\right) = d\left(\left(\sup_{x \in \mathbb{A}} g(x)\right)^\varphi, \left(\sup_{x \in \mathbb{A}_k} g(x)\right)^\varphi\right) = d\left(\sup_{x \in \mathbb{A}} g(x), \sup_{x \in \mathbb{A}_k} g(x)\right) \rightarrow 0,$$

as $k \rightarrow \infty$. From this, and from the fact that $(\sup_{x \in \mathbb{A}_k} g^\varphi(\eta, x))_{k \in \mathbb{N}}$ is monotone for all $\eta \in \overline{\Omega}_S$, we deduce that $\sup_{x \in \mathbb{A}_k} g^\varphi(x)$ converges $\overline{\mathbb{P}}_S$ -a.s. to $(\sup_{x \in \mathbb{A}} g(x))^\varphi$. On the other hand, the monotonicity also implies that

$$\lim_{k \rightarrow \infty} \sup_{x \in \mathbb{A}_k} g^\varphi(\eta, x) = \sup_{x \in \mathbb{A}} g^\varphi(\eta, x)$$

for all $\eta \in \overline{\Omega}_S$. Hence, it follows from continuity that $\sup_{x \in \mathbb{X}} g^\varphi(x)$ is $\overline{\mathbb{P}}_S$ -a.s. finite and

$$\sup_{x \in \mathbb{X}} g^\varphi(x) \mathbf{1}_{\{\sup_{y \in \mathbb{X}} g^\varphi(y) \in \mathbb{R}\}} \in \left(\sup_{x \in \mathbb{X}} g(x)\right)^\varphi.$$

□

8.3 Remark. Lemma 8.2 implies that if the assumptions of Lemma 7.1 are satisfied by $f \in \mathcal{L}_0(\overline{\Omega}_T, \Sigma_T^0, \overline{\mathbb{P}}_T; C(\mathbb{X}))$, then they are also satisfied by f^φ . This holds, since applying Lemma 8.2 restricted to a compact $K \subseteq \mathbb{X}$, we notice that if $\mathbb{E} \sup_{x \in K} |f(x)| < \infty$, then $\mathbb{E} \sup_{x \in K} |f^\varphi(x)| = \mathbb{E} \sup_{x \in K} |f(x)|$.

8.4 Lemma ([18, Remark 2.14]). *Let $X \in \mathcal{L}_0(\overline{\Omega}_T, \Sigma_T^0, \overline{\mathbb{P}}_T)$ such that $\int_0^T |X(t, \omega)| dt < \infty$ for all $\omega \in \overline{\Omega}$. Then for any representative $X^\varphi \in \mathcal{L}_0(\overline{\Omega}_T, \Sigma_T^\varphi, \overline{\mathbb{P}}_T)$ we have that $\overline{\mathbb{P}}\left(\int_0^T |X^\varphi(t)| dt < \infty\right) = 1$, and*

$$\int_0^T X^\varphi(t) \mathbf{1}_{\{\int_0^T |X^\varphi(s)| ds < \infty\}} dt \in \left(\int_0^T X(t) dt\right)^\varphi.$$

8.5 Lemma. *Let \mathbb{X} be locally σ -compact and let $f \in \mathcal{L}_0(\overline{\Omega}_T, \Sigma_T^0, \overline{\mathbb{P}}_T; C(\mathbb{X}))$ such that*

$$\overline{\mathbb{P}}\left(\int_0^T \sup_{x \in \mathbb{X}} |f(t, \omega, x)| dt < \infty\right) = 1.$$

Then there exists a representative h^φ of $|f^\varphi| \in L_0(\overline{\Omega}_T, \Sigma_T^\varphi, \overline{\mathbb{P}}_T; C(\mathbb{X}))$ such that

$$\int_0^T \sup_{x \in \mathbb{X}} |h^\varphi(t, x)| dt \in \left(\int_0^T \sup_{x \in \mathbb{X}} |f(t, x)| dt\right)^\varphi.$$

Proof. First note that

$$\overline{\mathbb{P}}\left(\int_0^T \sup_{x \in \mathbb{X}} |f(r, \omega, x)| dr < \infty\right) = 1$$

implies

$$\overline{\mathbb{P}}_T\left(\sup_{x \in \mathbb{X}} |f(t, \omega, x)| < \infty\right) = 1.$$

We may redefine f such that $\sup_{x \in \mathbb{X}} |f(t, \omega, x)| < \infty$ for all $(t, \omega) \in \overline{\Omega}_T$, and $\int_0^T \sup_{x \in \mathbb{X}} |f(r, \omega, x)| dr < \infty$ for all $\omega \in \overline{\Omega}$. It is a direct consequence of Proposition 4.3(ii) that $|f|^\varphi = |f^\varphi|$, so that we may look for a representative of $|f|^\varphi$ that satisfies the claim. Applying Lemma 8.2 to $|f|$ gives us a representative h^φ of $|f|^\varphi \in L_0(\overline{\Omega}_T, \Sigma_T^\varphi, \overline{\mathbb{P}}_T; C(\mathbb{X}))$ such that $\sup_{x \in \mathbb{X}} h^\varphi(x) \in (\sup_{x \in \mathbb{X}} |f(x)|)^\varphi$. Letting $X(t, \omega) := \sup_{x \in \mathbb{X}} |f(t, \omega, x)|$ for $(t, \omega) \in \overline{\Omega}_T$, we then have that $\sup_{x \in \mathbb{X}} h^\varphi(x)$ is a representative of X^φ . Hence, Lemma 8.4 implies that

$$\overline{\mathbb{P}} \left(\int_0^T \sup_{x \in \mathbb{X}} |h^\varphi(t, x)| dt < \infty \right) = 1,$$

and

$$\int_0^T \sup_{x \in \mathbb{X}} |h^\varphi(t, x)| 1_{\{\int_0^T \sup_{x \in \mathbb{X}} |h^\varphi(r, x)| dr < \infty\}} dt \in \left(\int_0^T \sup_{x \in \mathbb{X}} |f(t, x)| dt \right)^\varphi.$$

The representative of $|f|^\varphi \in L_0(\overline{\Omega}_T, \Sigma_T^\varphi, \overline{\mathbb{P}}_T; C(\mathbb{X}))$ that satisfies the claim, is $|h^\varphi| 1_{\{\int_0^T \sup_{x \in \mathbb{X}} |h^\varphi(r, x)| dr < \infty\}}$. \square

We are ready to prove the desired result. Recall that $\overline{\mathcal{G}}_s^t$ and $\overline{\mathcal{H}}_s^t$ were defined in (4.5) and $\xi^{(a,b)}$ in (4.4).

8.6 Proposition. *Assume that \mathbb{X} is locally σ -compact. Let $p \in [1, \infty)$, $0 \leq s < t \leq T$, $0 \leq u_1 < u_2 \leq T$, and $f \in \mathcal{L}_0(\overline{\Omega}_T, \Sigma_T^0, \overline{\mathbb{P}}_T; C(\mathbb{X}))$ such that $\int_{\overline{\Omega}_T} \sup_{x \in K} |f(x)| d\overline{\mathbb{P}}_T < \infty$ for every compact $K \subseteq \mathbb{X}$. If*

$$\left\| \int_{u_1}^{u_2} \sup_{x \in \mathbb{X}} |f(r, x) - (\mathbb{E}^{\overline{\mathcal{H}}_s^t} f)(r, x)| dr \right\|_p < \infty,$$

then $\overline{\mathbb{P}}$ -a.s.

$$\mathbb{E}^{\overline{\mathcal{G}}_s^t} \left(\int_{u_1}^{u_2} \sup_{x \in \mathbb{X}} |f(r, x) - f^{(s,t)}(r, x)| dr \right)^p \leq 2^p \mathbb{E}^{\overline{\mathcal{G}}_s^t} \left(\int_{u_1}^{u_2} \sup_{x \in \mathbb{X}} |f(r, x) - (\mathbb{E}^{\overline{\mathcal{H}}_s^t} f)(r, x)| dr \right)^p.$$

Conversely, if

$$\left\| \int_{u_1}^{u_2} \sup_{x \in \mathbb{X}} |f(r, x) - f^{(s,t)}(r, x)| dr \right\|_p < \infty,$$

then $\overline{\mathbb{P}}$ -a.s.

$$\mathbb{E}^{\overline{\mathcal{G}}_s^t} \left(\int_{u_1}^{u_2} \sup_{x \in \mathbb{X}} |f(r, x) - (\mathbb{E}^{\overline{\mathcal{H}}_s^t} f)(r, x)| dr \right)^p \leq \mathbb{E}^{\overline{\mathcal{G}}_s^t} \left(\int_{u_1}^{u_2} \sup_{x \in \mathbb{X}} |f(r, x) - f^{(s,t)}(r, x)| dr \right)^p.$$

8.7 Remark.

- (1) To define $\mathbb{E}^{\overline{\mathcal{H}}_s^t} f \in \mathcal{L}_0(\overline{\Omega}_T, \overline{\mathcal{H}}_s^t, \overline{\mathbb{P}}_T; C(\mathbb{X}))$ we apply Lemma 7.1, and this is why we need to assume that $\int_{\overline{\Omega}_T} \sup_{x \in K} |f(x)| d\overline{\mathbb{P}}_T < \infty$ for every compact $K \subseteq \mathbb{X}$.
- (2) The conclusion of Proposition 8.6 with $p = 1$ implies that

$$\frac{1}{2} \mathbb{E}^{\overline{\mathcal{H}}_s^t} \|f - f^{(s,t)}\|_{C(\mathbb{X})} \leq \mathbb{E}^{\overline{\mathcal{H}}_s^t} \|f - \mathbb{E}^{\overline{\mathcal{H}}_s^t} f\|_{C(\mathbb{X})} \leq \mathbb{E}^{\overline{\mathcal{H}}_s^t} \|f - f^{(s,t)}\|_{C(\mathbb{X})}.$$

Hence, Proposition 8.6 generalizes Lemma 4.6 from random variables $\xi : \Omega \rightarrow \mathbb{R}$ to function-space valued stochastic processes $f : \Omega_T \rightarrow C(\mathbb{X})$.

Proof of Proposition 8.6. We will use $\|\cdot\|_p$ for $\|\cdot\|_{L_p(\bar{\Omega})}$ and \mathbb{A} for a fixed dense countable subset of \mathbb{X} . Note that $\sup_{x \in \mathbb{X}} h(x) = \sup_{x \in \mathbb{A}} h(x)$ whenever h is continuous, so that we may replace \mathbb{X} by \mathbb{A} in the proof below. To simplify the notation in the proof, we assume that $u_1 = 0$ and $u_2 = T$.

Step 1: We will first show that if $g \in \mathcal{L}_0(\bar{\Omega}_T, \overline{\mathcal{H}}_s^t, \overline{\mathbb{P}}_T; C(\mathbb{X}))$ is such that

$$\left\| \int_0^T \sup_{x \in \mathbb{A}} |g(r, x) - f(r, x)| dr \right\|_p < \infty,$$

then

$$\left\| \int_0^T \sup_{x \in \mathbb{A}} |f(r, x) - f^{(s,t)}(r, x)| dr \right\|_p \leq 2 \left\| \int_0^T \sup_{x \in \mathbb{A}} |f(r, x) - g(r, x)| dr \right\|_p.$$

Fixing g as described above, Lemma 4.5(v) implies that $g \in g^{(s,t]}$, so Lemma 8.5 applied to $g - f$ in particular implies

$$\mathbb{E} \left(\int_0^T \sup_{x \in \mathbb{A}} |g(r, x) - f^{(s,t)}(r, x)| dr \right)^p = \mathbb{E} \left(\int_0^T \sup_{x \in \mathbb{A}} |g(r, x) - f(r, x)| dr \right)^p.$$

From this we deduce

$$\begin{aligned} & \left\| \int_0^T \sup_{x \in \mathbb{A}} |f(r, x) - f^{(s,t)}(r, x)| dr \right\|_p \\ & \leq \left\| \int_0^T \sup_{x \in \mathbb{A}} |f(r, x) - g(r, x)| dr \right\|_p + \left\| \int_0^T \sup_{x \in \mathbb{A}} |g(r, x) - f^{(s,t)}(r, x)| dr \right\|_p \\ & = 2 \left\| \int_0^T \sup_{x \in \mathbb{A}} |f(r, x) - g(r, x)| dr \right\|_p. \end{aligned}$$

Step 2: We assume that $\left\| \int_0^T \sup_{x \in \mathbb{A}} |f(r, x) - f^{(s,t)}(r, x)| dr \right\|_p < \infty$, and will show that

$$\left\| \int_0^T \sup_{x \in \mathbb{A}} |f(r, x) - (\mathbb{E}^{\overline{\mathcal{H}}_s^t} f)(r, x)| dr \right\|_p \leq \left\| \int_0^T \sup_{x \in \mathbb{A}} |f(r, x) - f^{(s,t)}(r, x)| dr \right\|_p.$$

We use W^0, W^1 to denote the canonical extensions of W, W' , respectively, and for $0 \leq a < b \leq T$ we work with the σ -algebras

$$\begin{aligned} \mathcal{H}_{a,b}^{W^0} & := \mathcal{B}([0, T]) \otimes \sigma(W_r^0 - W_a^0, r \in [a, b]), \\ \mathcal{H}_{a,b}^{W^1} & := \mathcal{B}([0, T]) \otimes \sigma(W_r^1 - W_a^1, r \in [a, b]), \\ \mathcal{H} & := \{\emptyset, [0, T]\} \otimes \sigma(W_r^0 - W_s^0, r \in [s, t]). \end{aligned}$$

Note that these are σ -algebras in $\bar{\Omega}_T$, and we have the inclusions

$$\begin{aligned} \mathcal{H}_{0,T}^{W^0} & \subseteq \Sigma_T^0, \\ \mathcal{H}_{0,s}^{W^0} \vee \mathcal{H}_{s,t}^{W^1} \vee \mathcal{H}_{t,T}^{W^0} & \subseteq \Sigma_T^{(s,t]}, \\ \mathcal{H}_{0,s}^{W^0} \vee \mathcal{H}_{t,T}^{W^0} & \subseteq \overline{\mathcal{H}}_s^t. \end{aligned}$$

Moreover, the inclusions are "up to nullsets", which in this context means that we have for example

$$\Sigma_T^0 = \mathcal{H}_{0,T}^{W^0} \vee (\mathcal{B}([0, T]) \otimes \overline{\mathcal{N}}),$$

where $\overline{\mathcal{N}}$ are the $\overline{\mathbb{P}}$ -nullsets. From this it follows that f and $\mathbb{E}^{\mathcal{H}_{0,T}^{W^0}} f$ given by Lemma 7.1 are indistinguishable. To keep the notations as light as possible, we simply say that using Lemma 8.1 and Lemma 7.1, we may assume that

- (1) f is $\mathcal{H}_{0,T}^{W^0}$ -measurable,
- (2) $f^{(s,t]}$ is $\mathcal{H}_{0,s}^{W^0} \vee \mathcal{H}_{s,t}^{W^1} \vee \mathcal{H}_{t,T}^{W^0}$ -measurable,
- (3) $\mathbb{E}^{\overline{\mathcal{H}}_s^i} f^{(s,t]}$ is $\mathcal{H}_{0,s}^{W^0} \vee \mathcal{H}_{t,T}^{W^0}$ -measurable.

Then the facts that for all $x \in \mathbb{A}$

$$\mathcal{H}_{0,s}^{W^0} \vee \mathcal{H}_{t,T}^{W^0} \vee \sigma(f^{(s,t]}(x)) \text{ is independent of } \mathcal{H},$$

$$f(x) \text{ is } \mathcal{H}_{0,s}^{W^0} \vee \mathcal{H}_{t,T}^{W^0} \vee \mathcal{H}\text{-measurable,}$$

are immediate. Hence, it follows from [28, 9.7(k)] that

$$\mathbb{E}^{\overline{\mathcal{H}}_s^i \vee \mathcal{H}} f^{(s,t]}(x) = \mathbb{E}^{\overline{\mathcal{H}}_s^i} f^{(s,t]}(x)$$

for all $x \in \mathbb{A}$. Since $f(x) \in L_1(\overline{\Omega}_T, \Sigma_T^0, \overline{\mathbb{P}}_T)$ for all $x \in \mathbb{A}$, it follows from Proposition 4.5(i) that $\mathbb{E}^{\overline{\mathcal{H}}_s^i} f(x) = \mathbb{E}^{\overline{\mathcal{H}}_s^i} f^{(s,t]}(x)$ for all $x \in \mathbb{A}$. Thus we have

$$\begin{aligned} \left\| \int_0^T \sup_{x \in \mathbb{A}} |f(r, x) - (\mathbb{E}^{\overline{\mathcal{H}}_s^i} f)(r, x)| dr \right\|_p &= \left\| \int_0^T \sup_{x \in \mathbb{A}} |f(r, x) - (\mathbb{E}^{\overline{\mathcal{H}}_s^i} f^{(s,t]})(r, x)| dr \right\|_p \\ &= \left\| \int_0^T \sup_{x \in \mathbb{A}} |f(r, x) - (\mathbb{E}^{\overline{\mathcal{H}}_s^i \vee \mathcal{H}} f^{(s,t]})(r, x)| dr \right\|_p \\ &= \left\| \int_0^T \sup_{x \in \mathbb{A}} |\mathbb{E}^{\overline{\mathcal{H}}_s^i \vee \mathcal{H}}(f(r, x) - f^{(s,t]}(r, x))| dr \right\|_p \\ &\leq \left\| \int_0^T \sup_{x \in \mathbb{A}} \mathbb{E}^{\overline{\mathcal{H}}_s^i \vee \mathcal{H}} |f(r, x) - f^{(s,t]}(r, x)| dr \right\|_p \\ &\leq \left\| \int_0^T \mathbb{E}^{\overline{\mathcal{H}}_s^i \vee \mathcal{H}} \left(\sup_{x \in \mathbb{A}} |f(r, x) - f^{(s,t]}(r, x)| \right) dr \right\|_p \\ &= \left\| \mathbb{E}^{\mathcal{F}^0} \left(\int_0^T \sup_{x \in \mathbb{A}} |f(r, x) - f^{(s,t]}(r, x)| dr \right) \right\|_p \\ &\leq \left\| \int_0^T \sup_{x \in \mathbb{A}} |f(r, x) - f^{(s,t]}(r, x)| dr \right\|_p. \end{aligned}$$

Step 3: The conditional claim follows from the result with the full expectation as in Lemma 4.6: assume that $f \in \mathcal{L}_0(\overline{\Omega}_T, \Sigma_T^0, \overline{\mathbb{P}}_T; C(\mathbb{X}))$ is such that $\int_{\overline{\Omega}_T} \sup_{x \in K} |f(x)| d\overline{\mathbb{P}}_T < \infty$ for every compact $K \subseteq \mathbb{X}$. Let $B \in \overline{\mathcal{G}}_s^t$ with $\overline{\mathbb{P}}(B) > 0$, and define

$$\tilde{f} \in \mathcal{L}_0(\overline{\Omega}_T, \Sigma_T^0, \overline{\mathbb{P}}_T; C(\mathbb{X})) \text{ by } \tilde{f}(r, \omega, x) := f(r, \omega, x) 1_B(\omega).$$

Fixing any representative of $f^{(s,t]}$, we have that

- (1) $\int_{\Omega_T} \sup_{x \in K} |\tilde{f}(x)| d\bar{\mathbb{P}}_T < \infty$ for every compact $K \subseteq \mathbb{X}$,
- (2) $1_B(\mathbb{E}^{\mathcal{H}_t^i} f)$ is a representative of $(\mathbb{E}^{\mathcal{H}_t^i} \tilde{f})$,
- (3) $1_B f^{(s,t]}$ is a representative of $\tilde{f}^{(s,t]}$,

so the claim follows by applying steps 1 and 2 with \tilde{f} . \square

9. Appendic C: A John-Nirenberg type theorem

We recall the result [17] (Theorem 9.1). Whereas in [17] càdlàg processes are considered, we only need the case of continuous processes. Fix $R > 0$, let $(\Omega, \mathcal{G}_R, \mathbb{P}, (\mathcal{G}_r)_{r \in [0,R]})$ be a stochastic basis such that $(\Omega, \mathcal{G}_R, \mathbb{P})$ is complete, $(\mathcal{G}_r)_{r \in [0,R]}$ is right-continuous, and \mathcal{G}_0 contains all nullsets, and let $A = (A_r)_{r \in [0,R]}$ be a *continuous*, adapted stochastic process with $A_0 \equiv 0$. Moreover, we assume that $(\Psi_r)_{r \in [0,R]}$ is a càdlàg $(\mathcal{G}_r)_{r \in [0,R]}$ -adapted stochastic process such that $\Psi_r(\omega) > 0$ for all $(r, \omega) \in \Omega_R$. Put

$$\mathcal{S}_{0,R}^{\mathcal{G}} := \left\{ \tau : \Omega \rightarrow [0, R] \mid \tau \text{ is a } (\mathcal{G}_r)_{r \in [0,R]} \text{ - stopping time} \right\}$$

and define

$$W_{\Psi}(B, \nu; \tau) := \mathbb{P} \left(B \cap \left\{ \sup_{u \in [\tau, R]} \Psi_u > \nu \right\} \right)$$

for $\nu > 0$, $\tau \in \mathcal{S}_{0,R}^{\mathcal{G}}$, and $B \in \mathcal{G}_{\tau}$. Recall that for $B \in \mathcal{G}_R$ of positive measure

$$\mathbb{P}_B(\cdot) := \frac{\mathbb{P}(B \cap \cdot)}{\mathbb{P}(B)}.$$

9.1 Theorem ([17, Theorem 1]). *Assume that there is an $\alpha \in (0, \frac{1}{2})$ such that*

$$\mathbb{P}_B(|A_R - A_{\tau}| > \nu) \leq \alpha + \frac{W_{\Psi}(B, \nu; \tau)}{\mathbb{P}(B)} \tag{9.1}$$

for all $\nu > 0$, $\tau \in \mathcal{S}_{0,R}^{\mathcal{G}}$, and $B \in \mathcal{G}_{\tau}$ of positive measure. Then there are constants $a, c > 0$, depending on α only, such that

$$\mathbb{P}_B \left(\sup_{u \in [\tau, R]} |A_u - A_{\tau}| > \lambda + a\mu\nu \right) \leq e^{1-\mu} \mathbb{P}_B \left(\sup_{u \in [\tau, R]} |A_u - A_{\tau}| > \lambda \right) + c \frac{W_{\Psi}(B, \nu; \tau)}{\mathbb{P}(B)}$$

for all $\lambda, \mu, \nu > 0$, $\tau \in \mathcal{S}_{0,R}^{\mathcal{G}}$, and $B \in \mathcal{G}_{\tau}$ of positive measure.

References

- [1] R. Avikainen: Convergence rates for approximations of functionals of SDEs. arXiv:0712.3635.
- [2] R. Bañuelos and A.G. Bennett: Paraproducts and Commutators of Martingale Transforms. *Proc. AMS* 103:1226-1234, 1988.

- [3] M.T. Barlow, M. Yor: Semi-martingale inequalities via the Garsia-Rodemich-Rumsey Lemma, and applications to local times. *Journal of Functional Analysis*, Vol.49, 198-229, 1982.
- [4] J-M. Bismut: Conjugate convex functions in optimal stochastic control. *Journal of Mathematical Analysis and Applications*, Vol.44(2), 384-404, 1973.
- [5] B. Bouchard, N. Touzi: Discrete Time Approximation and Monte-Carlo Simulation of Backward Stochastic Differential Equations. *Stoch. Proc. Appl.* Vol. 111(2), 175-206, 2004.
- [6] B. Bouchard, R. Elie, N. Touzi: Discrete-time approximation of BSDEs and probabilistic schemes for fully nonlinear PDEs. *Radon Series Comp. Appl. Math*, 8, 1-34, 2009.
- [7] P. Briand, B. Delyon, Y. Hu, E. Pardoux, L. Stoica: L_p solutions of backward stochastic differential equations. *Stoch. Proc. Appl.* 108:109-129, 2003.
- [8] P. Briand, R. Elie: A simple constructive approach to quadratic BSDEs with or without delay. *Stoch. Proc. Appl.* 2921-2939, 2013.
- [9] P. Briand, Y. Hu: BSDE with quadratic growth and unbounded terminal value. *Probab. Theory Related Fields*, 136(4), 604-618, 2006.
- [10] D.L. Burkholder: Explorations in martingale theory and its applications. École d'Été de Probabilités de Saint-Flour XIX 1989, Lecture Notes in Math. 1464, Springer, Berlin, 1991, pp. 1-66.
- [11] P. Cheridito, K. Nam: BSDEs with terminal conditions that have bounded Malliavin derivative. *J. Funct. Anal.*, 266, 1257-1285, 2014.
- [12] Delbaen, Hu, Richou: On the uniqueness of solutions to quadratic BSDEs with convex generators and unbounded terminal conditions. *Ann. Inst. H. Poincaré Probab. Statist.*, Vol.47(2), 559-574, 2011.
- [13] F. Delbaen, S. Tang: Harmonic analysis of stochastic equations and backward stochastic differential equations. *Probab. Theory Related Fields* 146:291-336, 2010.
- [14] J. Diestel, J.J. Uhl: Vector Measures. Math. Surveys, Number 15, AMS Providence, Rhode Island, 1977.
- [15] N. El Karoui, S. Peng, M.C. Quenez: Backward stochastic differential equations in finance. *Mathematical Finance* 7:1-71, 1997.
- [16] C. Geiss, S. Geiss, E. Gobet: Generalized fractional smoothness and L_p -variation of BSDEs with non-Lipschitz terminal condition. *Stoch. Proc. Appl.*, 122:2078-2116, 2012.
- [17] S. Geiss: Weighted BMO and discrete time hedging within the Black-Scholes model. *Probab. Theory Related Fields*, 132, 13-38, 2005.
- [18] S. Geiss, J. Ylisen: Decoupling on the Wiener space and applications to BSDEs. To appear in Memoir AMS.
- [19] Y. Hu, P. Imkeller, M. Müller: Utility maximization in incomplete markets. *Ann. Appl. Probab.*, Vol.15(3), 1691-1712, 2005.
- [20] I. Karatzas, S. Shreve: Brownian Motion and Stochastic Calculus. 2nd Edition. Springer, 1991.

- [21] N. Kazamaki: Continuous Exponential Martingales and BMO. *Lecture Notes in Mathematics 1579*, Springer, 1994.
- [22] M. Kobylanski: Backward stochastic differential equations and partial differential equations with quadratic growth. *Annals of Probability*, 28(2), 558-602, 2000.
- [23] M-A. Morlais: Quadratic BSDEs driven by a continuous martingale and applications to the utility maximization problem. *Finance and Stochastics*, Vol.13(1), 121-150, 2009.
- [24] E. Pardoux, S. Peng: Adapted Solution of a Backward Stochastic Differential Equation. *Systems Control Lett.*, 14, 55-61, 1990.
- [25] E. Pardoux, S. Peng: Backward Stochastic Differential Equations and Quasilinear Parabolic Partial Differential Equations. *Lecture Notes in CIS, vol.176*, Springer-Verlag, 200-217, 1992.
- [26] D. Revuz, M. Yor: Continuous martingales and Brownian motion. Springer, 1998
- [27] A. Richou: Markovian quadratic and superquadratic BSDEs with an unbounded terminal condition. *Stoch. Proc. Appl.*, 122:3173-3208, 2012.
- [28] D. Williams Probability with Martingales Cambridge University Press 1991.
- [29] J. Zhang: Some fine properties of backward stochastic differential equations. PhD thesis, Purdue university, 2001.
- [30] J. Zhang: A numerical scheme for BSDEs. *Ann. of Appl. Prob.* 14:459-488, 2004.