UNIVERSITY OF JYVÄSKYLÄ
DEPARTMENT OF MATHEMATICS
AND STATISTICS

REPORT 148

UNIVERSITÄT JYVÄSKYLÄ
INSTITUT FÜR MATHEMATIK
UND STATISTIK

BERICHT 148

DECOUPLING ON THE WIENER SPACE AND VARIATIONAL ESTIMATES FOR BSDEs

JUHA YLINEN

JYVÄSKYLÄ
2015

UNIVERSITY OF JYVÄSKYLÄ
DEPARTMENT OF MATHEMATICS
AND STATISTICS

REPORT 148

UNIVERSITÄT JYVÄSKYLÄ
INSTITUT FÜR MATHEMATIK
UND STATISTIK

BERICHT 148

DECOUPLING ON THE WIENER SPACE AND VARIATIONAL ESTIMATES FOR BSDEs

JUHA YLINEN

To be presented, with the permission of the Faculty of Mathematics and Science of the University of Jyväskylä, for public criticism in Auditorium MaA 211 on May 29th, 2015, at 12 o'clock noon.

JYVÄSKYLÄ
2015

Editor: Pekka Koskela
Department of Mathematics and Statistics
P.O. Box 35 (MaD)
FI–40014 University of Jyväskylä
Finland

ISBN 978-951-39-6442-9 (PDF)

ISBN 978-951-39-6193-0
ISSN 1457-8905

University Printing House
Jyväskylä 2015

Acknowledgements

I wish to express my gratitude to my advisor Stefan Geiss for suggesting interesting topics, correcting my errors, and his continuing help during my studies. Everyone at the Department of Mathematics and Statistics of the University of Jyväskylä deserves my thanks for making my studies enjoyable and educational. In particular I want to thank Tuula Blåfield, Taimi Närhi, Eeva Partanen and Hannele Säntti-Ahomäki for making sure I don't need to worry about the practical work-related matters. Also, the people at the Department of Mathematics of the University of Innsbruck made my many visits there delightful.

For financial support I am indebted to the project "Stochastic and Harmonic Analysis, interactions, and applications", No. 133914, of the Academy of Finland, as well as to the Vilho, Yrjö and Kalle Väisälä foundation of the Finnish Academy of Science and Letters.

I am grateful to "Bebetos of Pirttilä" for always providing facilities and opponents in any sport. Finally, I would like to thank my parents and brothers for keeping me as sane as possible.

Jyväskylä, April 2015

Juha Ylinen

This thesis consists of an introductory part and the following articles:

List of included articles

[A] S. Geiss, J. Ylinen: Decoupling on the Wiener space and applications to BSDEs. http://arxiv.org/abs/1409.5322v2

[B] J. Ylinen: Tales and tails of BSDEs. http://arxiv.org/abs/1501.01183

In the introductory part, these articles are referred to as [A] and [B], whereas the other references will be numbered as [1], [2], ...

The author of this dissertation has actively taken part in the research of the joint paper [A].

Contents

1 Introduction

The main subject of this thesis are Backward Stochastic Differential Equations, BSDEs from now on, of type

$$Y_t = \xi + \int_t^T f(r, Y_r, Z_r) dr - \int_t^T Z_r dW_r, \quad t \in [0, T], \tag{1}$$

where $T > 0$ is a fixed time horizon and W is a d-dimensional Brownian motion. The pair (ξ, f) is referred to as the data, and the pair (Y, Z) (or just the process Y) as the solution. We deduce different types of upper bounds for the variation of the solution on subintervals of $[0, T]$. These upper bounds are given, as is usual with BSDEs, in terms of the data. In [A] an L_p-quantity of the variation is upper bounded by an L_p-quantity of the data, with $p \in [2, \infty)$. The upper bound is obtained using a new decoupling technique, which also gives rise to anisotropic Besov spaces. These spaces include for example the Besov spaces obtained by real interpolation, and we also show a connection of a certain anisotropic Besov space to the Malliavin derivative. In [B] we show that the solution is of weighted bounded mean oscillation (weighted BMO), where the weight process is given in terms of the data. Using the theory of weighted BMO, the variation of the solution is shown to satisfy a tail estimate that is better than what would be obtained from an L_p-estimate.

2 BSDEs

We start by introducing our setting and fixing some notation. In this thesis we work on a complete stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \in [0,T]})$, where "the information" $(\mathcal{F}_t)_{t \in [0,T]}$ is given by a d-dimensional Brownian motion $W = (W_t)_{t \in [0,T]}$ with $d \geq 1$, and $\mathcal{F} = \mathcal{F}_T$. To be precise, $(\mathcal{F}_t)_{t \in [0,T]}$ is the augmented natural filtration of W. The predictable σ-algebra generated by left-continuous $(\mathcal{F}_t)_{t \in [0,T]}$-adapted processes is denoted by $\mathcal{P}$.

For $X \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$ and a sub-σ-algebra $\mathcal{G} \subseteq \mathcal{F}$ we will use the notation $\mathbb{E}^{\mathcal{G}} X = \mathbb{E}\left[X \mid \mathcal{G}\right]$ for the conditional expectation of X given $\mathcal{G}$. Inequalities concerning random variables, such as $\mathbb{E}^{\mathcal{G}} X \leq Y$, hold in general only almost surely even though this is not always explicitly mentioned. The notations 1_A and χ_A are reserved for the *indicator function* of a set A. That is,

$$1_A(x) = \chi_A(x) = \begin{cases} 1 & , \quad \text{if } x \in A, \\ 0 & , \quad \text{if } x \notin A. \end{cases}$$

In the BSDE (1) we are given a pair (ξ, f), where the terminal value ξ is an $\mathcal{F}_T$-measurable random variable (i.e. at time T you know the exact value of ξ). The generator $f : [0,T] \times \Omega \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ is assumed to be such that

- $(t, \omega) \mapsto f(t, \omega, y, z)$ is $(\mathcal{F}_t)_{t\in[0,T]}$-predictable for all $(y, z) \in \mathbb{R} \times \mathbb{R}^d$, and
- $(y, z) \mapsto f(t, \omega, y, z)$ is continuous for all $(t, \omega) \in [0, T] \times \Omega$.

The solution consists of the pair of stochastic processes (Y, Z), where $t \mapsto Y_t$ is continuous, Y is $(\mathcal{F}_t)_{t\in[0,T]}$-adapted, Z is $(\mathcal{F}_t)_{t\in[0,T]}$-predictable, and (Y, Z) satisfy (1) almost surely.[1]

The philosophy is that if we know the structure or mechanism of ξ, and the dynamics of the system (the generator f), then we want to find (Y, Z). The Y-process starts from a deterministic constant, and travels continuously in time into ξ, almost surely. An important point here is the adaptedness of Y; at time t we know the value of Y_t. The Z-process on the other hand acts as a control process that guides Y into ξ.

In the special case $f \equiv 0$, Y_t equals the conditional expectation of ξ given $\mathcal{F}_t$. Furthermore, if it happens to be that ξ belongs to the Malliavin Sobolev space $\mathbb{D}_{1,2}$, then the Clark-Ocone formula tells us that Z equals the predictable projection of the Malliavin derivative of Y.

The original motivation for studying BSDEs comes from optimal stochastic control theory [2], later connections to mathematical finance were discovered for example in [12] and [19]. BSDEs are also closely connected to a group of partial differential equations, as was already proven through a nonlinear Feynman-Kac theorem in the seminal work [27].

The research of BSDEs was initiated by Bismut, who introduced BSDEs with a generator f that is linear in y, z, i.e. linear BSDEs. Once the case of a uniformly Lipschitz f (in y, z) with data in L_2 was handled in [26], the amount of related papers hugely increased. Another important benchmark with a uniformly Lipschitz f is [5], where the case of data in L_p with $1 \le p < 2$ is handled. A typical estimate with BSDEs is the following "apriori estimate" from [5]:

$$\left\| \sup_{t\in[0,T]} |Y_t| \right\|_p + \left\| \left(\int_0^T |Z_r|^2 dr \right)^{\frac{1}{2}} \right\|_p \le C_p \left[\|\xi\|_p + \left\| \int_0^T |f(r, 0, 0)| dr \right\|_p \right], \quad (2)$$

where $C_p > 0$, $p > 1$, and $\| \cdot \|_p$ stands for $\| \cdot \|_{L_p(\Omega)}$.

For example in connection to utility maximization with exponential utility, the generator f may grow quadratically in the z-variable. These quadratic

[1] This is often abbreviated to $\mathbb{P}$-a.s., also to emphasize the measure used.

BSDEs are considered with a bounded ξ for example in [22] and [19], and with exponential moments in ξ for example in [7] and [9]. An important ingredient when proving the well-posedness in the quadratic case are BMO-martingales (see Section 5). In [10] the authors go a step further by considering martingales that are sliceable in BMO, and this concept is exploited in this thesis as well.

To give more insight into BSDEs, we consider a basic example of asset pricing as described in [12]:

Example 2.1. For simplicity of the presentation we assume here that the Brownian motion $(W_t)_{t\in[0,T]}$ is 1-dimensional. Our market model is that we have two assets: one riskless asset ("bank account") with price per unit P^0 governed by the equation

$$dP_t^0 = P_t^0 r_t dt, \tag{3}$$

where r is the interest rate, and one risky security (the stock) where the price process P^1 is modeled by the SDE

$$dP_t^1 = P_t^1 \left(b_t dt + \sigma_t dW_t\right), \tag{4}$$

where b is the stock appreciation rate and σ is the volatility. Moreover, the coefficients in (3) and (4) are "nice", i.e.

(i) r is a predictable non-negative bounded process,

(ii) b and σ are predictable and bounded,

(iii) $\sigma_t \neq 0$ for any $t \in [0,T]$ almost surely, and σ^{-1} is a bounded process,

(iv) there exists a predictable and bounded process θ (called the risk premium) such that $b_t - r_t = \sigma_t\theta_t$ for all $t \in [0,T]$ almost surely.

We consider a small investor who at time t decides what amount π_t of the wealth V_t to invest in the stock. Since his decisions can be based only on the current information (no insider trading), the processes π_t and $\pi_t^0 := V_t - \pi_t$ are predictable.

A *strategy* (V,π) is *self-financing* if the investor's wealth at time t consists of the initial wealth V_0 and the losses or gains that he has obtained using $(\pi_s)_{s\in[0,t]}$. That is, the wealth process satisfies

$$V_t = V_0 + \int_0^t \pi_s^0 \frac{dP_s^0}{P_s^0} + \int_0^t \pi_s \frac{dP_s^1}{P_s^1},$$

which, using equations (3) and (4), is equivalent to the wealth process satisfying the SDE

$$dV_t = r_t V_t dt + \pi_t \sigma_t [dW_t + \theta_t dt].$$

For this to make sense it has to also hold, $\mathbb{P}$-a.s., that

$$\int_0^T |\pi_t \sigma_t|^2 dt < \infty.$$

The strategy is called *feasible* if $V_t \geq 0$ for all $t \in [0, T]$, $\mathbb{P}$-a.s., i.e. no borrowing is allowed in the model.

Now we consider a non-negative European *contingent claim* $\xi \geq 0$ settled at time T. This is an $\mathcal{F}_T$-measurable random variable, and can be thought of as a contract that pays the amount ξ at maturity T. For example, $\xi = 1_{(K,\infty)}(P_T^1)$ with $K \in (0, \infty)$ is a European contingent claim. A buyer of this claim receives one unit of currency if the value of the stock at time T exceeds the value K. If the value of the stock at maturity is below or equal to K, then the buyer gets nothing.

How much should this claim cost at time 0? It seems fair that if we let V_0 be the price of the claim, then it should be possible to invest this amount into the assets P^0 and P^1 such that at time T we have $V_T = \xi$. This means that we can *replicate* the claim using the price as an initial endowment. Moreover, the fair price should be the smallest amount V_0 with which this can be done. This principle is the basis of *arbitrage-free pricing* of the claim.

We say that a *hedging strategy* against a non-negative contingent claim ξ is a feasible self-financing strategy (V, π) such that $V_T = \xi$.

With our assumptions, any square-integrable non-negative claim ξ can be hedged, i.e. there is a hedging strategy against ξ.

The *fair price* (at time 0) of the claim ξ is the smallest initial endowment needed to hedge ξ.

Now, with our assumptions, we have the following:

Theorem 2.2 ([12, Theorem 1.1])**.**
Let ξ be a non-negative square-integrable contingent claim. Then there exists a hedging strategy (V, π) against ξ such that

$$dV_t = r_t V_t dt + \pi_t \sigma_t \theta_t dt + \pi_t \sigma_t dW_t, \quad V_T = \xi, \tag{5}$$

and such that V_0 is the fair price of the claim.

Now we can finally write this as a BSDE. It follows from Theorem 2.2 that $(V, \pi\sigma)$ is a solution of the (linear) BSDE

$$Y_t = \xi + \int_t^T f(r, Y_r, Z_r)dr - \int_t^T Z_r dW_r, \quad t \in [0, T], \tag{6}$$

where $f(t, y, z) := r_t y + \theta_t z$.

The model used in Example 2.1 is very simple in this form. However, one can easily incorporate for example borrowing with a higher interest rate than r, consumption, and transaction costs to the BSDE-formulation.

3 Decoupling

In this Section we recall the new general functional mapping procedure that was introduced in [A].

We assume that there are two complete probability spaces where the randomness is, up to nullsets, induced by a countable family of random variables. Moreover, these two families should have the same finite-dimensional distributions. In particular we do not require this distribution to be Gaussian.

Then we can map equivalence classes of random variables from the first space to equivalence classes of random variables in the second space. This is done in such a way that we do not change the structure of the random variable in question.

More generally, the same procedure applies to stochastic processes taking values in spaces of continuous functions.

This procedure can be applied to a Wiener space as a basis for decoupling,[2] and as a factorization through a canonical space. Because our approach is distributional, canonical space here refers to the sequence space $\mathbb{R}^{\mathbb{N}}$, but as a by-product we can also map all random variables and processes to the standard Wiener space $C_0([0, T])$.

Some advantages of this functional mapping procedure are that the approach is robust, but also easy to use. It is robust, since we only assume that the randomness comes from a countable sequence of random variables. It is also easy to use since, as one can see from the results in Section 3.3 below, it preserves the structure of the random objects.

[2] Why are we interested in decoupling? See Section 3.1 below.

The origins of this procedure are in [13], where the L_p-variation of certain BSDEs is considered. One of the assumptions in [13] is, that there is an underlying diffusion $(X_r)_{r\in[0,T]}$ and the terminal value of the BSDE can be written as $\xi = g(X_{t_1}, \dots, X_{t_n})$, where g is a Borel-function and $0 \le t_1 < \dots < t_n \le T$. Moreover, the randomness of the generator f comes only from the diffusion, i.e. we have $f(r, \omega, y, z) = h(r, X_r(\omega), y, z)$ for a measurable function h. These assumptions can be dropped using our decoupling technique.

This decoupling also gives rise to Banach spaces of random variables known as anisotropic Besov spaces (see Section 4). These spaces measure the fractional smoothness of random variables.

3.1 Motivation

We start this Section by considering Example 2.1 in real life. Unfortunately it is not possible to trade continuously in time as is required in Theorem 2.2, but instead you only trade a stock at a finite number of time points $0 = t_0 < t_1 < \dots < t_N = T$, i.e. on a time-grid $\pi := \{t_0, \dots, t_N\}$. Trying to replicate the option with only a finite number of adjustments will most likely fail, but it is of interest to know how large the difference is.

The same phenomenon occurs if one wishes to simulate a (solution of a) BSDE; for the computer the system needs to be discretized. That is, we evaluate our processes only on a finite time-grid π as before. Because of this, there is a difference between the solution and the result of the simulation. This, or some norm of this, is the simulation error. Naturally, it is important to know how large the simulation error is.

To illustrate how the error caused by the time discretization can be estimated, we will follow the approach from [4, Chapter 2]: Let $T = 1$ and consider the FBSDE

$$\begin{aligned} X_t &= x + \int_0^t b(X_r)dr + \int_0^t \sigma(X_r)dW_r, \quad t \in [0,T], \\ Y_t &= g(X_T) + \int_t^T h(X_r, Y_r, Z_r)dr - \int_t^T Z_r dW_r, \quad t \in [0,T], \end{aligned} \tag{7}$$

where $x \in \mathbb{R}^d$, and the functions $b : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d}$, $g : \mathbb{R}^d \to \mathbb{R}$, $h : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ are assumed to be Lipschitz-continuous. Given a time-grid π as before, with modulus

$$|\pi| := \sup_{k=0,\dots,N-1} (t_{k+1} - t_k),$$

the Euler scheme of the forward process X is defined as

$$X_0^\pi = x, \quad X_{t_{k+1}}^\pi = b(X_{t_k}^\pi)\Delta t_k + \sigma(X_{t_k}^\pi)\Delta W_{t_k}$$

for $k < N$, where $\Delta t_k = t_{k+1} - t_k$ and $\Delta W_{t_k} = W_{t_{k+1}} - W_{t_k}$. To motivate the definition of the *backward Euler scheme* we first write

$$Y_{t_k} = Y_{t_{k+1}} + \int_{t_k}^{t_{k+1}} h(X_r, Y_r, Z_r)dr - \int_{t_k}^{t_{k+1}} Z_r dW_r,$$

and then formally approximate the righthand-side to arrive at:

$$Y_{t_k} \approx Y_{t_{k+1}} + h(X_{t_k}^\pi, Y_{t_k}, Z_{t_k})\Delta t_k - Z_{t_k}\Delta W_{t_k}. \tag{8}$$

First, by taking conditional expectation given $\mathcal{F}_{t_k}$ on both sides of (8), we get that

$$Y_{t_k} \approx \mathbb{E}\left[Y_{t_{k+1}} \,\middle|\, \mathcal{F}_{t_k}\right] + h(X_{t_k}^\pi, Y_{t_k}, Z_{t_k})\Delta t_k.$$

Secondly, by multiplying both sides of (8) by ΔW_{t_k}, and then taking conditional expectation given $\mathcal{F}_{t_k}$, we get by Itô's isometry that

$$0 \approx \mathbb{E}\left[Y_{t_{k+1}}\Delta W_{t_k} \,\middle|\, \mathcal{F}_{t_k}\right] - Z_{t_k}\Delta t_k.$$

These steps lead to the (implicit) backward Euler scheme

$$\overline{Z}_{t_k}^\pi = \frac{1}{\Delta t_k}\mathbb{E}\left[Y_{t_{k+1}}^\pi \Delta W_{t_k} \,\middle|\, \mathcal{F}_{t_k}\right], \tag{9}$$

$$Y_{t_k}^\pi = \mathbb{E}\left[Y_{t_{k+1}}^\pi \,\middle|\, \mathcal{F}_{t_k}\right] + h(X_{t_k}^\pi, Y_{t_k}^\pi, \overline{Z}_{t_k}^\pi)\Delta t_k, \tag{10}$$

with $k < N$, and $Y_T^\pi = g(X_T^\pi)$. We define the (squared) *simulation error* as

$$\mathrm{Err}(\pi)^2 := \max_{0\le k<N} \mathbb{E}\left[\sup_{r\in[t_k,t_{k+1}]} |Y_r - Y_{t_k}^\pi|^2\right] + \mathbb{E}\left[\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}} |Z_r - \overline{Z}_{t_k}^\pi|^2 dr\right].$$

Now, letting

$$\overline{Z}_{t_k} = \frac{1}{\Delta t_k}\mathbb{E}\left[\int_{t_k}^{t_{k+1}} Z_r dr \,\middle|\, \mathcal{F}_{t_k}\right],$$

we define the (squared) *modulus of regularity* of Y and Z as

$$\mathcal{R}_Y^2(\pi) \;:=\; \max_{0\le k<N} \mathbb{E}\left[\sup_{r\in[t_k,t_{k+1}]} |Y_r - Y_{t_k}|^2\right], \tag{11}$$

$$\mathcal{R}_Z^2(\pi) \;:=\; \mathbb{E}\left[\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}} |Z_r - \overline{Z}_{t_k}|^2 dr\right]. \tag{12}$$

Then it follows ([4, Proposition 2.2.1]) that there exists a constant $C > 0$, independent of π, such that

$$\mathrm{Err}(\pi) \le C\left(|\pi| + \mathcal{R}_Y^2(\pi) + \mathcal{R}_Z^2(\pi)\right)^{\frac{1}{2}}.$$

Moreover, in the case that $h \equiv 0$ it even follows ([4, Remark 2.2.4]) that there exists a constant $c > 0$ such that

$$c\left(\mathcal{R}_Y^2(\pi) + \mathcal{R}_Z^2(\pi)\right)^{\frac{1}{2}} \le \mathrm{Err}(\pi),$$

and up to a term depending on $|\pi|$ this holds also when $h \not\equiv 0$.

The above example indicates that the simulation error can be approximated by regularity of the exact solution itself. This is one reason why we want to find upper bounds of $|Y_t - Y_s|$ in some sense for all $0 \le s < t \le T$. It is also an interesting task in itself. In [A] we consider the L_p-quantity $\mathbb{E}|Y_t - Y_s|^p$ for $2 \le p < \infty$, and in [B] we consider the conditional L_p-quantity $\mathbb{E}^{\mathcal{F}_s}|Y_t - Y_s|^p$ for $2 \le p < \infty$.

The strategy in both cases was the same: we start with

$$\mathbb{E}^{\mathcal{F}_s}|Y_t - Y_s|^p \le c_p\left[\mathbb{E}^{\mathcal{F}_s}|Y_t - \mathbb{E}^{\mathcal{F}_s}Y_t|^p + \mathbb{E}^{\mathcal{F}_s}|\mathbb{E}^{\mathcal{F}_s}Y_t - Y_s|^p\right], \tag{13}$$

where $c_p > 0$. With $p = 2$ this inequality is actually an equality, with $c_2 = 1$. The second part in equation (13), $\mathbb{E}^{\mathcal{F}_s}|\mathbb{E}^{\mathcal{F}_s}Y_t - Y_s|^p$, can be estimated directly (see Theorem 6.5) using estimates that are mostly standard. The only non-standard argument, an extension of Fefferman's inequality (see Proposition 5.4 or [B, Corollary 2.10]), is used if f grows superlinearly in the z-variable.

Now we focus on the more difficult problem, upper bounding $\mathbb{E}^{\mathcal{F}_s}|Y_t - \mathbb{E}^{\mathcal{F}_s}Y_t|^p$. To explain the idea, consider first the case $s = 0$. Then, taking X_t to be an independent copy of Y_t, we have

$$\frac{1}{2^p}\mathbb{E}|Y_t - X_t|^p \le \mathbb{E}|Y_t - \mathbb{E}Y_t|^p \le \mathbb{E}|Y_t - X_t|^p, \tag{14}$$

where $p \ge 1$.

Naturally, we want X_t to be such that we can (without too much of an extra effort) upper bound the quantity $\mathbb{E}|Y_t - X_t|^p$ mentioned above. If X_t is indeed a copy of Y_t, then X_t itself should be a solution of a BSDE. This is essential for us; both X_t and Y_t are solutions of BSDEs at time t, so we can also write their difference $X_t - Y_t$ as a solution of a BSDE at time t. After this we use an apriori estimate (similar to equation (2)) to find an upper bound in terms of the data.

To indicate how we can handle the case $s > 0$, recall that our filtration is generated by the Brownian motion W, so it makes sense to first consider the case $Y_t = W_t$ (for simplicity we may think that W is 1-dimensional). In the case $s = 0$ we took X_t to be an independent copy of Y_t, and this leads us to assume that there exists $(W'_r)_{r\in[0,T]}$ which is a Brownian motion that is independent of W.

Let now $s > 0$, and let us try to use W' to find a random variable X_t such that

$$\frac{1}{2^p}\mathbb{E}^{\mathcal{F}_s}|W_t - X_t|^p \le \mathbb{E}^{\mathcal{F}_s}|W_t - \mathbb{E}^{\mathcal{F}_s}W_t|^p \le \mathbb{E}^{\mathcal{F}_s}|W_t - X_t|^p.$$

We simplify the setting once more by considering the case $p = 2$. Now, using properties of the Brownian motion, we have

$$\begin{aligned}\mathbb{E}^{\mathcal{F}_s}|W_t - \mathbb{E}^{\mathcal{F}_s}W_t|^2 &= \mathbb{E}^{\mathcal{F}_s}|W_t - W_s|^2 \\ &= \frac{1}{2}\mathbb{E}^{\mathcal{F}_s}|W_t - (W'_t - W'_s + W_s)|^2,\end{aligned}$$

where we used the fact that W and W' are independent. This leads us to choose $X_t = W'_t - W'_s + W_s$, or in terms of stochastic integrals,

$$X_t = W_t^{(s,t]} := \int_0^t 1 - 1_{(s,t]}(r)dW_r + \int_0^t 1_{(s,t]}(r)dW'_r.$$

Then X_t is a *conditionally independent copy of* W_t *given* $\mathcal{F}_s$.

For illustration we have a figure of the different Brownian motions:

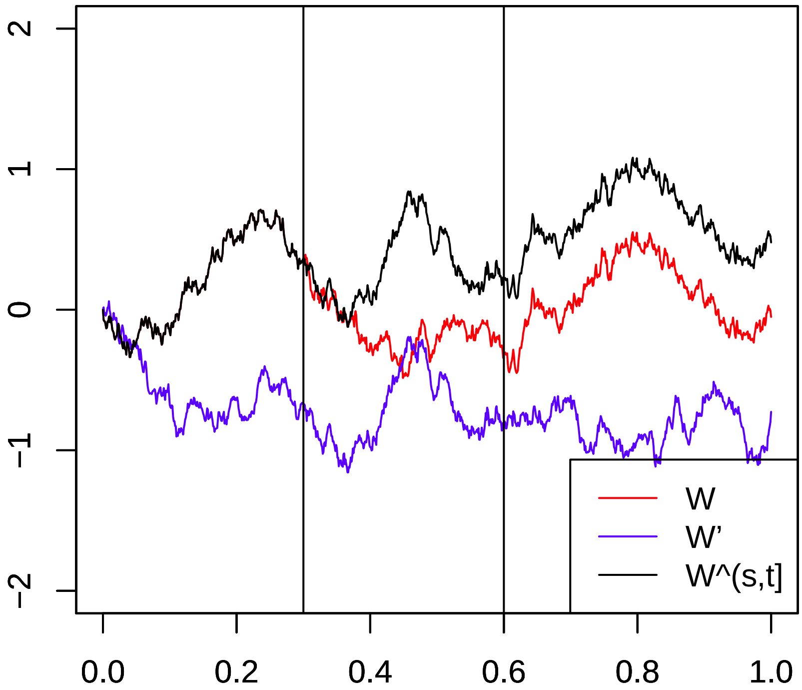

Figure 1: Brownian motions W,W' and $W^{(s,t]}$. Here $s = 0.3$, $t = 0.6$ and $T = 1$.

Can we do the same for any $\mathcal{F}_t$-measurable Y_t? Recall that $\mathcal{F}_t$ is (essentially) generated by the Brownian motion W until time t. This means that the randomness of Y_t comes, in some sense, from the underlying process $(W_r)_{r\in[0,t]}$. Next we explain how we can change the random variable Y_t by changing the underlying Brownian motion W into the new Brownian motion $W^{(s,t]}$.

3.2 Setting

Let us be precise on how we can find a Brownian motion W' that is independent of W. We start by taking another stochastic basis $(\Omega', \mathcal{F}', \mathbb{P}', (\mathcal{F}'_r)_{r\in[0,T]})$, where $(\mathcal{F}'_r)_{r\in[0,T]}$ is the augmented filtration of the d-dimensional Brownian motion W', and $\mathcal{F}' = \mathcal{F}'_T$.

To speak of independence, we must have W and W' defined on the $s$ame probability space. Thus, we let

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

and work on the probability space $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}})$. By considering the canonical extensions $W(\omega, \omega') := W(\omega)$ and $W'(\omega, \omega') := W'(\omega')$ for all $(\omega, \omega') \in \Omega \times \Omega'$, we then have that W and W', both defined on $\overline{\Omega}$, are independent d-dimensional Brownian motions. The augmented natural filtration[3] of the $2d$-dimensional Brownian motion $\overline{W} = (W, W')$ is denoted by $(\overline{\mathcal{F}}_t)_{t\in[0,T]}$. We also use the notation

$$\begin{aligned} (\Omega_T, \Sigma_T, \mathbb{P}_T) &:= ([0,T] \times \Omega, \mathcal{B}([0,T]) \otimes \mathcal{F}, \frac{\lambda}{T} \times \mathbb{P}), \\ (\overline{\Omega}_T, \overline{\Sigma}_T, \overline{\mathbb{P}}_T) &:= ([0,T] \times \overline{\Omega}, \mathcal{B}([0,T]) \otimes \overline{\mathcal{F}}, \frac{\lambda}{T} \times \overline{\mathbb{P}}), \end{aligned}$$

where λ is the Lebesgue-measure on $[0,T]$.

Now we let $\varphi : [0,T] \to [0,1]$ be Borel measurable, and define for all $r \in [0,T]$

$$W_r^\varphi = \int_0^r \sqrt{1 - \varphi^2(u)} dW_u + \int_0^r \varphi(u) dW'_u.$$

Because of Lévy's characterization of the Brownian motion, $(W_r^\varphi)_{r\in[0,T]}$ is a standard Brownian motion, and we use $(\mathcal{F}_r^\varphi)_{r\in[0,T]}$ to denote its augmented natural filtration and $\mathcal{F}^\varphi := \mathcal{F}_T^\varphi$. The predictable σ-algebra on the stochastic basis $(\overline{\Omega}, \mathcal{F}^\varphi, \overline{\mathbb{P}}, (\mathcal{F}_r^\varphi)_{r\in[0,T]})$ is denoted by $\mathcal{P}^\varphi$, and we will also make use of the notation $\Sigma_T^\varphi = \mathcal{B}([0,T]) \otimes \mathcal{F}^\varphi$.

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 $\overline{\mathbb{P}}$-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 the notation W^0 for the extension of W and $(\mathcal{F}_r^0)_{r\in[0,T]}$ for the corresponding filtration.

For our purposes the Brownian motion $W^{(s,t]} := W^{\chi_{(s,t]}}$, where $0 \le s < t \le T$, plays an essential role.

One may view the definition of W^φ, and the soon-to-be introduced techniques, as an extension of the techniques used to obtain Mehler's formula (see for example [25, Equation (1.67)]). The same technique that was used to obtain Mehler's formula has also been used in characterizing (isotropic) Besov spaces obtained by real interpolation in [18].

[3] Whenever we augment a filtration in $\overline{\Omega}$, we augment it by $\overline{\mathbb{P}}$-nullsets.

The difference between our definition and these cases is, that they only considered constant-valued functions φ. The reason why we want to have a covariance-function that does depend on time is, that we want to estimate $\mathbb{E}^{\mathcal{F}_s}|Y_t - \mathbb{E}^{\mathcal{F}_s}Y_t|^p$.

Now we describe how the randomness of a process $X \in \mathcal{L}_0(\Omega_T, \Sigma_T, \mathbb{P}_T)$ is changed to come from W^φ instead of W. Changing the randomness of a random variable $\xi \in \mathcal{L}_0(\Omega, \mathcal{F}, \mathbb{P})$ to come from W^φ instead of W is analogous.

1. For $X \in \mathcal{L}_0(\Omega_T, \Sigma_T, \mathbb{P}_T)$ take the canonical extension $X(t, \omega, \omega') := X(t, \omega)$, and consider the corresponding equivalence class of random variables $[X] \in L_0(\overline{\Omega}_T, \Sigma_T^0, \overline{\mathbb{P}}_T)$.

2. [4]Letting $(g_n)_{n\in\mathbb{N}} : \overline{\Omega} \to \mathbb{R}$ be the family of finite differences of W^0 generated by Haar functions, there exists a $\mathcal{B}([0,T]) \otimes \sigma(g_n, n \in \mathbb{N})$-measurable $X^0 \in [X]$.

3. [5]Defining $J_T : \overline{\Omega}_T \to [0,T] \times \mathbb{R}^{\mathbb{N}}, \quad J_T(t,\eta) = (t, g_n(\eta))_{n\in\mathbb{N}}$, there exists a random variable $\hat{X} : [0,T] \times \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ so that X^0 can be factorized through $[0,T] \times \mathbb{R}^{\mathbb{N}}$:

$$X^0 : \overline{\Omega}_T \xrightarrow{J_T} [0,T] \times \mathbb{R}^{\mathbb{N}} \xrightarrow{\hat{X}} \mathbb{R}.$$

4. [6] Defining J_T^φ analogously, but using W^φ instead of W^0, we have that $\hat{X}(J_T^\varphi) : \overline{\Omega}_T \to \mathbb{R}$ is a well-defined $\mathcal{B}([0,T]) \otimes \sigma(g_n^\varphi, n \in \mathbb{N})$-measurable random variable.

5. Finally, we let $X^\varphi \in L_0(\overline{\Omega}_T, \Sigma_T^\varphi, \overline{\mathbb{P}}_T)$ be the equivalence class that contains all Σ_T^φ-measurable random variables that are $\overline{\mathbb{P}}_T$-a.s. the same as $\hat{X}(J_T^\varphi)$.

In step 2 it is essential that the σ-algebras $\mathcal{F}^0$ and $\sigma(g_n, n \in \mathbb{N})$ differ only by nullsets. Following steps 1-5 we now have the well-defined functional mappings

$$\mathcal{C}_T : L_0(\overline{\Omega}_T, \Sigma_T^0, \overline{\mathbb{P}}_T) \to L_0(\overline{\Omega}_T, \Sigma_T^\varphi, \overline{\mathbb{P}}_T), \quad \mathcal{C}_T([X]) := X^\varphi,$$

and, analogously, $\mathcal{C}_0 : L_0(\overline{\Omega}, \mathcal{F}^0, \overline{\mathbb{P}}) \to L_0(\overline{\Omega}, \mathcal{F}^\varphi, \overline{\mathbb{P}})$.

[4][A, Lemma 2.2]
[5][A, Lemma 2.3(1)]
[6][A, Lemma 2.3(2)]

In fact, in [A] we start with a complete probability space $(\Omega^1, \mathcal{F}^1, \mathbb{P}^1)$, and we assume:

- There exists a sequence $(\xi_n^1)_{n\in\mathbb{N}}$ of random variables in Ω^1 such that $\mathcal{F}^1 = \sigma(\xi_n^1, n \in \mathbb{N}) \vee \mathcal{N}^1$, where $\mathcal{N}^1$ are the $\mathbb{P}^1$-nullsets,

- There exists a complete probability space $(\Omega^2, \mathcal{F}^2, \mathbb{P}^2)$ such that $\mathcal{F}^2 = \sigma(\xi_n^2, n \in \mathbb{N}) \vee \mathcal{N}^2$, where $\mathcal{N}^2$ are the $\mathbb{P}^2$-nullsets, and $(\xi_n^2)_{n\in\mathbb{N}}$ is a sequence of random variables in Ω^2,

- $(\xi_n^1)_{n\in\mathbb{N}}$ and $(\xi_n^2)_{n\in\mathbb{N}}$ have the same finite-dimensional distributions.

Then we follow steps 1-5, with $(g_n)_{n\in\mathbb{N}}$ replaced by $(\xi_n^1)_{n\in\mathbb{N}}$ and $(g_n^\varphi)_{n\in\mathbb{N}}$ by $(\xi_n^2)_{n\in\mathbb{N}}$, thus defining the functional mappings

$$\mathcal{C}_T : L_0(\Omega_T^1, \Sigma_T^1, \mathbb{P}_T^1) \to L_0(\Omega_T^2, \Sigma_T^2, \mathbb{P}_T^2)$$

and $\mathcal{C}_0 : L_0(\Omega^1, \mathcal{F}^1, \mathbb{P}^1) \to L_0(\Omega^2, \mathcal{F}^2, \mathbb{P}^2)$.

However, we restrict here ourselves to the mentioned case $(\Omega^1, \mathcal{F}^1, \mathbb{P}^1) = (\overline{\Omega}, \mathcal{F}^0, \overline{\mathbb{P}})$, and $(\Omega^2, \mathcal{F}^2, \mathbb{P}^2) = (\overline{\Omega}, \mathcal{F}^\varphi, \overline{\mathbb{P}})$. Then, for $X \in L_0(\overline{\Omega}_T, \Sigma_T^0, \overline{\mathbb{P}}_T)$, we have that $\mathcal{C}_T(X)$ and X are equivalence classes that consist of stochastic processes that are in the *same* probability space $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}})$.

In particular, this approach does not require continuous paths or a gaussian distribution. As such, the approach might be useful also in other situations. It should also be mentioned that a similar distributional approach is used in [23, Chapter V.1.6] (see also [20]) on Gaussian random variables to define the Gaussian Sobolev spaces (or the Malliavin Sobolev spaces).

This approach being so flexible and "general", the question arises whether it is strong enough to preserve some regularity structures of the random variables and stochastic processes. So far we have only positive answers.

3.3 Properties

We present some natural and expected properties that make our method applicable in various situations, for example with BSDEs.

To shorten the presentation, we sometimes work in Ω and Ω_T in parallel. This is done by considering $S \in \{0, T\}$, and using the notation $\Omega_0 := \Omega$ and $\Sigma_0 := \mathcal{F}$.

Also, we fix $\mathbb{X} \neq \emptyset$ to be a complete metric space that is locally σ-compact, i.e. there exist compact subsets $\emptyset \neq K_1 \subseteq K_2 \subseteq \ldots$, such that $\overline{\mathring{K}_n} = K_n$ and $\mathbb{X} = \bigcup_{n=1}^{\infty} \mathring{K}_n$.

The basic form of the results in this Section is the following: If the original random object $X \in \mathcal{L}_0(\Omega_S, \Sigma_S, \mathbb{P}_S)$ satisfies some property, then in the *equivalence class* $X^\varphi \in L_0(\overline{\Omega}_S, \Sigma_S^\varphi, \overline{\mathbb{P}}_S)$ we can find a *representative* that satisfies, in the proper sense, the same property.

Definition 3.1. *We let $f \in \mathcal{L}_0(\Omega_S, \Sigma_S, \mathbb{P}_S; C(\mathbb{X}))$ if $f : \Omega_S \times \mathbb{X} \to \mathbb{R}$ is such that*

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

The continuity of a stochastic process is preserved in the following sense:

Proposition 3.2 ([A, Lemma 2.9])**.**
If $f \in \mathcal{L}_0(\Omega_S, \Sigma_S, \mathbb{P}_S; C(\mathbb{X}))$, then there exists $f^\varphi \in \mathcal{L}_0(\overline{\Omega}_S, \Sigma_S^\varphi, \overline{\mathbb{P}}_S; C(\mathbb{X}))$ such that $f^\varphi(y) \in (f(y))^\varphi$ for all $y \in \mathbb{X}$. Given f_1^φ and f_2^φ with these two properties, it follows that $f_1^\varphi(\cdot) = f_2^\varphi(\cdot)$ ($\overline{\mathbb{P}}_S$-a.s.).

Given $f \in \mathcal{L}_0(\Omega_S, \Sigma_S, \mathbb{P}_S; C(\mathbb{X}))$, Proposition 3.2 is used to define f^φ as the equivalence class[7] of elements in $\mathcal{L}_0(\overline{\Omega}_S, \Sigma_S^\varphi, \overline{\mathbb{P}}_S; C(\mathbb{X}))$ such that $f^\varphi(y) \in (f(y))^\varphi$ for all $y \in \mathbb{X}$.

Predictability and adaptedness are transferred in the following sense:

Proposition 3.3 ([A, Lemma 3.1 and Theorem 2.8])**.**

(i) *If $\xi \in \mathcal{L}_0(\Omega, \mathcal{F}_t, \mathbb{P})$ for some $t \in [0, T]$, then all representatives of $\xi^\varphi \in L_0(\overline{\Omega}, \mathcal{F}^\varphi, \overline{\mathbb{P}})$ are $\mathcal{F}_t^\varphi$-measurable.*

(ii) *If $f \in \mathcal{L}_0(\Omega_T, \mathcal{P}, \mathbb{P}_T; C(\mathbb{X}))$[8], then there is a $\mathcal{P}^\varphi$-measurable[9] representative of $f^\varphi \in L_0(\overline{\Omega}_T, \Sigma_T^\varphi, \overline{\mathbb{P}}_T; C(\mathbb{X}))$.*

(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 $Y^\varphi \in L_0(\overline{\Omega}, \mathcal{F}^\varphi, \overline{\mathbb{P}}; C([0, T]))$ are $(\mathcal{F}_t^\varphi)_{t \in [0,T]}$-adapted.*

We summarize some further properties proven in [A]:

[7] We identify $f_1, f_2 \in \mathcal{L}_0(\overline{\Omega}_S, \Sigma_S^\varphi, \overline{\mathbb{P}}_S; C(\mathbb{X}))$ if $\overline{\mathbb{P}}_S\left(f_1(y) = f_2(y) \text{ for all } y \in \mathbb{X}\right) = 1$.

[8] i.e. $\eta \mapsto f(\eta, x)$ is $\mathcal{P}$-measurable for all $x \in \mathbb{X}$.

[9] i.e. $\eta \mapsto f^\varphi(\eta, x)$ is $\mathcal{P}^\varphi$-measurable for all $x \in \mathbb{X}$.

Proposition 3.4 ([A, Theorems 2.6, 2.11, and Lemma 3.2]).
Let $N \in \mathbb{N}$, $g : \mathbb{R}^N \to \mathbb{R}$ a Borel function, $S \in \{0, T\}$, $X, X_1, \ldots, X_N \in \mathcal{L}_0(\Omega_S, \Sigma_S, \mathbb{P}_S)$, $Y \in \mathcal{L}_1(\Omega_T, \Sigma_T, \mathbb{P}_T)$, $f \in \mathcal{L}_0(\Omega_S, \Sigma_S, \mathbb{P}_S; C(\mathbb{R}^N))$, and $Z \in \mathcal{L}_2(\Omega_T, \mathcal{P}, \mathbb{P}_T)$. Then

(i) $X \overset{d}{=} X^\varphi$.

(ii) $(g(X_1, \ldots, X_N))^\varphi = g(X_1^\varphi, \ldots, X_N^\varphi)$.

(iii) $(f(X_1, \ldots, X_N))^\varphi = f^\varphi(X_1^\varphi, \ldots, X_N^\varphi)$.

(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.[10]

Our next result states that if we have a strong solution of an SDE in the first space, then changing the randomness of the solution results into a strong solution of another SDE. Naturally, the randomness of the data is changed. It is noteworthy that we do not assume uniqueness of the solution.

Proposition 3.5 ([A, Theorem 3.3]).
Assume that $f, g_i \in \mathcal{L}_0(\Omega_T, \mathcal{P}, \mathbb{P}_T; C(\mathbb{R}^{1+d}))$, $Z_i \in \mathcal{L}_0(\Omega_T, \mathcal{P}, \mathbb{P}_T)$, $i = 1, \ldots, d$, 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

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

If $\xi \in \mathcal{L}_0(\Omega, \mathcal{F}, \mathbb{P})$ and

$$Y_u = \xi + \int_u^T f(r, Y_r, Z_r) dr - \int_u^T g(r, Y_r, Z_r) dW_r \tag{15}$$

for $u \in [0, T]$, $\mathbb{P}$-a.s., then fixing any predictable representatives of $f^\varphi, g_i^\varphi, Z_i^\varphi$, and $(\mathcal{F}_t^\varphi)_{t \in [0,T]}$-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

$$Y_u^\varphi = \xi^\varphi + \int_u^T f^\varphi(r, Y_r^\varphi, Z_r^\varphi) dr - \int_u^T g^\varphi(r, Y_r^\varphi, Z_r^\varphi) dW_r^\varphi \tag{16}$$

for $u \in [0, T]$, $\overline{\mathbb{P}}$-a.s.

[10] By Proposition 3.3(ii) there exists such a representative.

We want to obtain conditional estimates in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ from estimates obtained using the coupling technique. Here the next result is vital:

Lemma 3.6 ([B, Lemma 4.9])**.**
Let $p \geq 1$, $0 \leq s < t \leq T$, $\xi \in \mathcal{L}_p(\overline{\Omega}, \mathcal{F}^0, \overline{\mathbb{P}})$, and put

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

Then

$$\frac{1}{2^p}\mathbb{E}^{\mathcal{F}_s^0}|\xi - \xi^{(s,t]}|^p \leq \mathbb{E}^{\mathcal{F}_s^0}|\xi - \mathbb{E}^{\mathcal{G}_s^t}\xi|^p \leq \mathbb{E}^{\mathcal{F}_s^0}|\xi - \xi^{(s,t]}|^p. \tag{17}$$

Moreover, if ξ is $\mathcal{F}_t^0$-measurable, then

$$\frac{1}{2^p}\mathbb{E}^{\mathcal{F}_s^0}|\xi - \xi^{(s,t]}|^p \leq \mathbb{E}^{\mathcal{F}_s^0}|\xi - \mathbb{E}^{\mathcal{F}_s^0}\xi|^p \leq \mathbb{E}^{\mathcal{F}_s^0}|\xi - \xi^{(s,t]}|^p.$$

Remark 3.7. *Recall what we discussed in Section 3.1: our goal is to upper bound $\mathbb{E}^{\mathcal{F}_s}|Y_t - \mathbb{E}^{\mathcal{F}_s}Y_t|^p$, where Y_t is a solution of a BSDE at time t. Lemma 3.6 tells us that*

$$\mathbb{E}^{\mathcal{F}_s}|Y_t - \mathbb{E}^{\mathcal{F}_s}Y_t|^p \leq \mathbb{E}^{\mathcal{F}_s^0}|Y_t - Y_t^{(s,t]}|^p \quad \overline{\mathbb{P}}\text{-a.s.},$$

where $\mathbb{E}^{\mathcal{F}_s}|Y_t - \mathbb{E}^{\mathcal{F}_s}Y_t|^p$ stands for the canonical extension. Using Proposition 3.5 we notice that $Y^{(s,t]}$ solves a BSDE just like Y did. We again emphasize that we only need $(y,z) \mapsto f(r,\omega,y,z)$ to be continuous for all $(r,\omega) \in \Omega_T$, and we need to know that Y is a solution of the BSDE. In particular the solution need not be unique, the existence is enough.

Now it remains to use an apriori estimate to upper bound $\mathbb{E}^{\mathcal{F}_s^0}|Y_t - Y_t^{(s,t]}|^p$. For this we need to introduce a lot of assumptions that guarantee for us a natural upper bound. To understand some of these assumptions better, it is best to introduce the concept of bounded mean oscillation (BMO) before we state our main results concerning BSDEs. But first we show what else can be done with our decoupling technique; we introduce anisotropic Besov spaces.

4 Anisotropic Besov Spaces

In this Section we recall the family of Banach spaces describing functional fractional smoothness of random variables, as introduced in [A].

To define these spaces we only need to use the decoupling technique from Section 3. Although we do not use Malliavin Calculus, in Theorem 4.9 a certain anisotropic Besov space is characterized using Malliavin derivatives.

We mention also some connections to previously studied Besov spaces obtained by real interpolation.

We first define the metric space of parameter functions that we use:

Definition 4.1. *We define the metric space (Δ, δ) as the equivalence classes of the pseudo-metric space $(\mathfrak{D}, \delta)$ with*

$$\mathfrak{D} = \{\psi \in \mathcal{L}_2((0,T]) : 0 \le \psi \le 1\} \quad \text{and} \quad \delta(\varphi, \psi) = \|\varphi - \psi\|_{L_2((0,T])}.$$

Let $p \in (0, \infty)$, $\xi \in \mathcal{L}_p(\Omega, \mathcal{F}, \mathbb{P})$, and $\varphi \in \Delta$. Then we exploit our notion of decoupling to study "the sensitivity of ξ to the direction φ", by measuring the map $F_{\xi,p}(\varphi) = \|\xi - \xi^\varphi\|_{L_p(\overline{\Omega})}$ in different ways. First we note that $\varphi \to \|\xi - \xi^\varphi\|_p$ is a continuous map:

Lemma 4.2 ([A, Lemma 4.9])**.** *For $p \in (0, \infty)$ and $\xi \in \mathcal{L}_p(\Omega, \mathcal{F}, \mathbb{P})$ the map $F_{\xi,p} : \Delta \to [0, \infty)$ defined by $F_{\xi,p}(\varphi) = \|\xi - \xi^\varphi\|_p$ is continuous.*

Next we define the concept of anisotropic Besov spaces. This is done by measuring the map $F_{\xi,p}$ using an admissible functional:

Definition 4.3. *Let $C^+(\Delta)$ be the space of all non-negative continuous functions $F : \Delta \to [0, \infty)$. A functional $\Phi : C^+(\Delta) \to [0, \infty]$ is called* admissible *provided that*

$(C1)$ $\Phi(F + G) \le \Phi(F) + \Phi(G)$,

$(C2)$ $\Phi(\lambda F) = \lambda \Phi(F)$ *for* $\lambda \ge 0$,

$(C3)$ $\Phi(F) \le \Phi(G)$ *for* $0 \le F \le G$,

$(C4)$ $\Phi(F) \le \limsup_n \Phi(F_n)$ *for* $\sup_{\varphi \in \Delta} |F_n(\varphi) - F(\varphi)| \to_n 0$.

Definition 4.4. For $p \in (0, \infty)$, $\xi \in \mathcal{L}_p(\Omega)$, and an admissible $\Phi : C^+(\Delta) \to [0, \infty]$ we let $\xi \in \mathbb{B}_p^\Phi$ provided that $\Phi(\varphi \to \|\xi - \xi^\varphi\|_p) < \infty$ and set

$$\|\xi\|_{\mathbb{B}_p^\Phi} := \left[\mathbb{E}|\xi|^p + \|\xi\|_{\Phi,p}^p\right]^{\frac{1}{p}} \quad \text{with} \quad \|\xi\|_{\Phi,p} := \Phi(\varphi \to \|\xi - \xi^\varphi\|_p).$$

This definition yields a Banach space.

Theorem 4.5 ([A, Proposition 4.13])**.**
For $p \in [1, \infty)$ the space (of equivalence classes) $\mathbb{B}_p^\Phi$ is a Banach space.

4.1 Interpolation spaces as (an)isotropic Besov spaces

We give two examples of admissible functionals that correspond to Besov spaces obtained by real interpolation method.

Definition 4.6. *For* $0 = r_0 < r_1 < \cdots r_L = T$, $\theta_l \in (0,1)$, $q_l \in [1,\infty]$, *and* $F \in C^+(\Delta)$ *we let*

$$\Phi^{(\theta_1,q_1),\ldots,(\theta_L,q_L)}_{r_1,\ldots,r_L}(F) := \sup_{l=1,\ldots,L} \left\| (r_l - t)^{-\theta_l/2} F(\chi_{(t,r_l]}) \right\|_{L_{q_l}([r_{l-1},r_l), \frac{dt}{r_l - t})}.$$

Using the notation γ_d for the standard d-dimensional gaussian distribution, the spaces $\mathbb{B}^\theta_{p,q}(\mathbb{R}^d,\gamma_d)$ are interpolation spaces between L_p and the Malliavin Sobolev space $\mathbb{D}_{1,p}$. Here $\theta \in (0,1)$ describes the fractional smoothness, and $q \in [1,\infty]$ is the finetuning-index. For more information see for example [17] and [A, Sections 7.1-7.2].

Proposition 4.7 ([A, Proposition 4.16])**.**
For $\theta \in (0,1)$, $p \in [2,\infty)$, $q \in [1,\infty]$, *and* $f(W_1) \in \mathcal{L}_p$ *one has*

$$f \in \mathbb{B}^\theta_{p,q}(\mathbb{R}^d,\gamma_d) \quad \textit{if and only if} \quad f(W_1) \in \mathbb{B}_p^{\Phi_1^{(\theta,q)}}.$$

Definition 4.8. *Let* $K : [0,1] \to \mathbb{R}$ *be non-negative and Borel-measurable,* $q \in [1,\infty)$, *and let* $\varphi_r : (0,T] \to \mathbb{R}$ *be given by*

$$\varphi_r \equiv r \quad \textit{for} \quad r \in [0,1].$$

Then we define

$$\Phi^{(K,q)}(F) := \left(\int_0^1 K(r) |F(\varphi_r)|^q dr \right)^{\frac{1}{q}}.$$

The definition above means that we use the map $\xi \to \xi^{\varphi_r}$ that exchanges in an isotropic way the full Brownian motion W by its mixture $\sqrt{1-r^2}W + rW'$. Using the notation

$$\xi(W) = \xi \quad \text{and} \quad \xi(\sqrt{1-r^2}W + rW') = \xi^{\varphi_r}$$

this yields to the expression

$$\left(\int_0^1 K(r) \|\xi(W) - \xi(\sqrt{1-r^2}W + rW')\|_p^q dr \right)^{\frac{1}{q}}.$$

Using the particular kernel

$$K(r) := \frac{2r}{1-r^2}\left(\ln\frac{1}{1-r^2}\right)^{-1-\frac{\theta q}{2}},$$

for $\theta \in (0,1)$ and $p = q \in (1,\infty)$ this gives

$$\int_0^\infty t^{-1-\frac{\theta p}{2}} \|\xi(W) - \xi(e^{-\frac{t}{2}}W + \sqrt{1-e^{-t}}W')\|_p^p dt.$$

Spaces based on this type of expression were considered in [18, Remark on p. 428] and identified as interpolation spaces. The same idea was also used in [25, Section 1.4.1] to characterize the Ornstein-Uhlenbeck semigroup.

4.2 Connection to Malliavin derivatives

Here we give an example of another anisotropic Besov space, and characterize it using Malliavin derivatives.

We study the space $\mathbb{B}_p^{\Phi_2}$, where the functional $\Phi_2 : C^+(\Delta) \to [0,\infty]$ is given by

$$\Phi_2(F) := \sup_{0\le s<t\le T} \frac{F(\chi_{(s,t]})}{\sqrt{t-s}}. \tag{18}$$

This means that the Besov space $\mathbb{B}_p^{\Phi_2}$ consists of $\xi \in L_p(\Omega,\mathcal{F},\mathbb{P})$ such that $\sup_{0\le s<t\le T} \frac{\|\xi-\xi^{(s,t]}\|_p}{\sqrt{t-s}} < \infty$. In other words, $\xi \in \mathbb{B}_p^{\Phi_2}$ if and only if there exists $C > 0$ such that

$$\|\xi - \xi^{(s,t]}\|_p \le C\sqrt{t-s}$$

for all $0 \le s < t \le T$. In a sense ξ is Lipschitz.

To describe these spaces we let, for $\xi \in \mathbb{D}_{1,2}$ and D being the Malliavin derivative operator,

$$\begin{aligned}
\|D\xi\|_{L_\infty([0,T];L_p(\Omega))} &:= \operatorname{esssup}_{s\in[0,T]} \|D_s\xi\|_p, \\
\|D\xi\|_{L_p^*(\Omega;L_2([0,T]))} &:= \sup_{0\le a<b\le T} \left\| \left(\frac{1}{b-a}\int_a^b |D_s\xi|^2 ds\right)^{\frac{1}{2}} \right\|_p.
\end{aligned}$$

By the Lebesgue differentiation theorem one has that

$$\|D\xi\|_{L_\infty([0,T];L_2(\Omega))} = \|D\xi\|_{L_2^*(\Omega;L_2([0,T]))}.$$

To formulate our main result, the notation $A \sim_c B$, where $A, B \ge 0$ and $c \ge 1$, stands for $(1/c)A \le B \le cA$. Our main result is that $\mathbb{B}_2^{\Phi_2} \subsetneq \mathbb{D}_{1,2}$, and

Theorem 4.9 ([A, Theorem 4.19]).

(1) *For $p \in [2, \infty)$ and $\xi \in \mathbb{D}_{1,2} \cap L_p$ one has*

$$\|\xi\|_{\Phi_2,p} \sim_{c_1} \|D\xi\|_{L_p^*(\Omega;L_2([0,T]))},$$

where $c_1 > 0$ depends on p only.

(2) *For $p \in (1, 2)$ and $\xi \in \mathbb{D}_{1,2}$ one has*

$$\frac{1}{c_2} \|D\xi\|_{L_\infty([0,T];L_p(\Omega))} \leq \|\xi\|_{\Phi_2,p} \leq c_2 \|D\xi\|_{L_p^*(\Omega;L_2([0,T]))},$$

where $c_2 > 0$ depends on p only.

(3) *There is a $\xi \in \mathbb{D}_{1,2}$ such that for all $p \in [1, \infty)$ one has $\xi \in L_p(\Omega)$, $D\xi \in L_p(\Omega; L_2([0,T]))$, and $\xi \notin \mathbb{B}_p^{\Phi_2}$.*

In deriving the upper bounds of Theorem 4.9(1) and (2), the following generalization of *Stein's lemma* is used:

Lemma 4.10 ([A, Lemma A.6]).
Let $p \in (1, \infty)$. Assume a process $a = (a_t)_{t\in[0,1]} \subseteq L_p$ with values in ℓ_2^N that has left-continuous paths for all $\omega \in \Omega$, a filtration $(\mathcal{H}_t)_{t\in[0,1]}$, and an $(\mathcal{H}_t)_{t\in[0,1]}$-adapted process $(b_t)_{t\in[0,1]} \subseteq L_p$ with values in ℓ_2^N that has left-continuous paths for all $\omega \in \Omega$ as well and such that $b_t = \mathbb{E}(a_t|\mathcal{H}_t)$ a.s. for all $t = k/2^n$ with $n = 0, 1, 2, \ldots$ and $k = 0, \ldots, 2^n - 1$. If $\sup_t |a_t| \in L_p$, then one has that

$$\left\| \left(\int_0^1 |b_t|^2 dt \right)^{\frac{1}{2}} \right\|_p \leq c_p \left\| \left(\int_0^1 |a_t|^2 dt \right)^{\frac{1}{2}} \right\|_p$$

where the constant $c_p > 0$ depends at most on p.

To further justify the term *an*isotropic, we note that the admissible functional in equation (18) can be generalized by letting $r \geq 2$, $0 \leq A < B \leq T$, and

$$\Phi_r^{A,B}(F) := \sup_{A\leq s<t\leq B} \frac{F(\chi_{(s,t]})}{(t-s)^{\frac{1}{r}}}. \tag{19}$$

Moreover, it is possible to study different fractional smoothness on different intervals. This can be done for example by putting

$$
\begin{aligned}
A_1 &:= \left\{\varphi \in \Delta \,\middle|\, \varphi = \chi_{(s,t]}, \quad 0 \le s < t \le \frac{T}{2}\right\}, \\
A_2 &:= \left\{\varphi \in \Delta \,\middle|\, \varphi = \chi_{(s,t]}, \quad \frac{T}{2} \le s < t \le T\right\},
\end{aligned}
$$

and

$$
\alpha(\varphi) = \begin{cases} (t-s)^{\frac{1}{2}}, & \varphi = \chi_{(s,t]} \in A_1, \\ (t-s)^{\frac{1}{4}}, & \varphi = \chi_{(s,t]} \in A_2. \end{cases}
$$

Then, putting $A = A_1 \cup A_2$, one can study the admissible functional

$$
\Phi(F) := \sup_{\varphi \in A} \frac{F(\varphi)}{\alpha(\varphi)}.
$$

Next we indicate how one can verify using decoupling that a random variable belongs to a certain anisotropic Besov space. The following is a special case of [A, Theorem 4.23].

Example 4.11. Consider $\xi := \chi_{[K,\infty)}(|X_T|)$, where $0 < K < \infty$ and $(X_r)_{r\in[0,T]}$ is the solution of

$$
X_r = x_0 + \int_0^r b(u, X_u)du + \int_0^r \sigma(u, X_u)dW_u, \quad r \in [0,T].
$$

Here $x_0 \in \mathbb{R}^d$, and the coefficients $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ satisfy

$(A_{b,\sigma})$ $b, \sigma \in C_b^{0,2}([0,T] \times \mathbb{R}^d)$, where the derivatives up to order two are taken with respect to the space-variables and, for some $\gamma \in (0,1]$, are assumed to be γ-Hölder continuous (w.r.t. the parabolic metric) on all compact subsets of $[0,T] \times \mathbb{R}^d$. Moreover, letting $A = \sigma\sigma^*$, there is a $\delta > 0$ such that $\langle Ax, x\rangle \ge \delta |x|^2$ for $x \in \mathbb{R}^d$ and b and σ are $\frac{1}{2}$-Hölder continuous in time, uniformly in space.

Then it follows from [13, Proposition B.3] that X_T has a bounded density f_{X_T}. Moreover, using Proposition 3.4(ii) and [1, Lemma 3.4] we have that for any $0 < p, q < \infty$

$$
\begin{aligned}
\|\xi - \xi^{(s,t]}\|_p^p &= \|\chi_{[K,\infty)}(|X_T|) - \chi_{[K,\infty)}(|X_T|)^{(s,t]}\|_1 \\
&= \|\chi_{[K,\infty)}(|X_T|) - \chi_{[K,\infty)}(|X_T^{(s,t]}|)\|_1 \\
&\le 3\|f_{X_T}\|_\infty^{\frac{q}{q+1}} \|X_T - X_T^{(s,t]}\|_q^{\frac{q}{q+1}}.
\end{aligned}
$$

Since [13, Theorem 3] implies that there exists $C > 0$ depending at most on (q, T, b, σ) such that

$$\left\| \sup_{r \in [0,T]} |X_r - X_r^{(s,t]}| \right\|_q \leq C\sqrt{t-s},$$

we have that there exists a constant $c = c(p, T, q, f_{X_T}, b, \sigma) > 0$ such that

$$\|\xi - \xi^{(s,t]}\|_p \leq c(\sqrt{t-s})^{\frac{q}{q+1}\frac{1}{p}}.$$

This means that given $p \in [2, \infty)$ we have $\xi \in \mathbb{B}_p^{\Phi_r}$ for any $r > 2p$.

5 Bounded Mean Oscillation

This Section reviews some of the fine properties that martingales of bounded mean oscillation, BMO-martingales from now on, satisfy, and we also recall the results from the theory of weighted BMO that we used. The results that are new (Section 5.2) are easy observations, but they were useful to us. We present them here as they might be of independent interest. For simplicity, we present all definitions and results in the setting that was fixed in the beginning of Section 2.

5.1 BMO and weighted BMO

First we recall what it means that a martingale is of bounded mean oscillation.

Definition 5.1. *A martingale $M = (M_t)_{t \in [0,T]}$ is a* BMO*-martingale provided that $M_0 \equiv 0$ and there is constant $c > 0$ such that for all stopping times $\tau : \Omega \to [0, T]$ one has that*

$$\mathbb{E}(|M_T - M_{\tau-}|^2 | \mathcal{F}_\tau) \leq c^2, \tag{20}$$

where

$$M_{\tau-} := \lim_{n \to \infty} M_{\left((\tau - \frac{1}{n}) \vee 0\right)}.$$

We let $\|M\|_{\mathrm{BMO}} := \inf c$ where the infimum is taken over all $c > 0$ as above.

In our setting every martingale is continuous, and the probability space is complete. It follows from optional stopping theorem that:

Lemma 5.2. *Let $M = (M_t)_{t\in[0,T]}$ a martingale with $M_0 \equiv 0$. Then M is a BMO-martingale if and only if there is a constant $c > 0$ such that*

$$\mathbb{E}(|M_T - M_t|^2|\mathcal{F}_t) \leq c^2$$

for all $t \in [0, T]$. Moreover, $\|M\|^2_{\mathrm{BMO}} = \sup_{t\in[0,T]} \|\mathbb{E}^{\mathcal{F}_t}|M_T - M_t|^2\|_\infty$.

At this point an educated reader might be worried of the notation "BMO"; why is it not emphasized that we study $\mathbb{E}^{\mathcal{F}_t}|M_T - M_t|^2$ instead of $\mathbb{E}^{\mathcal{F}_t}|M_T - M_t|^p$ for some $p \in [1, \infty)$? The answer is remarkable:

$$\sup_{t\in[0,T]} \|\mathbb{E}^{\mathcal{F}_t}|M_T - M_t|^q\|_\infty < \infty \quad \text{for some } q \in (0, \infty)$$

if and only if

$$\sup_{t\in[0,T]} \|\mathbb{E}^{\mathcal{F}_t}|M_T - M_t|^p\|_\infty < \infty \quad \text{for all } p \in (0, \infty).$$

This follows from the celebrated John-Nirenberg inequality [21, Theorem 2.1], and is contained in [21, Corollary 2.1].
For $p \in [1, \infty]$ we denote by H_p the space of martingales that satisfy

$$\|M\|_{H_p} := \|\langle M\rangle_T^{\frac{1}{2}}\|_{L_p} < \infty,$$

where $\langle M\rangle$ is the quadratic variation process of M. If M is of the form $(M_t)_{t\in[0,T]} = (\int_0^t \alpha_s dB_s)_{t\in[0,T]}$, where B is a Brownian motion, then

$$\|M\|_{H_p} = \left\| \left(\int_0^T |\alpha_s|^2 ds \right)^{\frac{1}{2}} \right\|_{L_p}.$$

One important property of BMO is, that it can be characterized as the dual of H_1. One part of this result, the fact that BMO $\subseteq H_1^*$, is proven in [21, Theorem 2.5]. A special case of this result can be written as follows:

Theorem 5.3. *(Fefferman's inequality)*
Assume that $B = (B_r)_{r\in[0,T]}$ is a one-dimensional Brownian motion, that $M = (M_t)_{t\in[0,T]} = (\int_0^t \alpha_s dB_s)_{t\in[0,T]} \in BMO$, and that $N = (N_t)_{t\in[0,T]} = (\int_0^t \beta_s dB_s)_{t\in[0,T]} \in H_1$. Then

$$\mathbb{E}\left[\int_0^T |\alpha_s\beta_s| ds\right] \leq \sqrt{2}\|M\|_{BMO}\|N\|_{H_1}. \tag{21}$$

In [10, Lemma 1.6] the inequality (21) is generalized to include the following case:

Proposition 5.4. *Let $p \in [1, \infty)$, assume that $B = (B_r)_{r\in[0,T]}$ is a one-dimensional Brownian motion, $M = (M_t)_{t\in[0,T]} = (\int_0^t \alpha_s dB_s)_{t\in[0,T]} \in BMO$ and that $N = (N_t)_{t\in[0,T]} = (\int_0^t \beta_s dB_s)_{t\in[0,T]} \in H_p$. Then*

$$\left(\mathbb{E}\left(\int_0^T |\alpha_s\beta_s| ds\right)^p\right)^{1/p} \leq \sqrt{2}p\|M\|_{BMO}\|N\|_{H_p}. \tag{22}$$

In [15] the traditional BMO is generalized by replacing the constant c in equation (20) by an adapted weight process. We will use $\mathcal{CL}_0$ for the set of càdlàg adapted processes $A = (A_t)_{t\in[0,T]}$ with $A_0 = 0$, and $\mathcal{CL}^+$ for the set of càdlàg adapted processes A with $A_t(\omega) > 0$ for all $(t, \omega) \in \Omega_T$.

Definition 5.5 ([15, Definition 1]).
Let $A \in \mathcal{CL}_0$, and $\Phi \in \mathcal{CL}^+$. We define for $p \in (0, \infty)$

$$\|A\|_{\mathrm{BMO}_p^\Phi} := \sup_\sigma \left\| \mathbb{E}\left[\frac{|A_T - A_{\sigma-}|^p}{\Phi_\sigma^p} \;\middle|\; \mathcal{F}_\sigma\right]\right\|_{L_\infty}^{1/p},$$

where the supremum is taken over all stopping times $\sigma : \Omega \to [0, T]$.

The main use of this theory for us is, that it gives nice tail estimates. The traditional BMO can even be characterized by a tail estimate; the following result is mentioned in [21, page 26] and goes back to Emery:

Proposition 5.6. *A uniformly integrable martingale $M = (M_t)_{t\in[0,\infty)}$ is of bounded mean oscillation if and only if there exist two constants $a \geq 0$ and $0 < \epsilon < 1$ such that*

$$\mathbb{P}\left(\sup_{0\leq t<\infty} |M_{\sigma+t} - M_\sigma| > a \;\middle|\; \mathcal{F}_\sigma\right) \leq 1 - \epsilon$$

for any stopping time σ.

With weighted BMO we do not necessarily have a characterization, but we do have an additive upper bound consisting of an exponential part and a part that is a tail estimate of the weight process. We will make use of the notation

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

for $B \in \mathcal{F}_T$ of positive measure. The first step towards a tail estimate is the following:

Proposition 5.7 ([15, Example 1 and proof of Corollary 1(a)]).
Let $p \in (0,\infty)$ and assume that $Y \in \mathcal{CL}_0$ and $\Phi \in \mathcal{CL}^+$ are such that

$$\|Y\|_{\mathrm{BMO}_p^\Phi} \le C,$$

where $C > 0$. Then, letting $\theta \in (0, \frac{1}{2})$, and defining $A \in \mathcal{CL}_0$ by $A := \frac{Y\theta^{1/p}}{C}$, we have for any $\nu > 0$, stopping time σ, and $B \in \mathcal{F}_\sigma$ that

$$\mathbb{P}_B\left(|A_T - A_{\sigma-}| > \nu\right) \le \theta + \mathbb{P}_B(\Phi_\sigma > \nu).$$

Now we consider $\Phi \in \mathcal{CL}^+$, and let

$$W_\Phi(B,\nu;\sigma) := \mathbb{P}\left(B \cap \left\{\sup_{u\in[\sigma,T]} \Phi_u > \nu\right\}\right),$$

for $\nu > 0$, a stopping time σ, and $B \in \mathcal{F}_\sigma$. After this we can proceed with:

Theorem 5.8 ([15, Theorem 1]).
Assume that $A \in \mathcal{CL}_0$, $\Phi \in \mathcal{CL}^+$, and that there is $\theta \in (0, \frac{1}{2})$ such that

$$\mathbb{P}_B(|A_T - A_\sigma| > \nu) \le \theta + \frac{W_\Psi(B,\nu;\sigma)}{\mathbb{P}(B)} \tag{23}$$

for all $\nu > 0$, stopping times σ, and $B \in \mathcal{F}_\sigma$ of positive measure. Then there are constants $a, \alpha > 0$, depending on θ only, such that

$$\begin{aligned}\mathbb{P}_B\left(\sup_{u\in[\sigma,T]} |A_u - A_\sigma| > \lambda + a\mu\nu\right) \quad &\le \quad e^{1-\mu}\mathbb{P}_B\left(\sup_{u\in[\sigma,T]} |A_u - A_\sigma| > \lambda\right) \\ &\quad + \alpha\frac{W_\Psi(B,\nu;\sigma)}{\mathbb{P}(B)}\end{aligned}$$

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

BMO-martingales have many nice properties as explained in the excellent lecture notes [21], we will now concentrate on one that is relevant with BSDEs. In our apriori estimate concerning BSDEs (Lemma 6.1), we remove the superlinear (up to quadratic) drift term using a Girsanov transformation. This technique was already used in [19] to prove uniqueness of the solution to certain quadratic BSDE. In [6] it is nicely presented how to proceed to obtain an apriori estimate, which then results into an existence result. The fact that the Girsanov transformation is well-defined follows from the Reverse Hölder inequality:

Definition 5.9. *Assume a martingale* $M=(M_t)_{t\in[0,T]}$ *with* $M_0 \equiv 0$ *such that* $\mathcal{E}(M)$ *with*

$$\mathcal{E}(M)_t = e^{M_t - \frac{1}{2}\langle M\rangle_t}$$

for $t \in [0,T]$ *is a martingale as well. For* $\beta \in (1,\infty)$ *we let* $\mathcal{E}(M) \in \mathcal{RH}_\beta$ *provided that there is a constant* $c > 0$ *such that for all stopping times* $\tau : \Omega \to [0,T]$ *one has that*

$$\mathbb{E}(|\mathcal{E}(M)_T|^\beta | \mathcal{F}_\tau)^{\frac{1}{\beta}} \le c\mathcal{E}(M)_\tau \ a.s.$$

The smallest possible $c \ge 0$ *is denoted by* $\mathcal{RH}_\beta(\mathcal{E}(M))$.

The reason why this leads to BMO-martingales is the following result:

Proposition 5.10 ([21, Theorems 2.4 and 3.4])**.**
Let M *be a martingale with* $M_0 \equiv 0$ *such that* $\mathcal{E}(M)$ *is a martingale. Then* $M \in \mathrm{BMO}$ *if and only if* $\mathcal{E}(M) \in \bigcup_{p\in(1,\infty)} \mathcal{RH}_p$.

In our application of the reverse Hölder inequalities this is still a bit unsatisfactory, as we would like to know for which $p \in (1,\infty)$ we have $\mathcal{E}(M) \in \mathcal{RH}_p$. For this purpose we let

$$\Phi : (1,\infty) \to (0,\infty), \quad \Phi(p) = \left(1 + \frac{1}{p^2}\log\left(1 + \frac{1}{2p-2}\right)\right)^{1/2} - 1, \tag{24}$$

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

$$\begin{aligned} \Psi \quad &: \quad \Big\{(\gamma,p) \in [0,\infty)\times(1,\infty) : 0 \le \gamma < \Phi(p) < \infty\Big\} \to [0,\infty), \\ \Psi(\gamma,p) \quad &:= \quad \left(\frac{2}{1 - \frac{2p-2}{2p-1}e^{p^2[\gamma^2+2\gamma]}}\right)^{\frac{1}{p}}. \end{aligned}$$

Then, according to [21, Proof of Theorem 3.1], we have that

$$\Psi(\gamma_1,p) \le \Psi(\gamma_2,p) \quad \text{for} \quad 0 \le \gamma_1 \le \gamma_2 < \Phi(p),$$

and

$$\|M\|_{\mathrm{BMO}} < \Phi(p) \quad \text{implies} \quad \mathcal{RH}_p(\mathcal{E}(M)) \le \Psi(\|M\|_{\mathrm{BMO}},p). \tag{25}$$

5.2 New results

We will improve (25) by considering the approach of *sliceable numbers*. To formulate it, we recall the notation ${}^{\sigma}M^{\tau} := (M_{\tau\wedge t} - M_{\sigma\wedge t})_{t\in[0,T]}$ for stopping times σ, τ.

Definition 5.11 ([A, Definition 5.2]).
For a BMO-martingale $M = (M_t)_{t\in[0,T]}$ and $N \geq 1$ we let

$$\mathrm{sl_N(M)} := \inf \varepsilon,$$

where the infimum is taken over all $\varepsilon > 0$ such that there are stopping times $0 = \tau_0 \leq \tau_1 \leq \cdots \leq \tau_N = T$ with

$$\sup_{k=1,\dots,N} \|{}^{\tau_{k-1}}M^{\tau_k}\|_{\mathrm{BMO}} \leq \varepsilon.$$

Moreover, we let

$$\mathrm{sl_\infty(M)} := \lim_{\mathrm{N}} \mathrm{sl_N(M)}.$$

We call $\mathrm{sl_N(M)}$ *the N-*sliceable number *of M. The BMO-martingale M is called* sliceable *provided that* $\mathrm{sl_\infty(M)} = 0$.

Some properties of these numbers are:

Lemma 5.12 ([A, Lemma 5.4]).
For $M, M_1, M_2 \in \mathrm{BMO}$ one has the following:

(i) $\mathrm{sl_1(M)} = \|\mathrm{M}\|_{\mathrm{BMO}}$.

(ii) $\mathrm{sl_1(M)} \geq \mathrm{sl_2(M)} \geq \cdots \geq 0$.

(iii) $\mathrm{sl_{N_1+N_2-1}(M_1 + M_2)} \leq \mathrm{sl_{N_1}(M_1)} + \mathrm{sl_{N_2}(M_2)}$.

(iv) *For $\epsilon > 0$ we have $\mathrm{sl_\infty(M)} < \epsilon$ if and only if $d_{\mathrm{BMO}}(M, H_\infty) < \epsilon$, where $d_{\mathrm{BMO}}(M, H_\infty)$ is the distance of M to H_∞ with respect to the* BMO-*norm.*

Part (iv) of Lemma 5.12 follows from [29, Theorem 1.1]. Parts (i),(ii), and (iii) on the other hand imply that sliceable numbers are a relative of the s-numbers from operator-theory, see for example [28, Chapter 2.2].

Now we can improve the index p in (25) by the following

Theorem 5.13 ([A, Theorem 5.8]).
If $\mathrm{sl_N(M)} < \Phi(\mathrm{p})$, *then* $\mathcal{RH}_p(\mathcal{E}(M)) \leq \left[\Psi(\mathrm{sl_N(M)}, \mathrm{p})\right]^{\mathrm{N}}$. *In particular, if M is sliceable then* $\mathcal{E}(M) \in \bigcap_{p\in(1,\infty)} \mathcal{RH}_p$.

In our approach to BSDEs it is useful to consider BMO and sliceable numbers for processes as well.

Definition 5.14 ([A, Definition 5.9])**.**
For $m \in \mathbb{N}$ and an $\mathbb{R}^m$-valued predictable process $Z = (Z_t)_{t\in[0,T]}$ we let

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

This is quantified using for any $N \geq 1$

$$\mathrm{sl_N(Z)} := \inf \varepsilon,$$

where the infimum is taken over all $\varepsilon > 0$ such that there are stopping times $0 = \tau_0 \leq \tau_1 \leq \cdots \leq \tau_N = T$ with

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

Moreover, we let $\mathrm{sl}_\infty(\mathrm{Z}) := \lim_\mathrm{N} \mathrm{sl_N(Z)}$.

We illustrate two possible ways of exploiting the martingale setting:
(1) For a d-dimensional process Z it follows from Itô's isometry that $\|Z\|_{\mathrm{BMO}(S_2)} = \| \int_0^\cdot Z_r dW_r\|_{\mathrm{BMO}}$, so that for example Theorem 5.13 is applicable for this type of processes.
(2) Assume that the process Z is m-dimensional for some $m \in \mathbb{N}$, and such that $\||Z|^\theta\|_{\mathrm{BMO}(S_2)} < \infty$ for some $\theta \in [0, \infty)$. Then we let $(B_t)_{t\in[0,T]}$ be (for example) the first component of the d-dimensional Brownian motion $(W_t)_{t\in[0,T]}$, and consider the martingale

$$M = (M_t)_{t\in[0,T]} := \left(\int_0^t |Z_r|^\theta dB_r \right)_{t\in[0,T]}.$$

Using Itô's isometry we once again have that $\|M\|_{\mathrm{BMO}} < \infty$, so that for example Proposition 5.4 and Lemma 5.12 are applicable for this type of processes.
Note that by considering $\||Z|^\theta\|_{\mathrm{BMO}(S_2)} < \infty$ for $\theta \in (1, \infty)$ we obtain a condition that is stronger than the classical BMO condition $\|Z\|_{\mathrm{BMO}(S_2)} < \infty$, whereas for $\theta \in (0, 1)$ the condition gets weaker.
If the generator of our BSDE is locally Lipschitz with parameter $\theta \in [0, 1]$, then we need to assume that the control process Z satisfies $\||Z|^\theta\|_{\mathrm{BMO}(S_2)} < \infty$. At the same time we need to have $\left(\int_0^T |Z_r|^2 dr\right)^{\frac{1}{2}} \in L_q$,

where $q \geq 2$ is determined by the sliceability-number $\mathrm{sl}_\infty(|\mathrm{Z}|^\theta)$ (see assumption (A_1) in subsection 6.2 below).
To see that all these conditions can hold at the same time, we have an example of a process Z that satisfies the additional condition that $\left(\int_0^T |Z_t|^2 dt\right)^{1/2} \in \mathcal{L}_{\exp}$, where the Orlicz space $L_{\exp}$ is given by

$$\|F\|_{L_{\exp}} := \inf\left\{\lambda > 0 : \mathbb{E}e^{\frac{|F|}{\lambda}} \leq 2\right\}$$

for a random variable F taking values in $\mathbb{R}$.

Example 5.15 ([A, Example 5.10]).
For each $\eta \in (0,1)$ there is a predictable process $Z = (Z_t)_{t\in[0,T]}$ such that

1. $\left(\int_0^T |Z_t|^2 dt\right)^{1/2} \in \mathcal{L}_{\exp}$,
2. $|Z|^\eta \in \mathrm{BMO}(S_2)$,
3. $Z \notin \mathrm{BMO}(S_2)$.

The case that $|Z|^\theta$ is sliceable is important to us, as then we do not need to impose any additional integrability on $\left(\int_0^T |Z_r|^2 dr\right)^{\frac{1}{2}}$. Thus, we present two cases for $\mathrm{sl}_\infty(|\mathrm{Z}|^\theta) = 0$:

Example 5.16 ([A, Remark 6.5]).

(i) *For $\theta = 0$ (with the convention $0^0 := 1$) we have that*

$$\mathrm{sl}_{\mathrm{N}}\left(|\mathrm{Z}|^\theta\right) \leq \sqrt{\frac{\mathrm{T}}{\mathrm{N}}}$$

if we take equidistant time-nets.

(ii) *Let $0 < \theta < \eta \leq 1$ and assume that $\||Z|^\eta\|_{\mathrm{BMO}(S_2)} < \infty$. Then we obtain for any $0 \leq a < b \leq T$ that*

$$\|(\chi_{(a,b]}(t)|Z_t|^\theta)_{t\in[0,T]}\|_{\mathrm{BMO}(S_2)}^{\frac{1}{\theta}} \leq (b-a)^{\frac{1}{2\theta}-\frac{1}{2\eta}}\|(\chi_{(a,b]}(t)|Z_t|^\eta)_{t\in[0,T]}\|_{\mathrm{BMO}(S_2)}^{\frac{1}{\eta}}$$

and, by using equidistant grids, that

$$\mathrm{sl}_{\mathrm{N}}\left(|\mathrm{Z}|^\theta\right) \leq \left(\frac{\mathrm{T}}{\mathrm{N}}\right)^{\frac{1}{2}\left(1-\frac{\theta}{\eta}\right)} \||\mathrm{Z}|^\eta\|_{\mathrm{BMO(S_2)}}^{\frac{\theta}{\eta}}.$$

Example 5.16(ii) can be seen as an embedding theorem: Let $(B_r)_{r\in[0,T]}$ be a 1-dimensional Brownian motion, let $(\alpha_r)_{r\in[0,T]}$ be a 1-dimensional predictable process such that $\mathbb{E}\int_0^T |\alpha_r|^2 dr < \infty$, and define

$$T_\theta\left(\int_0^\cdot \alpha_r dB_r\right) := \int_0^\cdot |\alpha_r|^\theta sgn(\alpha_r) dB_r,$$

for $\theta \in [0,1]$. Then it follows from Example 5.16(ii) and Lemma 5.12(iv) that $T_\theta(\int_0^\cdot \alpha_r dW_r) \in \overline{H_\infty}^{\mathrm{BMO}}$, whenever $\| |\alpha|^\eta \|_{\mathrm{BMO}(S_2)} < \infty$ and $0 \le \theta < \eta \le 1$.

6 Main results concerning BSDEs

We begin this Section with the apriori estimate Lemma 6.1. The proof of this result uses the reverse Hölder inequalities, and it is a continuation of the arguments from [6, Proposition 2.3]. Their apriori estimate upper bounded an L_p-quantity of the solution by an L_{2p}-quantity of the data. Using Proposition 5.4 we were able to improve this so that an L_p-quantity of the solution is upper bounded by an L_p-quantity of the data.

Using the apriori estimate together with decoupling, especially with Proposition 3.5, we deduce the stability result Theorem 6.3. Applying this result we obtain results on L_p-variation of the solution of BSDE (26), as well as an embedding theorem with respect to anisotropic Besov spaces. Moreover, the conditional version of the stability result (Proposition 6.9) implies that the solution is in a weighted BMO-space where the weight depends only on the data. As mentioned in Section 5, this yields to tail estimates (like in the John-Nirenberg theorem) for the variation of the Y-process of our BSDE.

6.1 Apriori estimate

In this section we follow the ideas of [6, Proof of Proposition 2.3] but adapt and extend the ideas for our purpose. Let $B = (B_t)_{t\in[0,T]}$ be an n-dimensional standard Brownian motion (where all paths are continuous) on a basis $(A, \mathcal{A}, \mu, (\mathcal{A}_t)_{t\in[0,T]})$, where $(A, \mathcal{A}, \mu)$ is complete, $(\mathcal{A}_t)_{t\in[0,T]}$ the augmentation of the natural filtration of B, and $\mathcal{A}_T = \mathcal{A}$. We consider the two backward equations

$$\begin{aligned} Y_t^0 &= \xi^0 + \int_t^T f^0(s, Y_s^0, Z_s^0) ds - \int_t^T Z_s^0 dB_s, \quad t \in [0,T], \\ Y_t^1 &= \xi^1 + \int_t^T f^1(s) ds - \int_t^T Z_s^1 dB_s, \quad t \in [0,T], \end{aligned}$$

where we assume the following conditions:

(D1) The processes f^1, Z^0 and Z^1 are predictable and the processes Y^0 and Y^1 continuous and adapted,

(D2) $\mathbb{E}|\xi^i|^2 < \infty$ and $\mathbb{E}\int_0^T |Z_s^i|^2 ds < \infty$ for $i = 0, 1$,

(D3) $\mathbb{E}\left|\int_0^T |f^0(s, Y_s^0, Z_s^0)| ds\right|^2 < \infty$ and $\mathbb{E}\left|\int_0^T |f^1(s)| ds\right|^2 < \infty$,

(D4) the generator $f^0 : \Omega_T \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ is such that $(t, \omega) \mapsto f^0(t, \omega, y, z)$ is predictable for all (y, z), $(y, z) \to f^0(t, \omega, y, z)$ is continuous for all (t, ω), and there is an $L_Y \geq 0$ such that, for all (t, ω, y_0, y_1, z),

$$|f^0(t, \omega, y_0, z) - f^0(t, \omega, y_1, z)| \leq L_Y |y_0 - y_1|.$$

We let $\Delta\xi := \xi^1 - \xi^0$, and for $s \in [0, T]$,

$$\begin{aligned}
\Delta Y_s &:= Y_s^1 - Y_s^0, \\
\Delta Z_s &:= Z_s^1 - Z_s^0, \\
a_s &:= f^1(s) - f^0(s, Y_s^1, Z_s^1), \\
c_s &:= \frac{f^0(s, Y_s^0, Z_s^1) - f^0(s, Y_s^0, Z_s^0)}{|\Delta Z_s|^2} \chi_{\{\Delta Z_s \neq 0\}} \Delta Z_s, \\
\Xi_s &:= |\Delta\xi| + \int_s^T |a_r| dr.
\end{aligned}$$

Lemma 6.1 ([A, Lemma 6.17]).
Assume that $c = (c_t)_{t \in [0,T]}$ is n-dimensional with $\|c\|_{\mathrm{BMO}(S_2)} \leq \beta < \infty$, and that $\lambda_t := \exp(\int_0^t c_s dB_s - \frac{1}{2}\int_0^t |c_s|^2 ds)$ and $p_0 \in (1, \infty)$ are such that $\mathcal{RH}_{p_0'}(\lambda) \leq \rho < \infty$ with $1 = \frac{1}{p_0} + \frac{1}{p_0'}$. Assume $p \in [2, \infty)$ with $p > p_0$ such that

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

Then there is a $c_{(6.1)} \in (0, \infty)$, depending at most on $(T, L_Y, p, p_0, \beta, \rho, n)$, such that for all $t \in [0, T]$ one has that

$$\left\| \sup_{s \in [t,T]} |\Delta Y_s| \right\|_p + \left\| \left(\int_t^T |\Delta Z_s|^2 ds\right)^{\frac{1}{2}} \right\|_p \leq c_{(6.1)} \|\Xi_t\|_p.$$

6.2 Basic result

We start by listing the assumptions that will be used throughout this Section. These conditions are imposed on random objects defined on Ω, even though we use our decoupling technique that is defined for objects on $\overline{\Omega}$. However, in [A] and [B] we have shown that these conditions hold for appropriate elements of $\overline{\Omega}$ as well. The validity of these assumptions, as well as examples when they hold, are considered in [A, Section 18].

(A_1) There exists a solution (Y, Z) to the equation

$$Y_t = \xi + \int_t^T f(r, Y_r, Z_r)dr - \int_t^T Z_r dW_r, \quad t \in [0,T], \tag{26}$$

where we assume that there exists $p \in [2,\infty)$ and $\theta \in [0,1]$ such that the following conditions are satisfied:

(i) The process Y is $(\mathcal{F}_r)_{r\in[0,T]}$-adapted and path-wise continuous, and the process Z is $(\mathcal{F}_r)_{r\in[0,T]}$-predictable

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

$$\begin{aligned}|f(t,\omega,y_0,z_0) - f(t,\omega,y_1,z_1)| \\ \leq L_y|y_0 - y_1| + L_z[1 + |z_0| + |z_1|]^\theta |z_0 - z_1|\end{aligned}$$

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

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

(v) We assume

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

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

$$\mathrm{sl_N}(|\mathrm{Z}|^\theta) \leq \mathrm{s_N},$$

and put $s_\infty := \lim_N s_N$

(vi) If $s_\infty > 0$, then the constant $p \in [2,\infty)$ satisfies additionally

$$p > \frac{\Phi^{-1}(2\sqrt{2}L_z s_\infty)}{\Phi^{-1}(2\sqrt{2}L_z s_\infty) - 1} \in (1,\infty),$$

where $\Phi : (1,\infty) \to (0,\infty)$ is defined in equation (24).

Remark 6.2. Assumption (A_1)(v) is an implicit fractional BMO assumption on the control process Z. This assumption is used to remove the (up to quadratic) drift term of the generator in the z-variable. Here we need to use the reverse Hölder inequalities, which by Theorem 5.13 hold for all $q \in (1, \infty)$ if $s_\infty = 0$. In case $s_\infty > 0$, we only have the reverse Hölder inequalities for small enough q, and this is why we need to assume some more integrability as imposed by (A_1)(vi). Two examples of the case $s_\infty = 0$ were discussed in Example 5.16, now we explain its impact regarding Lipschitz and quadratic BSDEs.

(i) Assume that f satisfies (A_1)(ii) and (A_1)(iii) with $\theta = 0$ and $1 < p < \infty$, and that $\xi \in L_p$. Then there exists a unique solution (Y, Z) of BSDE (26), and (A_1) is satisfied with $\theta = 0$. This follows for example from [5, Theorem 4.2]. Note that since $\theta = 0$, we have $s_\infty = 0$.

(ii) Assume that f satisfies (A_1)(ii) with $\theta = 1$, that $\int_0^T |f(s,0,0)| ds \in \mathcal{L}_\infty$, and that $\xi \in L_\infty$. Then there exists a unique solution (Y, Z) of BSDE (26), and (A_1) is satisfied with $\theta = 1$ and all $p \in [2, \infty)$. The solution is unique in the class $S^\infty \times L_2(\Omega_T)$, where S^∞ consists of bounded continuous processes. This follows for example from [24, Theorem 2.6 and Lemma 3.1]. However, it might be that $s_\infty > 0$, so that our results only hold when $p \geq 2$ is *large enough* as imposed by (A_1)(vi).

(iii) Assume that f satisfies (A_1)(ii) with $0 < \theta < 1$, that $\int_0^T |f(s,0,0)| ds \in \mathcal{L}_\infty$, and that $\xi \in L_\infty$. Then, as above, we have a unique solution and (A_1) is satisfied with $\theta = 1$ and all $p \in [2, \infty)$. Now it follows as in Example 5.16(ii) that $s_\infty = 0$, so that our results hold for all $p \geq 2$.

Now we can state our basic result:

Theorem 6.3 ([A, Theorem 6.4]).
Assume (A_1). Then we have for all $t \in [0, T]$ that

$$\begin{aligned} &\left\| \sup_{s \in [t,T]} |Y_s^\varphi - Y_s^\psi| \right\|_p \\ &\quad + \left\| \left(\int_t^T D\big[\varphi(s), \psi(s)\big] \, |Z_s|^2 ds \right)^{\frac{1}{2}} \right\|_p + \left\| \left(\int_t^T |Z_s^\varphi - Z_s^\psi|^2 ds \right)^{\frac{1}{2}} \right\|_p \\ &\leq c_{(6.3)} \left[\|\xi^\varphi - \xi^\psi\|_p + \left\| \int_t^T |f^\varphi(s, Y_s^\psi, Z_s^\psi) - f^\psi(s, Y_s^\psi, Z_s^\psi)| ds \right\|_p \right], \quad (27) \end{aligned}$$

where $\varphi, \psi \in \Delta$, $D[\nu,\eta] := 1 - \sqrt{1-\nu^2}\sqrt{1-\eta^2} - \nu\eta$, and $c_{(6.3)} > 0$ depends at most on $(L_y, L_z, T, (s_N)_{N=1}^{\infty}, p, d)$.

Remark 6.4. *The function $D[\nu,\eta] : [0,1]^2 \to [0,1]$ measures the distance between ν and η, by projecting the vector $(\nu, \sqrt{1-\nu^2})$ onto $(\eta, \sqrt{1-\eta^2})$ and comparing the projection to $(\eta, \sqrt{1-\eta^2})$. In particular, $D[\nu,\eta] = 0$ if and only if $\nu = \eta$.*

6.3 L_p-variation

As it was explained in Section 3.1, it is important to estimate the L_p-variation of the solution of a BSDE. Our main result in this respect is

Theorem 6.5 ([A, Theorem 6.16])**.**
Assume (A_1). Then there is a constant $c_{(6.5)} > 0$ such that for all $0 \le s < t \le T$ one has

$$\|Y_t - Y_s\|_p \le c_{(6.5)} \left[\left\| \int_s^t [1 + |f(r,0,0)|] dr \right\|_p + \|\xi - \xi^{(s,t]}\|_p \right.$$
$$\left. + \left\| \int_s^T |f(r,Y_r,Z_r) - f^{(s,t]}(r,Y_r,Z_r)| dr \right\|_p \right].$$

Proof. Fix $0 \le s < t \le T$. Since Y_t is $\mathcal{F}_t$-measurable, we obtain

$$\begin{aligned}
\|Y_t - Y_s\|_p &\le \|Y_t - \mathbb{E}^{\mathcal{F}_s} Y_t\|_p + \|\mathbb{E}^{\mathcal{F}_s} Y_t - Y_s\|_p \\
&\le \|Y_t - Y_t^{(s,t]}\|_p + \|\mathbb{E}^{\mathcal{F}_s} Y_t - Y_s\|_p \\
&\le c_{(6.3)} \left[\|\xi - \xi^{(s,t]}\|_p + \left\| \int_t^T |f(r,Y_r,Z_r) - f^{(s,t]}(r,Y_r,Z_r)| dr \right\|_p \right] \\
&\quad + \|\mathbb{E}^{\mathcal{F}_s} Y_t - Y_s\|_p,
\end{aligned}$$

where we applied Theorem 6.3 with the pair $(0, \chi_{(s,t]}) \in \Delta \times \Delta$. The last term can be upper bounded by

$$\begin{aligned}
&\|\mathbb{E}^{\mathcal{F}_s} Y_t - Y_s\|_p \\
= &\left\| \mathbb{E}^{\mathcal{F}_s} \int_s^t f(r,Y_r,Z_r) dr \right\|_p \\
\le &\left\| \int_s^t |f(r,0,0)| dr \right\|_p + L_y \left\| \int_s^t |Y_r| dr \right\|_p + L_z \left\| \int_s^t [1 + |Z_r|]^\theta |Z_r| dr \right\|_p .
\end{aligned}$$

It follows from our assumptions that $\sup_{r\in[0,T]}|Y_r| \in \mathcal{L}_p$ so that $\sup_{r\in[0,T]}\|Y_r\|_p < \infty$. On the other hand, we can use Proposition 5.4 (with B one of the components of W) and Theorem 6.3 to deduce

$$\begin{aligned}
&\left\|\int_s^t [1+|Z_r|]^\theta |Z_r| dr\right\|_p \\
\leq\ & \sqrt{2}p \left\|\left(\int_s^t |Z_r|^2 dr\right)^{\frac{1}{2}}\right\|_p \sup_{r\in[0,T]} \left\|\mathbb{E}\left(\int_r^T \chi_{(s,t]}(u)[1+|Z_u|]^{2\theta} du \,\middle|\, \mathcal{F}_r\right)\right\|_{L_\infty}^{\frac{1}{2}} \\
\leq\ & \sqrt{2}pc_{(6.3)} \left[\|\xi-\xi^{(s,t]}\|_p + \left\|\int_s^T |f(r,Y_r,Z_r) - f^{(s,t]}(r,Y_r,Z_r)| dr\right\|_p\right] \times \\
& \times \left[\sqrt{t-s} + \sup_{r\in[0,T]} \left\|\mathbb{E}\left(\int_r^T |Z_u|^{2\theta} du \,\middle|\, \mathcal{F}_r\right)\right\|_{L_\infty}^{\frac{1}{2}}\right].
\end{aligned}$$

This concludes the proof. □

6.4 Anisotropic Besov Spaces

In this Section we present our non-linear embedding theorem for BSDEs and anisotropic Besov spaces. The main result is Theorem 6.8, which states that if the data (ξ, f) is in certain anisotropic Besov spaces, then so is the solution (Y, Z).

Definition 6.6. *For $q, r \in [1,\infty)$, a predictable process $(A_t)_{t\in[0,T]}$ with*

$$\left\|\left(\int_0^T |A_s|^r ds\right)^{\frac{1}{r}}\right\|_q < \infty,$$

for $t \in [0,T]$, and for an admissible functional[11] *$\Phi : C^+(\Delta) \to [0,\infty]$ we let*

$$\|A\|_{\Phi,q}^{r,t} := \Phi\left(\varphi \to \left\|\left(\int_t^T |A_s - A_s^\varphi|^r ds\right)^{\frac{1}{r}}\right\|_q\right).$$

First we note that this definition is possible:

Lemma 6.7 ([A, Lemma 6.13])**.** *The map*

$$\varphi \to \left\|\left(\int_t^T |A_s - A_s^\varphi|^r ds\right)^{\frac{1}{r}}\right\|_q$$

is continuous as a map from Δ into $[0,\infty)$.

[11] In the sense of Definition 4.3.

Theorem 6.8 ([A, Corollary 6.14]).
Let $t \in [0,T]$, assume (A_1), and that there are predictable processes $(V_s^l)_{s\in[t,T]}$ such that, for all $\varphi \in \Delta$,

$$\left\| \int_t^T |f(s,Y_s^\varphi,Z_s^\varphi) - f^\varphi(s,Y_s^\varphi,Z_s^\varphi)|ds \right\|_p \le \sum_{l=1}^L \|V_\cdot^l - (V_\cdot^l)^\varphi\|_{L_{q_l}(L_{r_l}([t,T]))}$$

for some $q_l \in [p,\infty)$ and $r_l \in [1,\infty)$[12]*. Let Φ be admissible. Then we have that*

$$\|Y_t\|_{\Phi,p} + \|Z\|_{\Phi,p}^{2,t} \le 2c_{(6.3)} \left[\|\xi\|_{\Phi,p} + \sum_{l=1}^L \|V^l\|_{\Phi,q_l}^{r_l,t} \right].$$

Proof. Follows directly from Theorem 6.3 applied to the pair $(0,\varphi)$. □

Examples how to obtain processes V in Theorem 6.8 are discussed in [A, Section 19].

6.5 Weighted BMO

In this Section we show how we can exploit the idea of self-iterating inequalities that are used to prove exponential tail estimates under assumptions on the mean oscillation of functions or processes in the context of BSDEs. Roughly speaking, we consider our results conditionally and obtain via iteration better tail estimates than L_p-estimates would give.

We use the notation $(\overline{\mathcal{F}}_r)_{r\in[0,T]}$ for the natural filtration of the $2d$-dimensional Brownian motion (W,W'), augmented by all $\overline{\mathbb{P}}$-nullsets. The conditional version of Theorem 6.3 is as follows:

Proposition 6.9 ([B, Proposition 5.4]).
Assume (A_1). Then for any $0 \le s < t \le T$ and $u \in [0,T]$ we have

$$\begin{aligned}
&\mathbb{E}^{\overline{\mathcal{F}}_u} \sup_{r\in[u,T]} |Y_r - Y_r^{(s,t]}|^p \\
&+\mathbb{E}^{\overline{\mathcal{F}}_u} \left(\int_u^T 1_{(s,t]}(r)[|Z_r|^2 + |Z_r^{(s,t]}|^2] + [1 - 1_{(s,t]}(r)]|Z_r - Z_r^{(s,t]}|^2 dr \right)^{\frac{p}{2}} \\
\le\ & c_{(6.9)}^p \mathbb{E}^{\overline{\mathcal{F}}_u} \left(|\xi - \xi^{(s,t]}| + \int_u^T |f(r,Y_r,Z_r) - f^{(s,t]}(r,Y_r,Z_r)|dr \right)^p,
\end{aligned}$$

where $c_{(6.9)} > 0$ depends at most on $(T,d,p,L_y,L_z,(s_N)_N)$, and s_N is taken from $(A_1)(v)$.

[12] The V^l may depend on (ξ,f,Y,Z,p,q_l,r_l).

To be able to use the theory of weighted BMO, we assume that ξ and f satisfy a certain weighted BMO-assumption on a subinterval $[s,t] \subseteq [0,T]$. If this assumption holds on a subinterval $[s,t]$, then *on this interval* we will have that the solution of our BSDE is of weighted BMO, and this gives us a tail estimate for the variation of the solution.

(A_2) There are $0 \le s < t \le T$ such that there exist càdlàg $(\mathcal{F}_r)_{r\in[0,T]}$-supermartingales $(w^{\xi}_{p,s,u,t})_{u\in[s,t]}$ and $(w^{f}_{p,s,u,t})_{u\in[s,t]}$, whose canonical extensions $(\overline{w}^{\xi}_{p,s,u,t})_{u\in[s,t]}$ and[13] $(\overline{w}^{f}_{p,s,u,t})_{u\in[s,t]}$, satisfy for any $u \in [s,t]$

(i) $\mathbb{E}^{\mathcal{F}^0_u}|\xi - \xi^{(u,t]}|^p \le \overline{w}^{\xi}_{p,s,u,t}$,

(ii) $\mathbb{E}^{\mathcal{F}^0_u}\left(\int_u^T |f(r,Y_r,Z_r) - f^{(u,t]}(r,Y_r,Z_r)|dr\right)^p \le \overline{w}^{f}_{p,s,u,t}$.

Assumption (A_2) is in fact a condition imposed in the product probability space $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}})$, and we used, as usual, the notation (ξ, f, Y, Z) for the canonical extensions.
We made an exception with the weight processes w^{ξ}_p and w^{f}_p to emphasize the fact that our main result, a weighted John-Nirenberg-type theorem for BSDEs, is a result in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Theorem 6.10 ([B, Theorem 3.10 and Theorem 3.11])**.**
Assume (A_1) and (A_2). Then there exists $c_{(6.10)} > 0$ depending at most on $(T, d, p, L_y, L_z, (s_N)_N)$ such that for any stopping time $\sigma \in [s,t]$

$$\mathbb{E}^{\mathcal{F}_\sigma}|Y_t - Y_\sigma|^p \le (c_{(6.10)} w_{p,s,\sigma,t})^p, \tag{28}$$

where

$$\begin{aligned} w^p_{p,s,u,t} &= \left(w^{\xi}_{p,s,u,t} + w^{f}_{p,s,u,t}\right) + \mathbb{E}^{\mathcal{F}_u}\left(\int_u^t |f(r,0,0)|dr\right)^p \\ &\quad + (t-u)^p\left[\mathbb{E}^{\mathcal{F}_u}\left(|\xi| + \int_t^T |f(r,0,0)|dr\right)^p\right]. \end{aligned}$$

Consequently, there exists $c > 0$ such that for any stopping time $\sigma \in [s,t]$ one has

$$\begin{aligned} \mathbb{P}_B\left(\sup_{u\in[\sigma,t]} \frac{|Y_u - Y_\sigma|}{c_{(6.10)}} > \lambda + c\mu\nu\right) &\le e^{1-\mu}\mathbb{P}_B\left(\sup_{u\in[\sigma,t]} \frac{|Y_u - Y_\sigma|}{c_{(6.10)}} > \lambda\right) \\ &\quad + c\mathbb{P}_B\left(\sup_{u\in[\sigma,t]} w_{p,s,u,t} > \nu\right), \end{aligned}$$

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

[13] These processes are càdlàg $(\mathcal{F}^0_r)_{r\in[0,T]}$-supermartingales.

Proof. The first claim is a conditional version of Theorem 6.5, and can be proven similarly with use of assumption (A_2). The consequently-part follows from the first part of Theorem 6.10 applied together with Proposition 5.7 and Theorem 5.8 on the subinterval $[s,t] \subseteq [0,T]$. $\square$

For simplification, we may take $\lambda \to 0$ in the second part of Theorem 6.10, so that for all stopping times $\sigma \in [s,t]$ we have

$$\mathbb{P}_B\left(\sup_{u\in[\sigma,t]} \frac{|Y_u - Y_\sigma|}{c_{(6.10)}} > c\mu\nu\right) \le e^{1-\mu} + c\mathbb{P}_B\left(\sup_{u\in[\sigma,t]} w_{p,s,u,t} > \nu\right),$$

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

As an example of the case when (A_1) and (A_2) hold, we consider for $x \in \mathbb{R}^d$ the decoupled FBSDE

$$\begin{aligned} X_t &= x + \int_0^t b(r, X_r)dr + \int_0^t \sigma(r, X_r)dW_r, \quad t \in [0,T], \\ Y_t &= g(X_T) + \int_t^T h(r, X_r, Y_r, Z_r)dr - \int_t^T Z_r dW_r, \quad t \in [0,T]. \quad (29) \end{aligned}$$

For the functions b, σ, h, and g we assume that σ is uniformly bounded in addition to the usual conditions of joint continuity and uniform Lipschitz condition in the state variables (the precise set of conditions is given in [B, Section 3.1]). Then it follows that (A_1) holds with $\theta = 0$ and all $p \in [2,\infty)$, and (A_2) holds on all subintervals $[s,t] \subseteq [0,T]$ with the weight functions

$$w^f_{p,s,u,t} = w^\xi_{p,s,u,t} = c^p (t-u)^{\frac{p}{2}},$$

where $c > 0$ depends at most on $(T, d, p, g, h, \sigma, b)$.

Now, from Theorem 6.10 and from the properties of the forward process X, we obtain the following result:

Theorem 6.11 ([B, Theorem 3.1 and Theorem 3.2]).
Let the assumptions stated for FBSDE (29) *hold, and let* $p \ge 2$. *Then there exists* $c_{(6.11)} = c_{(6.11)}(p, b, \sigma, g, h) > 0$, *such that for all* $0 \le s < t \le T$ *and all stopping times* $\tau : \Omega \to [s,t]$ *we have*

$$\mathbb{E}\left(\frac{|Y_t - Y_\tau|^p}{1 + |X_\tau|^p (t-\tau)^{p/2}} \,\Big|\, \mathcal{F}_\tau\right) \le c^p_{(6.11)} (t-\tau)^{p/2}.$$

Consequently, there exists $c > 0$ such that for any $p \geq 2$, $0 \leq s < t \leq T$ and any stopping time $\tau \in [s,t]$ we have

$$\mathbb{P}_B\left(\sup_{u\in[\tau,t]} \frac{|Y_u - Y_\tau|}{c_{(6.11)}\sqrt{t-s}} > c\mu\nu\right) \leq e^{1-\mu} + c\mathbb{P}_B\left(\sup_{u\in[\tau,t]} |X_u|^p (t-u)^{\frac{p}{2}} > \nu^p - 1\right),$$

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

It should be mentioned that Theorem 6.11 holds in this form as g is Lipschitz, and the Lipschitz property is scaling invariant. Moreover, it is shown in [B, Example 3.3] that the weight process of Theorem 6.11 is sharp. That is, there exist b, σ, h, g that satisfy our assumptions, and the solution of FBSDE (29) satisfies

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

for all $p \geq 2$, all $0 \leq s < t \leq T$ and all stopping times $\tau : \Omega \to [s,t]$.

7 Some perspectives

We conclude with open questions and possible topics for future research.

1. Which parts of this thesis can be extended to the case where the driving process is a Lévy-process other than a Brownian motion? In particular, does Proposition 3.2 still hold when "continuity" is replaced by "càdlàg-property"? Some related work in this direction is in [14].

2. The condition (A_1)(v) is needed for the proof of Theorem 6.3, as it implies that for a certain martingale M, the Doléan-Dade exponential $\mathcal{E}(M)$ satisfies the reverse Hölder inequalities for $q \in (1,\infty)$ that are small enough. The critical index is defined in terms of the sliceability number $\mathrm{sl}_\infty(|\mathrm{Z}|^\theta)$, and if $\mathrm{sl}_\infty(|\mathrm{Z}|^\theta) = 0$, we have that the reverse Hölder inequalities are satisfied for all $q \in (1,\infty)$. However, if $M \in \mathrm{BMO}$, then [21, Theorem 3.8] provides the characterization

$$\mathcal{E}(\lambda M) \in \bigcap_{q\in(1,\infty)} \mathcal{RH}_q \text{ for all } \lambda \in \mathbb{R} \Leftrightarrow M \in \overline{L_\infty}^{\mathrm{BMO}}.$$

   Hence, it would be interesting to investigate whether the condition $(\int_0^t |Z_r|^\theta dB_r)_{t\in[0,T]} \in \overline{L_\infty}^{\mathrm{BMO}}$ for a 1-dimensional Brownian motion B implies that our results hold for all $p \in [2,\infty)$, i.e. that assumption (A_1)(vi) could be dropped in this case.

3. Is there a counterpart of Theorem 4.9 for Φ_r (see (19))? Similarly to [16, page 932], this might hold with the Derivative operator replaced by a Riemann-Liouville operator.

4. For the proof of the tail estimate of Theorem 6.10 on $[s,t]$ we only need an upper bound of $\mathbb{E}^{\mathcal{F}_u}|Y_t - \mathbb{E}^{\mathcal{F}_u} Y_t|^p$ or, equivalently, of $\mathbb{E}^{\mathcal{F}_u}|Y_t - Y_t^{(u,t]}|^p$ with $u \in [s,t]$. However, the result we used (Proposition 6.9) gives us an upper bound of $\mathbb{E}^{\mathcal{F}_u}\left(\sup_{r\in[t,T]}|Y_r - Y_r^{(u,t]}|^p\right)$. This raises the question whether it would suffice in (A_2), for example, to upper bound $\mathbb{E}^{\mathcal{F}_u}|\mathbb{E}^{\mathcal{F}_t}\xi - \mathbb{E}^{\mathcal{F}_u}\xi|^p$ instead of $\mathbb{E}^{\mathcal{F}_u}|\xi - \mathbb{E}^{\mathcal{G}_u^t}\xi|^p$, and to use this as an assumption in a modified version of Proposition 6.9. The difference of these concepts is that the latter one additionally measures how much ξ changes on the interval $[t,T]$ if it is perturbed on the interval $[u,t]$. My conjecture is that this stability is needed in order to find an upper bound of $\mathbb{E}^{\mathcal{F}_u}|Y_t - \mathbb{E}^{\mathcal{F}_u} Y_t|^p$, but since I have not found a counterexample, it remains open whether the first quantity is enough to give us a tail estimate or not.

References

[1] R. Avikainen: On irregular functionals of SDEs and the Euler scheme. *Finance Stoch.* 13, 381-401, 2009.

[2] J-M. Bismut: Conjugate convex functions in optimal stochastic control. *Journal of Mathematical Analysis and Applications*, Vol.44(2), 384-404, 1973.

[3] B. Bouchard, N. Touzi: Discrete Time Approximation and Monte-Carlo Simulation of Backward Stochastic Differential Equations. *Stochastic Process. Appl.* Vol. 111(2), 175-206, 2004.

[4] B. Bouchard, R. Elie, N. Touzi: Discrete-time approximation of BS-DEs and probabilistic schemes for fully nonlinear PDEs. *Radon Series Comp. Appl. Math*, 8, 1-34, 2009.

[5] P. Briand, B. Delyon, Y. Hu, E. Pardoux, L. Stoica: L^p solutions of Backward Stochastic Differential Equations. *Stochastic Process. Appl.* Vol. 108(1), 109-129, 2003.

[6] P. Briand, R. Elie: A simple constructive approach to quadratic BSDEs with or without delay. *Stoch. Proc. App.*, Vol. 123(8), 2921-2939, 2013.

[7] P. Briand, Y. Hu: BSDE with quadratic growth and unbounded terminal value. *Probab. Theory Related Fields*, 136(4), 604-618, 2006.

[8] R. Buckdahn, J. Li: Stochastic Differential Games and Viscosity Solutions of Hamilton-Jacobi-Bellman-Isaacs Equations. *SIAM J. Control Optim.*, Vol.47(1), 444-475, 2008.

[9] F. Delbaen, Y. Hu, A. 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.

[10] F. Delbaen and S. Tang: Harmonic analysis of stochastic equations and backward stochastic differential equations. *Probability Theory and Related Fields*, 146, 291-336, 2010.

[11] G. Dos Reis: On some properties of solutions of quadratic growth BSDE and applications in finance and insurance. PhD Thesis, Humboldt University in Berlin, 2010.

[12] N. El Karoui, S. Peng, M.C. Quenez: Backward Stochastic Differential Equations in Finance. *Mathematical Finance*, Vol.7(1), 1-71, 1997.

[13] C. Geiss, S. Geiss, E. Gobet: Generalized fractional smoothness and L_p-variation of BSDEs with non-Lipschitz terminal condition. *Stoc. Proc. and Appl.*, Vol.122(5), 2078-2216, 2012.

[14] C. Geiss, A. Steinicke: L_2-variation of Lévy driven BSDEs with non-smooth terminal condition. http://arxiv.org/abs/1212.3420v3

[15] S. Geiss: Weighted BMO and discrete time hedging within the Black-Scholes model. *Probab. Theory Relat. Fields*, 132, 13-38, 2005.

[16] S. Geiss, A. Toivola: Weak convergence of error processes in discretizations of stochastic integrals and Besov spaces. *Bernoulli*, Vol.15(4), 925-954, 2009.

[17] S. Geiss, A. Toivola: On fractional smoothness and L_p-approximation on the Gaussian space. *Ann. Probab.*, 2014.

[18] F. Hirsch: Lipschitz functions and fractional Sobolev spaces. *Potential Analysis*, 11, 415-429, 1999.

[19] Y. Hu, P. Imkeller, M. Müller: Utility maximization in incomplete markets. *Ann. Appl. Probab.*, Vol.15(3), 1691-1712, 2005.

[20] P. Imkeller: Malliavin's calculus and applications in stochastic control and finance. *IM PAN Lecture Notes*, Warsaw 2009.

[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] P. Malliavin: Integration and Probability. *Grad. Texts in Math.*, 157, Springer, New York, 1995.

[24] 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.

[25] D. Nualart: The Malliavin Calculus and Related Topics. 2nd edition, Springer, 2006.

[26] E. Pardoux, S. Peng: Adapted Solution of a Backward Stochastic Differential Equation. *Systems Control Lett.*, 14, 55-61, 1990.

[27] 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.

[28] A. Pietsch: Eigenvalues and s-Numbers. *Monograph, Akademische Verlagsgesellschaft Geest & Portig K.-G., Leipzig*, 1987.

[29] W. Schachermayer: A characterisation of the closure of H^∞ in BMO. *Séminaire de probabilités (Strasbourg)*, Vol.30, 344-356, 1996.

114. LUIRO, HANNES, Regularity properties of maximal operators. (11 pp.) 2008
115. VIHOLAINEN, ANTTI, Prospective mathematics teachers' informal and formal reasoning about the concepts of derivative and differentiability. (86 pp.) 2008
116. LEHRBÄCK, JUHA, Weighted Hardy inequalities and the boundary size. (21 pp.) 2008
117. NISSINEN, KARI, Small area estimation with linear mixed models from unit-level panel and rotating panel data. (230 pp.) 2009
118. BOJARSKI, B.V., Generalized solutions of a system of differential equations of the first order and elliptic type with discontinuous coefficients. (64 pp.) 2009
119. RAJALA, TAPIO, Porosity and dimension of sets and measures. (22 pp.) 2009
120. MYLLYMÄKI, MARI, Statistical models and inference for spatial point patterns with intensity-dependent marks. (115 pp.) 2009
121. AVIKAINEN, RAINER, On generalized bounded variation and approximation of SDEs. (18 pp.) 2009
122. ZÜRCHER, THOMAS, Regularity of Sobolev–Lorentz mappings on null sets. (13 pp.) 2009
123. TOIVOLA, ANNI, On fractional smoothness and approximations of stochastic integrals. (19 pp.) 2009
124. VIHOLA, MATTI, On the convergence of unconstrained adaptive Markov chain Monte Carlo algorithms. (29 pp.) 2010
125. ZHOU, YUAN, Hajłasz–Sobolev extension and imbedding. (13 pp.) 2010
126. VILPPOLAINEN, MARKKU, Recursive set constructions and iterated function systems: separation conditions and dimension. (18 pp.) 2010
127. JULIN, VESA, Existence, uniqueness and qualitative properties of absolute minimizers. (13 pp.) 2010
128. LAMMI, PÄIVI, Homeomorphic equivalence of Gromov and internal boundaries. (21 pp.) 2011
129. ZAPADINSKAYA, ALEKSANDRA, Generalized dimension distortion under Sobolev mappings. (18 pp.) 2011
130. KEISALA, JUKKA, Existence and uniqueness of $p(x)$-harmonic functions for bounded and unbounded $p(x)$. (56 pp.) 2011
131. SEPPÄLÄ, HEIKKI, Interpolation spaces with parameter functions and L_2-approximations of stochastic integrals. (18 pp.) 2011
132. TUHOLA-KUJANPÄÄ, ANNA, On superharmonic functions and applications to Riccati type equations. (17 pp.) 2012
133. JIANG, RENJIN, Optimal regularity of solutions to Poisson equations on metric measure spaces and an application. (13 pp.) 2012
134. TÖRMÄKANGAS, TIMO, Simulation study on the properties of quantitative trait model estimation in twin study design of normally distributed and discrete event-time phenotype variables. (417 pp.) 2012
135. ZHANG, GUO, Liouville theorems for stationary flows of generalized Newtonian fluids. (14 pp.) 2012
136. RAJALA, TUOMAS, Use of secondary structures in the analysis of spatial point patterns. (27 pp.) 2012
137. LAUKKARINEN, EIJA, On Malliavin calculus and approximation of stochastic integrals for Lévy processes. (21 pp.) 2012
138. GUO, CHANGYU, Generalized quasidisks and the associated John domains. (17 pp.) 2013
139. ÄKKINEN, TUOMO, Mappings of finite distortion: Radial limits and boundary behavior. (14 pp.) 2014
140. ILMAVIRTA, JOONAS, On the broken ray transform. (37 pp.) 2014
141. MIETTINEN, JARI, On statistical properties of blind source separation methods based on joint diagonalization. (37 pp.) 2014
142. TENGVALL, VILLE, Mappings of finite distortion: Mappings in the Sobolev space $W^{1,n-1}$ with integrable inner distortion. (22 pp.) 2014
143. BENEDICT, SITA, Hardy-Orlicz spaces of quasiconformal mappings and conformal densities. (16 pp.) 2014
144. OJALA, TUOMO, Thin and fat sets: Geometry of doubling measures in metric spaces. (19 pp.) 2014
145. KARAK, NIJJWAL, Applications of chaining, Poincaré and pointwise decay of measures. (14 pp.) 2014
146. JYLHÄ, HEIKKI, On generalizations of Evans and Gangbo's approximation method and L^∞ transport. (20 pp.) 2014
147. KAURANEN, AAPO, Space-filling, energy and moduli of continuity. (16 pp.) 2015

ISBN 978-951-39-6193-0
ISSN 1457-8905