DISS. ETH No. 19705

# FUNCTIONAL DIFFERENTIAL APPROACHES TO BACKWARD STOCHASTIC EQUATIONS

A dissertation submitted to
ETH ZURICH

for the degree of
Doctor Of Sciences

presented by

MATTEO CASSERINI
Dipl. Math. ETH
born 19 October 1983
citizen of Cerentino (TI), Switzerland

accepted on the recommendation of

| | |
|---|---|
| Prof. Dr. Martin Schweizer | examiner |
| Prof. em. Dr. Freddy Delbaen | co-examiner |
| Prof. Dr. Terry Lyons | co-examiner |
| Prof. Dr. Halil Mete Soner | co-examiner |
| Prof. Dr. Josef Teichmann | co-examiner |

2011

*To my family*

# Abstract

This thesis is mainly concerned with the study of backward stochastic differential equations (BSDEs), introduced by Pardoux and Peng [58], and of their generalizations. Recently, Liang, Lyons and Qian [49] developed a new approach to Lipschitz BSDEs. Their method is based on the analysis of a particular class of functional differential equations, where the driver of the equation does not depend only on the present, but also on the terminal value of the solution.

The first part is dedicated to the introduction of various classes of functional differential equations, associated to several types of backward systems. After first studying BSDEs of quadratic growth with respect to the variable $z$, we then consider systems of fully coupled forward-backward stochastic differential equations (FBSDEs). Both types of equation have numerous applications in many areas of mathematical finance (for instance utility maximization problems or hedging problems for large investors). Afterwards, we analyze a numerical scheme for the approximation of decoupled systems of functional differential equations, which is based on the local iteration approach introduced in the fully coupled case.

In the last part, we introduce complexification techniques for stochastic processes, which are closely related to Hermite polynomials. These techniques allow to consider real-valued stochastic processes as projections of corresponding complex-valued processes, and we believe they could provide an important tool in the study of stochastic differential equations. We conclude by presenting an application of these techniques to a result obtained by Widder [70] concerning a class of Brownian martingales.

# Sommario

L'argomento principale di questa tesi è lo studio di equazioni differenziali stocastiche regressive (o backward), introdotte da Pardoux e Peng [58], e delle loro generalizzazioni. Recentemente, Liang, Lyons e Qian [49] hanno sviluppato un nuovo approccio ad equazioni regressive aventi continuità lipschitziana. Il loro metodo si basa sull'analisi di una particolare classe di equazioni differenziali funzionali, in cui il coefficiente dell'equazione dipende non solo dal valore presente, ma anche da quello terminale della soluzione.

La prima parte è dedicata all'introduzione di varie classi di equazioni differenziali funzionali, associate a diversi tipi di sistemi regressivi. In primo luogo, studiamo equazioni regressive aventi crescita quadratica rispetto alla variabile $z$, per poi dedicarci a sistemi accoppiati di equazioni differenziali stocastiche progressive-regressive (o forward-backward). Entrambi questi tipi di equazione hanno numerose applicazioni in molti ambiti di matematica finanziaria (ad esempio problemi di massimizzazione dell'utilità attesa, o problemi di hedging per grandi investitori). In seguito, analizziamo uno schema numerico per l'approssimazione di sistemi disaccoppiati di equazioni differenziali funzionali, basato sull'approccio di iterazione locale introdotto per sistemi accoppiati.

Nell'ultima parte, introduciamo tecniche di estensione allo spazio complesso per processi stocastici, strettamente correlate ai polinomi di Hermite. Tali tecniche permettono di considerare processi stocastici a valori reali come proiezioni di corrispondenti processi a valori complessi, e crediamo che possano costituire uno strumento importante nello studio di equazioni differenziali stocastiche. Terminiamo presentando un'applicazione di queste tecniche a un risultato ottenuto da Widder [70] riguardo a una classe di martingale Browniane.

# Acknowledgments

First of all, I would like to express my deepest and sincere gratitude to my 'de facto' supervisor, Prof. Freddy Delbaen, for his excellent support during my doctoral studies. I am grateful for the numerous stimulating discussions and for his guidance through the academic world. His deep knowledge of mathematics and his advices have been priceless, and his suggestions substantially improved this thesis.

Sincere thanks are also due to Prof. Schweizer, for agreeing to act as the official supervisor of this thesis and for his help with the administrative matters, and to Prof. Soner and Prof. Teichmann, who readily accepted to be my internal co-examiners. I am also indebted to Prof. Lyons, who kindly agreed to act as external reporter and hosted a research stay at Oxford University: I greatly benefited from the several interesting discussions and from the inspiring and pleasant atmosphere at the Mathematical Institute.

A special thanks goes to my coauthor, Gechun Liang, for hosting a visit to the Oxford-Man Institute and for the excellent collaboration. Further thanks are also due to my friends and colleagues at and outside ETH, especially to the members of Group 3 for creating an enjoyable and stimulating working environment. Moreover, I gratefully acknowledge the financial support by ETH Zurich.

Finally, I wish to express my deep gratitude to my mother, my brothers, my grandmother and Chiara for their continuous confidence, understanding and encouragement during the years leading up to this work. This thesis is dedicated to them.

# Contents

# Chapter 1

# Introduction

Backward stochastic differential equations, forward-backward stochastic differential equations and their generalizations have been subject of extensive research in the last twenty years. These equations have quickly become a central tool in stochastic analysis, especially due to their intimate relation with stochastic control, mathematical finance and partial differential equations: the applications in these fields are countless. After giving a brief overview of backward stochastic differential equations and related approaches in Section 1.1 and Section 1.2, we present the main results of the thesis in Section 1.3.

## 1.1 Backward stochastic differential equations

In the following, let $T > 0$. Assume that $(W_t)_{t\in[0,T]}$ is an $m$-dimensional Brownian motion defined on a probability space $(\Omega, \mathcal{F}, P)$, and denote by $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ the natural filtration of $(W_t)_{t\in[0,T]}$ and augmented by the $P$-nullsets of $\mathcal{F}$. We recall that filtrations are normally used in stochastic analysis to model flows of informations (in particular, $\mathcal{F}_t$ represents for $t \in [0, T]$ the history of $W$ on $[0, t]$ and thus the information available up to time $t$). A stochastic process $(X_t)_{t\in[0,T]}$ is said to be $\mathbb{F}$-adapted if, for every $t \in [0, T]$, $X_t$ is $\mathcal{F}_t$-measurable: this means that, given the information in the $\sigma$-field $\mathcal{F}_t$, the historical behaviour of $X$ up to $t$ is known.

Assume now that $f : \Omega \times [0, T] \times \mathbb{R}^d \times \mathbb{R}^{d\times m} \to \mathbb{R}^d$ is a random function and $\xi$ is an $\mathcal{F}_T$-measurable random variable. Backward stochastic differential equations (BSDEs) are equations of the form

$$\begin{cases} dY_t = -f(t, Y_t, Z_t)dt + Z_t dW_t, & t \in [0, T], \\ Y_T = \xi. \end{cases}$$

This differential formulation of the BSDE is just a compact notation for the following equivalent integral formulation:

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

A solution of the BSDE is then a pair of *adapted* processes $(Y, Z)$ satisfying appropriate measurability and integrability assumptions and such that the latter integral formulation is fulfilled for all $t \in [0, T]$. In this case, $Y$ is called the solution part and $Z$ the control part.

The problem is completely determined by the pair $(f, \xi)$, which is thus called the generator of the BSDE. $f$ is often referred to as the driver of the BSDE, while $\xi$ is called the terminal condition of the BSDE: this is due to the fact that the solution process $Y$ should satisfy $Y_T = \xi$ (as can be seen by taking $t = T$ in the above integral formulation).

Moreover, let us observe that the necessity of the control process $Z$ is closely related to the randomness of the equation: if the driver $f$ and the terminal condition $\xi$ are deterministic, we can choose $Z \equiv 0$, and the solution part $Y$ can be obtained by solving an ordinary differential equation.

The reader familiar with the theory of ordinary differential equations may wonder why BSDEs should be more difficult to solve with respect to forward stochastic differential equations (SDEs): namely, it is well known that ordinary differential equations with given terminal conditions can be converted into equations with initial conditions via time inversion. This approach, however, cannot be used for BSDEs mainly because of the adaptedness constraint, which prevents us from using time inversion.

This is best illustrated by the following basic example, which exhibits once more the need for an auxiliary control process $Z$. Consider the simple BSDE

$$\begin{cases} dY_t = 0, \\ Y_T = \xi. \end{cases}$$

Obviously, the only possible solution is the constant process $Y \equiv \xi$. This process, however, is adapted to the filtration only if the terminal condition $\xi$ is $\mathcal{F}_0$-measurable: in other words, the above BSDE has a solution only if $\xi$ is deterministic. This leads us to a natural question: if $\xi$ is just $\mathcal{F}_T$-measurable, how can we find an adapted process with terminal value $\xi$? From the point of view of martingale theory, the most simple choice is to associate to $\xi$ its corresponding martingale $(E[\xi|\mathcal{F}_t])_{t\in[0,T]}$. But the well known Itô representation theorem for Brownian martingales implies (under suitable integrability assumptions) that $E[\xi|\mathcal{F}_t]$ is of the form $\int_0^t \widetilde{Z}_s dW_s$ for some process $\widetilde{Z}$. Therefore, $(E[\xi|\mathcal{F}_t], \widetilde{Z}_t)_{t\in[0,T]}$ is the solution of the following BSDE:

$$\begin{cases} dY_t = Z_t dW_t, \\ Y_T = \xi. \end{cases}$$

If we compare this BSDE to the previous one, we see that the adaptedness constraint introduces a diffusion part, expressed in terms of the control process $Z$, while the drift part remains 0. Then, the general formulation of the BSDE is obtained by allowing the drift to depend on $(\omega, t)$ as well as on the solution $(Y, Z)$.

As already anticipated, BSDEs have found many applications in mathematical finance: in numerous problems, the terminal condition can namely be specified in terms of a random payoff due at the termination of the contract, while the dynamics of the BSDE are given by the modeling assumptions and can then be calibrated by observation of the historical behaviour. This is best illustrated by considering the classical problem of replicating an European call option in a complete Black-Scholes market. Assume that the price of the underlying asset $S_t$ follows a geometric Brownian motion: the payoff of an European call option with maturity $T$ and strike price $K$ is given by

$\max\{S_T - K, 0\}$. Moreover, assume that we have an investor who wants to hedge this call option: he can do this by investing in the market to replicate the option payoff, and he is interested in the trading strategy he has to adopt as well as the initial capital required.

We can then model the replicating portfolio via a BSDE: we explain briefly the intuition (for more details, the reader is referred to the survey article [32]). Since the portfolio should replicate the call option, its terminal value should equal the final payoff of the option, i.e. $Y_T = \max\{S_T - K, 0\}$. On the other hand, the dynamics can be easily deduced if we assume, for instance, that the investor trades in a self-financing way (i.e. all portfolio changes are financed with the bank account, without injecting or draining extra funds). The solution part $Y$ of the BSDE then represent the value of the replicating portfolio, while the control part $Z$ represents the trading strategy.

BSDEs and their applications have been studied in hundreds of articles and several books during the last two decades, and it is not possible to give an exhaustive list of the literature on the subject. We limit ourselves to a brief historical summary, and refer the reader to the books [31, 51] and the surveys [30, 32] for a more extensive overview.

BSDEs have been first introduced in the linear case by Bismut [6] in 1973. Later, his work has been extended by Pardoux and Peng [58] to the general non-linear case, and BSDEs have since then gained considerable attention. Pardoux and Peng were able to obtain the existence and uniqueness of solutions by relying on Itô's representation theorem and on Picard iteration arguments, when the terminal condition is square-integrable and the driver $f$ satisfies a Lipschitz condition with respect to $y$ and $z$. The connection with partial differential equations (PDEs) was soon recognized by the same authors [57, 59], who derived a probabilistic representation for solutions of semilinear parabolic PDEs in terms of BSDEs: this is often referred to as the non-linear Feynman-Kac representation. This connection, together with the numerous financial applications, motivated many authors to develop probabilistic numerical schemes for the approximation of Lipschitz BSDEs (see for example [4, 7, 16, 35, 73]).

There have been several different generalizations of the non-linear BSDEs of Pardoux and Peng. For instance, a natural extension is to consider systems where a BSDE is coupled with a forward SDE: such a system is called forward-backward stochastic differential equation (FBSDE). While simple types of decoupled FBSDEs had already been considered by Pardoux and Peng, the study of fully coupled FBSDEs has been initiated by Antonelli [1]. Fully coupled FBSDEs were later studied by several authors, see for instance [40, 50, 60, 61, 71]: an extensive account on the subject can be found in the book [51].

A second possible generalization of Pardoux and Peng's BSDEs consists in weakening the Lipschitz assumption on the driver $f$. The problem has been tackled by several authors: probably the most remarkable works in this direction are those of Kobylanski [46] and Lepeltier and San Martín [48], who studied scalar BSDEs when the terminal condition is bounded and the driver has at most quadratic growth with respect to the variable $z$. These BSDEs are usually called quadratic BSDEs, and arise naturally in many problems in mathematical finance and stochastic control. Recently, the results on quadratic BSDEs have been substantially improved [9, 10, 21, 22].

Finally, we point out that, while most authors limit themselves to the Brownian setting, it is also possible to consider BSDEs defined on general probability spaces: in this case, it is not possible to rely on Itô's representation theorem anymore, and the natural extension is to consider other types of martingale representations, given by orthogonal decompositions. For more details, we refer to the book [31] and the survey [32] for Lipschitz BSDEs, and to the works of Morlais [55] and Tevzadze [69] for quadratic BSDEs.

## 1.2 The approach via functional differential equations

Recently, Liang, Lyons and Qian [49] introduced a new approach for Lipschitz BSDEs on a general filtration: their method is purely probabilistic and does not rely a priori on the existence of martingale representations. We give a brief overview of their techniques, as they will be useful throughout the thesis. We begin with an illustrative example: for $T > 0$, assume that we have a special semimartingale $(Y_t)_{t\in[0,T]}$ on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P)$ satisfying the usual assumptions, and let the terminal value $Y_T = \xi \in L^1(\mathcal{F}_T)$ be given. Furthermore, assume that the canonical decomposition of $Y$ is given by

$$Y_t = M_t - V_t,$$

where $M$ is a martingale and $V$ a predictable process of finite variation with $V_0 = 0$. Then, if $V_T$ is integrable, it is easy to verify that, for all $t \in [0,T]$,

$$\begin{aligned} M_t &= E[M_T|\mathcal{F}_t] = E[\xi + V_T|\mathcal{F}_t] =: \mathcal{M}(\xi, V)_t, \\ Y_t &= M_t - V_t = E[\xi + V_T|\mathcal{F}_t] - V_t =: \mathcal{Y}(\xi, V)_t, \end{aligned} \tag{1.2.1}$$

where the operators $\mathcal{M}$ and $\mathcal{Y}$ are defined on $L^1(\mathcal{F}_T) \times \mathscr{C}$, $\mathscr{C}$ denoting the class of $\mathbb{R}^d$-valued adapted processes $V$ on $[0,T]$ such that $V_T \in L^1(\mathcal{F}_T)$. In other words, the semimartingale $Y$ and the martingale $M$ can be expressed as functionals of the terminal value $\xi$ and the finite variation process $V$.

We show now, with the help of some intuitive arguments, how these operators $\mathcal{M}$ and $\mathcal{Y}$ can be used to derive an alternative approach to Lipschitz BSDEs. Assume for the moment that we are on a probability space $(\Omega, \mathcal{F}, P)$ with a $m$-dimensional Brownian motion $W = (W_t)_{t\in[0,T]}$ on it. Let $(\mathcal{F}_t)_{t\in[0,T]}$ be its augmented filtration, and consider a classical BSDE of the form

$$\begin{cases} dY_t = -f(t, Y_t, Z_t)dt + Z_t dW_t, \\ Y_T = \xi, \end{cases} \tag{1.2.2}$$

where $\xi$ is $\mathcal{F}_T$-measurable and $f : \Omega \times [0,T] \times \mathbb{R}^d \times \mathbb{R}^{d\times m} \to \mathbb{R}^d$ satisfies the usual measurability assumptions. Then, by the definition of the operators $\mathcal{Y}$ and $\mathcal{M}$, it seems plausible to associate the above BSDE to the forward differential equation

$$V_t = \int_0^t f(s, \mathcal{Y}(\xi, V)_s, \mathcal{Z}(\xi, V)_s)ds, \tag{1.2.3}$$

where $\mathcal{Z}(\xi, V)$ is defined implicitly as the integrand process in the Itô representation of the martingale $\mathcal{M}(\xi, V)$, i.e.

$$\mathcal{M}(\xi, V)_T = E[\mathcal{M}(\xi, V)_T] + \int_0^T \mathcal{Z}(\xi, V)_s dW_s. \tag{1.2.4}$$

The reader can easily verify that (1.2.3) has a solution if and only if the BSDE (1.2.2) is solvable. Namely, if (1.2.3) has a solution $V$, then a simple computation yields that the couple $(\mathcal{Y}(\xi, V), \mathcal{Z}(\xi, V))$ given by (1.2.1) and (1.2.4) solves (1.2.2). Conversely, if $(Y, Z)$ solves (1.2.2), then we can construct a solution of (1.2.3) via the canonical decomposition of $Y$. This leads us to the equivalence between classical, Lipschitz BSDEs and forward equations of the type (1.2.3).

The peculiarity of the forward differential equation (1.2.3) consists in the fact that the driver $f$ depends not only on the behaviour of the process up to the present value, but also on the terminal value $V_T$ of the solution: such stochastic differential equations are not standard, and for this reason we will often refer to them as *functional differential equations* (note that the term "functional differential equations" is often used in the literature to refer to stochastic delay differential equations: however, contrary to the latter, the drivers of the equations studied here do not have any delay in the past, but rather in the future).

As shown by Liang, Lyons and Qian, such an interpretation of BSDEs can be made rigorous and extended to a much more general framework. The authors mainly considered two extensions: first of all, they worked on a general filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P)$ satisfying only the usual assumptions of right-continuity and completeness. Second, they assumed that the driver $f$ depends on some general process $\mathcal{L}(M)$ instead that on $Z$, where $\mathcal{L}$ is an abstract functional mapping martingales into some space of adapted processes. This is an important extension, since it allows both to take into account the generality of the filtration $\mathbb{F}$ and to treat other types of backward equations not fitting in the classical framework.

For $\xi$ $\mathcal{F}_T$-measurable and $f : \Omega \times [0, T] \times \mathbb{R}^d \times \mathbb{R}^p \to \mathbb{R}^d$, *backward stochastic* *dynamics* are equations of the form

$$\begin{cases} dY_t = -f(t, Y_t, \mathcal{L}(M)_t)dt + dM_t, \\ Y_T = \xi. \end{cases} \tag{1.2.5}$$

Similarly to BSDEs, a solution to (1.2.5) is then a pair of adapted processes $(Y, M)$ with appropriate integrability and measurability assumptions, and satisfying the integral formulation of (1.2.5). We note that the problem is completely determined by the generators $(f, \xi, \mathcal{L})$ and, as in the previous Brownian example, it can be reformulated as a functional differential equation by replacing $Y$, $M$ with the operators $\mathcal{Y}$, $\mathcal{M}$ given by (1.2.1). This leads us to

$$dV_t = f(t, \mathcal{Y}(\xi, V)_t, \mathcal{L}(\mathcal{M}(\xi, V))_t)dt, \quad V_0 = 0. \tag{1.2.6}$$

Liang, Lyons and Qian studied the problem of existence and uniqueness of solutions to the functional differential equation (1.2.6) in a $L^2$-framework, when the driver $f$ satisfies the standard Lipschitz assumption and $\xi$ is square-integrable, but for general functionals $\mathcal{L}$. This is obtained in two steps: first of all, they obtained the existence of a unique solution when the time horizon $T$ is sufficiently small, provided that the functional $\mathcal{L}$ satisfies some boundedness and Lipschitz conditions with respect to appropriate $L^2$-spaces. These conditions turn out to be rather mild and allow to treat many different types of functionals (even non-local ones). The main result of Liang, Lyons and Qian can be stated as follows:

**Theorem 1.2.1** (Liang et al. [49]). *Assume that $(f, \xi)$ satisfies the standard Lipschitz assumptions, and let $\mathcal{L}$ be bounded and Lipschitz with respect to appropriate $L^2$-spaces. Then there is a constant $\ell > 0$, depending only on the Lipschitz constants of $f$ and $\mathcal{L}$, such that the functional differential equation* (1.2.6) *admits a unique square-integrable solution $V$ for $T < \ell$.*

The theorem is shown via a Picard iteration argument. After proving the existence of a unique solution locally, the next step consists in extending Theorem 1.2.1 to any time horizon $T > 0$. However, the boundedness and Lipschitz assumptions on the functional $\mathcal{L}$ are not enough to guarantee the existence of solutions for arbitrary time horizons $T$ and general coefficients $(f, \xi)$ (mainly due to the possible non-locality of $\mathcal{L}$). This poses natural boundaries to the extension of Theorem 1.2.1.

In order to solve this problem, several methods are possible: a possibility is to consider only some concrete example of functional $\mathcal{L}$, and to then develop tailor-made techniques for each $\mathcal{L}$ in order to identify conditions on $(f, \xi)$ sufficient to obtain global solutions (for example, this is the approach applied by Delong and Imkeller [26] to BSDEs with time delayed generators).

Liang, Lyons and Qian chose instead to keep the functional $\mathcal{L}$ abstract. The extension to global intervals is obtained by imposing additional conditions on $\mathcal{L}$, while keeping the standard assumptions on $(f, \xi)$ intact, and the global solution to (1.2.6) is constructed by patching together the local solutions. More exactly, it can be shown that the desired extension holds if the functional $\mathcal{L}$ satisfies two conditions, called local-in-time property and differential property. Roughly speaking, the local-in-time property means that $\mathcal{L}(M)$ is defined locally: $\mathcal{L}(M)_t$ should only depend on $(M_s)_{s\in(t-\varepsilon,t+\varepsilon)}$ for $\varepsilon > 0$ arbitrarily small. The differential property tells us instead that $\mathcal{L}(M)$ should depend only on the increments of the martingale $M$. These properties then lead us to the following result:

**Theorem 1.2.2** (Liang et al. [49])**.** *Assume that the assumptions of Theorem 1.2.1 hold true, and let $\mathcal{L}$ satisfy the local-in-time and differential properties. Then, the functional differential equation* (1.2.6) *has a unique square-integrable solution $V$ for any $T > 0$.*

**Remark 1.2.3.** We conclude this section by suggesting a possible new direction of research, inspired by the introduction of this type of functional differential equations. Considering equation (1.2.6), it might be interesting to study the case where the operators $\mathcal{Y}$ and $\mathcal{M}$ are not necessarily given by the expressions (1.2.1), but rather by other more general functionals of $\xi$ and $V$. Such a generalization opens the door to a new class of stochastic delay differential equations, where the coefficients are delayed not only to the past, but also to the future behaviour of the solution process: a general study of this class of equations may bring exciting new insights in stochastic analysis.

The idea is best illustrated by considering the example of BSDEs with time delayed generators, which have been introduced by Delong and Imkeller [26]. Recently, Dos Reis et al. [27] studied a particular subclass of backward delayed equations, which can be written in the form

$$\begin{cases} dY_t = -f\Big(t, \int_{-t}^{0} Y_{t+s}\alpha_Y(ds), \int_{-t}^{0} Z_{t+s}\alpha_Z(ds)\Big)dt + Z_t dW_t, \\ Y_T = \xi, \end{cases}$$

where $\alpha_Y$ and $\alpha_Z$ are some given non-random finite measures supported on $[-T, 0]$. It is not difficult to reinterpret these backward equations as forward SDEs delayed to the whole path behaviour of the solution: namely, we can write

$$dV_t = f(t, \mathbb{Y}(\xi, V)_t, \mathcal{L}(\mathcal{M}(\xi, V))_t)dt, \quad V_0 = 0,$$

where the functional $\mathbb{Y}$ and $\mathcal{L}$ are given by

$$\begin{aligned} \mathbb{Y}(\xi, V)_t :&= \int_{-t}^{0} \mathcal{Y}(\xi, V)_{t+s}\alpha_Y(ds) \\ &= \int_{-t}^{0} \Big(E[\xi + V_T|\mathcal{F}_{t+s}] - V_{t+s}\Big)\alpha_Y(ds), \\ \mathcal{L}(\mathcal{M}(\xi, V))_t :&= \int_{-t}^{0} \mathcal{Z}(\xi, V)_{t+s}\alpha_Z(ds), \quad t \in [0, T], \end{aligned}$$

and $\mathcal{Z}$ denotes the operator introduced in (1.2.4).

## 1.3 Results of the thesis

In the following, we briefly summarize the main results and contributions of this thesis.

**Backward stochastic dynamics of quadratic growth.** Motivated by the approach of Liang, Lyons and Qian, we consider a general filtered probability space, and study functional differential equations of the form

$$V_t = \int_0^t \big(f\big(s, \mathcal{Y}(\xi, V)_s, \mathcal{L}(\mathcal{M}(\xi, V))_s\big) - f(s, 0, 0)\big)ds, \quad t \geq 0,$$

where the operators $\mathcal{M}$ and $\mathcal{Y}$ are defined as in (1.2.1), $\mathcal{L}$ is a general abstract functional, $\xi$ is bounded, and $f = f(\omega, t, y, z)$ satisfies, instead of the classical Lipschitz condition, a particular quadratic growth condition with respect to $z$. The quadratic growth assumption forces us to study these equations in a framework different from the usual $L^2$-setting. We thus introduce an appropriate $BMO$-space of continuous stochastic processes, which are not necessarily martingales.

By working in this space, the existence of a unique solution to the quadratic functional differential equation can then be derived as follows. First of all, we show the existence and uniqueness of solutions for small terminal conditions $\xi$, by applying first a contraction argument on the $BMO$-space we introduced, and then a change of measure. This approach is related to the work of Tevzadze [69], and gives us the following result:

**Lemma 1.3.1** (Chapter 2, Lemma 2.2.5). *Let $f$ satisfy appropriate quadratic growth assumptions, and assume that $\|\xi\|_\infty \leq \Theta$, where $\Theta$ depends on $T$ and on the growth constants of $f$. Then, the quadratic functional differential equation has a unique $BMO$-solution $V$.*

Moreover, under additional compatibility conditions on $f$, we can extend the solvability to arbitrary bounded terminal conditions $\xi$. Namely, by rewriting $\xi$ as $\xi = \sum_{i=1}^n \xi^i$ with $\|\xi^i\|_\infty \leq \Theta$, we can consider for each $i$ a subproblem with terminal condition $\xi^i$, each being solvable via Lemma 1.3.1. The subproblems can then be combined to give a solution to the original equation, and the uniqueness follows by comparison arguments.

**Fully coupled forward-backward dynamics.** Motivated by the applications of fully coupled FBSDEs, we then devote ourselves to the study of fully coupled systems of functional differential equations on general filtered probability spaces. In order to better reflect the coupling between the equations of such systems, we have to consider modified versions of the operators $\mathcal{Y}$ and $\mathcal{M}$. These are given, for $t \in [0, T]$, by

$$\mathcal{M}^\phi(X, V)_t := E[\phi(X_T) + V_T | \mathcal{F}_t], \quad \mathcal{Y}^\phi(X, V)_t := \mathcal{M}^\phi(X, V)_t - V_t,$$

where $\phi : \Omega \times [0,T]$ is a function which expresses the terminal condition of $(\mathcal{Y}^\phi(X,V)_t)_{t\geq 0}$. We write for simplicity $\mathcal{M} = \mathcal{M}^\phi$, $\mathcal{Y} = \mathcal{Y}^\phi$, and study fully coupled systems of the form

$$\begin{cases} dX_t = \mu(t, X_t, \mathcal{Y}(X,V)_t, \mathcal{L}^1(\mathcal{M}(X,V))_t)dt + \sigma(t, X_t, \mathcal{Y}(X,V)_t, \mathcal{L}^2(\mathcal{M}(X,V))_t)dW_t, \\ dV_t = f(t, X_t, \mathcal{Y}(X,V)_t, \mathcal{L}^3(\mathcal{M}(X,V))_t)dt, \\ X_0 = x, \quad V_0 = 0, \end{cases}$$

where $\mu$, $\sigma$, $f$ are random functions, and $\mathcal{L}^1$, $\mathcal{L}^2$, $\mathcal{L}^3$ are general abstract functionals. Our attention is focused on the existence and uniqueness of solutions to these systems of equations, by working in an appropriate $L^2$-framework. First of all, we obtain the existence of a unique solution for general coefficients and under very weak assumptions of functionals $\mathcal{L}^i$, by assuming that the time interval is sufficiently small. This is the main result we obtain for fully coupled systems, and can be stated as follows:

**Theorem 1.3.2** (Chapter 3, Theorem 3.2.3)**.** *Let $\mu$, $\sigma$, $f$ and $\phi$ satisfy appropriate Lipschitz and monotonicity conditions, and assume that $\mathcal{L}^i$, $i = 1,2,3$ are bounded and Lipschitz with respect to appropriate $L^2$-spaces. Then there is a constant $\ell > 0$, depending only on the Lipschitz constants of the coefficients, such that the system admits a unique square-integrable solution $(X,V)$ for $T < \ell$.*

The generality of the condition on the functionals $\mathcal{L}^i$ is an important result, as it allows to treat, within the same framework, many different types of coupled forward-backward systems not appearing in the classical literature: this is shown with the help of various examples, which also have applications to mathematical finance as well as interesting connections to parabolic integro-partial differential equations.

The next step then consists in extending the above result to any time interval. This seems however to be very difficult without knowing explicitly the functionals $\mathcal{L}^i$, and the need to treat this extension separately for each choice of $\mathcal{L}^i$ seems to be unavoidable. We thus focus on the case where the filtration is Brownian and the functionals $\mathcal{L}^i$ are given by Itô's representation: in this case, it is possible to derive the desired extension with the help of uniform Lipschitz estimates for solutions of classical FBSDEs.

**Numerical analysis of functional differential equations.** After the study of fully coupled systems, we focus our attention on decoupled systems in a Brownian setting of the form

$$\begin{cases} X_0 = x, & dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \\ V_0 = 0, & dV_t = f\big(t, X_t, \mathcal{Y}(X,V)_t, \mathcal{Z}(X,V)_t\big)dt, \end{cases}$$

and analyze a new numerical scheme for the approximation of these systems. This is based on a time discretization combined with a local Picard iteration approach, which is motivated by the previous contraction arguments.

We begin by introducing, for a partition $\pi = (t_0, \dots, t_N)$ of $[0,T]$, an implicit Euler scheme $(X^\pi_{t_i}, V^\pi_{t_i})_{i=0,\cdots,N}$. This is defined by setting $X^\pi_0 = x$, $V^\pi_0 = 0$, and for $0 \le i \le N-1$,

$$\begin{aligned} X^\pi_{t_{i+1}} &= X^\pi_{t_i} + \mu(t_i, X^\pi_{t_i})\Delta t_{i+1} + \sigma(t_i, X^\pi_{t_i})\Delta W_{t_{i+1}}, \\ V^\pi_{t_{i+1}} &= V^\pi_{t_i} + f\big(t_i, X^\pi_{t_i}, \mathcal{Y}(X^\pi, V^\pi)_{t_i}, \mathcal{Z}^\pi(X^\pi, V^\pi)_{t_i}\big)\Delta t_{i+1}, \end{aligned}$$

where $\mathcal{Z}^\pi$ is a discrete version of the operator $\mathcal{Z}$. We can then prove that the solution of this implicit scheme converges to the true solution when the mesh of the partition goes to zero, with the same rate of convergence as for classical Euler schemes (see [7, 45, 73]).

However, since the operators $\mathcal{Y}$ and $\mathcal{Z}^\pi$ depend on the terminal value $V^\pi_T$, the above Euler scheme is not explicitly solvable. To solve this problem, we approximate the solution of the implicit Euler scheme via a local Picard iteration procedure: we can then show that the rate of convergence remains unchanged, provided that the number of iterations is sufficiently large. This requires a careful local analysis of the iteration scheme, and finally leads to an implementable scheme with a satisfactory convergence rate:

**Theorem 1.3.3** (Chapter 4, Corollary 4.3.8). *Let $V^{\pi,(p_1,\cdots,p_N)}$ denote the approximation of $V^\pi$ via local iterations, and let the number of iterations $p_i$ be sufficiently large. Then, for a constant $\theta > 0$ and for $|\pi|$ sufficiently small,*

$$\max_{1 \le i \le N} \|X^\pi_{t_i} - X_{t_i}\|_2 + \|V^{\pi,(p_1,\cdots,p_N)}_{t_i} - V_{t_i}\|_2 \le \theta\sqrt{|\pi|}.$$

**Predictable projections of conformal stochastic integrals.** In the final part of the thesis, we introduce complexification techniques for stochastic processes: these techniques, although not related to the functional differential equation approach previously discussed, allow to represent real-valued processes as appropriate projections of corresponding conformal stochastic processes, and might have interesting applications in the study of stochastic differential equations. We begin by studying predictable projections on the real component of stochastic integrals with respect to a conformal Brownian motion. These projections turn out to have a particularly nice behaviour: namely, for a $d$-dimensional conformal Brownian motion $Z = X + iY$, we obtain the following result.

**Theorem 1.3.4** (Chapter 5, Theorem 5.1.2)**.** *Denote by* $(\cdot)^{\mathcal{P}^X}$ *the predictable projection with respect to* $X$*, and by* $\Pi^X$ *the orthogonal projection on* $\mathscr{H}^2(X)$*. Assume that* $H$ *is* $L^2$*-integrable with respect to* $Z$*, and that* $\left(\int H dZ\right)^{\mathcal{P}^X}$ *exists. Then, for all* $t \geq 0$*,* $\left(\int H dZ\right)_t^{\mathcal{P}^X} = \int_0^t \Pi^X(H)\, dX$ $P$*-a.s..*

In particular, this implies that the Hermite polynomials $H_\alpha(t, X_t)$, $\alpha \in \mathbb{N}^d$, can be seen as predictable projections of the corresponding powers $Z^\alpha$ of the conformal Brownian motion $Z$. In order to extend this explicit representation of $H_\alpha(t, X_t)$ to a much wider class of Brownian martingales, we thus investigate the $L^p$-convergence of series of Hermite polynomials and their connection to analytic functions. In particular, with the help of the well known hypercontractivity of the Wiener chaos, we obtain:

**Theorem 1.3.5** (Chapter 5, Theorem 5.2.3)**.** *Let* $p > 1$*, and assume that* $(M_t)_{t\in[0,T]}$ *is an* $L^p$*-martingale of the form* $M_t = g(t, X_t)$*. Then, there is a constant* $C = C(p) \in (0, 1)$ *such that, for* $S < CT$ *and* $s \in [0, S]$*, the series* $\sum_{\alpha\in\mathbb{N}^d} b_\alpha H_\alpha(s, X_s)$ *converges unconditionally in* $L^p$ *to* $M_s$*. Moreover, the analytic function* $f(z) := \sum_{\alpha\in\mathbb{N}^d} b_\alpha z^\alpha$*,* $z \in \mathbb{C}^d$*, defines an* $L^p$*-martingale* $(f(Z_s))_{s\in[0,S]}$ *such that, for all* $s \in [0, S]$*,* $\big(f(Z_s)\big)^{\mathcal{P}^X} = M_s$*.*

The condition that $p > 1$ is essential in the above result: we hence handle the special case $p = 1$ by relating the unconditional convergence in $L^1$ of an Hermite series to that in $L^p$, $p > 1$.

We conclude by applying the previous results to Widder's representation [70], which allows to express several Brownian martingales as Laplace-Stieltjes integrals with respect to some signed measure $\mu$: in particular, we characterize the existence of quadratic exponential moments of Widder's measure $\mu$ in terms of the unconditional convergence in $L^1$ of the corresponding Hermite series. In the case where $\mu$ is positive, our characterization actually leads to an equivalence between the existence of these moments and the aforementioned $L^1$-convergence.

**Organization of the thesis.** The thesis is organized as follows. Chapter 2 is devoted to functional differential equations of quadratic growth, and corresponds to the article [11]. In Chapter 3, we consider coupled systems of functional differential equations, while Chapter 4 is dedicated to the numerical analysis of decoupled systems: these two chapters are based on a joint article with Gechun Liang [13]. Finally, Chapter 5 concerns predictable projections on the real line of conformal stochastic processes and related complexification techniques: its content corresponds to the article [12].

# Chapter 2

# Backward stochastic dynamics of quadratic growth

This chapter is devoted to the application to quadratic BSDEs of the functional differential equation approach introduced by Liang, Lyons and Qian: since quadratic BSDEs are usually best treated in a $BMO$-framework, we will introduce a $BMO$-type of norm for continuous processes (for the deep connection between quadratic BSDEs and $BMO$-martingales, see for instance [3, 23, 39, 69]).

The study of quadratic BSDEs was initiated in the scalar, Brownian setting by Kobylanski [46] and Lepeltier and San Martín [48], who obtained conditions for existence and uniqueness of solutions when the terminal condition is bounded: such quadratic BSDEs became quickly of central importance in mathematical finance and stochastic control, for instance in connection to the utility maximization problem (see for instance Hu et al. [39]). Later, Kobylanski's results were extended by Briand, Hu [9, 10] and Delbaen et al. [22] to the case of convex drivers with unbounded terminal condition. Morlais [55] introduced a different extension of Kobylanski's article by studying quadratic BSDEs driven by general continuous martingales.

More recently, Bao et al. [21] investigated the possibility of further extending these results, allowing the driver to have super-quadratic growth in the $Z$-component. The answer is rather surprising: even in the Brownian setting, with convex drivers and bounded terminal conditions, the super-quadratic BSDEs are seriously ill-posed, and it is not possible to obtain neither existence nor uniqueness of solutions.

We observe however that, in opposition to the case of BSDEs with Lipschitz driver, none of the above approaches use contraction arguments to obtain existence and uniqueness: such contraction results are important, for instance, in order to derive numerical approximations. A first attempt to use contraction arguments for quadratic BSDEs was made by Tevzadze [69]. He constructed a contraction mapping when the $L^\infty$-bound on the terminal condition is sufficiently small: the result is then extended to arbitrary bounded terminal values by using tricky $BMO$-arguments. Such an extension however requires stronger conditions than Kobylanski's, in particular differentiability of the driver is needed.

Despite the stronger assumptions needed, the latter approach seems to be more suitable for a direct treatment via functional differential equations. Indeed, all the methods derived from Kobylanski's article rely on a transformation of the original quadratic BSDE: such a transformation depends on the special structure of the process $Z$, and therefore it does not seem possible to consider other operators. In the sequel, we will see how Tevzadze's techniques can be adapted to our framework in order to construct solutions for a class of quadratic backward stochastic dynamics. As for Lipschitz backward dynamics, such a functional differential approach will have some advantages. First of all, it allows us to work in a general filtration without additional difficulties. Second, the operator $Z$ in the driver can be substituted by a more general operator $\mathcal{L}(M)$, where $\mathcal{L}$ is an abstract functional: however, because we consider a $BMO$-framework, $\mathcal{L}$ has to satisfy boundedness and Lipschitz assumptions different from the ones for Lipschitz backward dynamics [49].

This chapter is organized as follows: in Section 2.1, we introduce the necessary conditions on the driver as well as appropriate solution spaces. Section 2.2 is dedicated to the solution of the functional differential equation when the terminal condition is bounded by a sufficiently small constant. Finally, the existence and uniqueness is extended in Section 2.3 to arbitrary bounded terminal conditions.

## 2.1 A class of quadratic backward stochastic dynamics

We begin by introducing an appropriate class of quadratic functional differential equations. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space, fix $T > 0$, and assume that we have a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ with the usual assumptions of right-continuity and completeness and such that all martingales with respect to $\mathbb{F}$ are continuous (observe that this implies that the corresponding optional and predictable $\sigma$-fields are identical). For $d \in \mathbb{N}$, we denote by $\mathscr{C} = \mathscr{C}([0,T], \mathbb{R}^d)$ the space of $\mathbb{R}^d$-valued adapted processes $V$ on $[0,T]$ such that $V_T \in L^1(\mathcal{F}_T)$. Then, we define the linear operators $\mathcal{M}$ and $\mathcal{Y}$, for $t \geq 0$, by

$$\begin{aligned} \mathcal{M} : L^1(\mathcal{F}_T) \times \mathscr{C} \to \mathscr{C}, \quad & \mathcal{M}(\xi, V)_t := E[\xi + V_T|\mathcal{F}_t], \\ \mathcal{Y} : L^1(\mathcal{F}_T) \times \mathscr{C} \to \mathscr{C}, \quad & \mathcal{Y}(\xi, V)_t := E[\xi + V_T|\mathcal{F}_t] - V_t. \end{aligned} \tag{2.1.1}$$

As discussed in the Introduction, if $\xi \in L^1(\mathcal{F}_T)$ and $V$ is a predictable process of finite variation in $\mathscr{C}$ such that $V_0 = 0$, then $\mathcal{Y}(\xi, V)$ is the unique special semimartingale $Y$ with drift part $-V$ and $Y_T = \xi$; its martingale part is then given (up to a $\mathcal{F}_0$-measurable random variable) by $\mathcal{M}(\xi, V)$. Moreover, we denote by $\mathcal{L}$ a (possibly non-linear) abstract functional, defined on the space of $d$-dimensional continuous martingales on $[0,T]$ and taking values, for some $p \in \mathbb{N}$, in the space of $p$-dimensional adapted processes (the codomain of $\mathcal{L}$ will be further specified later).

In this chapter, we study functional differential equations of the form

$$V_t = \int_0^t \big( f\big(s, \mathcal{Y}(\xi, V)_s, \mathcal{L}(\mathcal{M}(\xi, V))_s\big) - f(s, 0, 0)\big) ds, \tag{2.1.2}$$

where $\xi$ is $\mathcal{F}_T$-measurable, and $f : \Omega \times [0,T] \times \mathbb{R}^d \times \mathbb{R}^p \to \mathbb{R}^d$ satisfies the required measurability assumptions. Our main result is the existence and uniqueness of solutions to (2.1.2) when $\xi$ is bounded and $\mathcal{F}_T$-measurable, and $f$ satisfies the following quadratic condition:

**Assumption (A1):** *The function* $f : \Omega \times [0,T] \times \mathbb{R}^d \times \mathbb{R}^p \to \mathbb{R}^d$ *satisfies Assumption* **(A1)** *if:*

(A1.1) *For* $(y,z) \in \mathbb{R}^d \times \mathbb{R}^p$, $f(\cdot,y,z)$ *is* $\mathcal{P}$*-measurable, where* $\mathcal{P}$ *is the predictable* $\sigma$*-field with respect to* $\mathbb{F}$.

(A1.2) *For* $(\omega,t,y,z) \in \Omega \times [0,T] \times \mathbb{R}^d \times \mathbb{R}^p$, $f(\omega,t,\cdot,\cdot)$ *is twice continuously differentiable, and there are constants* $C, \theta > 0$ *such that*

$$|\nabla_y f(t,y,z)| \le C, \quad |\nabla_z f(t,y,z)| \le C + \theta|z|,$$
$$|H_{yy} f(t,y,z)| \le C^2, \quad |\nabla_z \nabla_y f(t,y,z)| \le C\theta, \quad |H_{zz} f(t,y,z)| \le \theta^2$$

*where* $H_{yy}f$, $H_{zz}f$ *denote the corresponding Hessian matrices of* $f$.

As discussed in the Introduction, the functional differential equation (2.1.2) is then closely related to the backward stochastic dynamics

$$dY_t = -\big(f(t,Y_t,\mathcal{L}(M)_t) - f(t,0,0)\big)dt + dM_t, \quad Y_T = \xi.$$

We now need to introduce some further notation. For $k$, $l \in \mathbb{N}$, we denote the Euclidean norm on $\mathbb{R}^k$ by $|\cdot|$, while on $\mathbb{R}^{k\times l}$ the notation $|\cdot|$ is used to designate the Hilbert-Schmidt norm, i.e. $|z| = \sqrt{\mathrm{Tr}(z^{\mathrm{T}}z)}$ for $z \in \mathbb{R}^{k\times l}$: we will often identify the space $\mathbb{R}^{k\times l}$ with $\mathbb{R}^{k\cdot l}$. For a probability measure $Q \approx P$, we denote by $E^Q$ the expectation with respect to $Q$: note that the functionals $\mathcal{Y}$, $\mathcal{M}$ in (2.1.1) depend on $Q$, and we will therefore write $\mathcal{Y}^Q$, $\mathcal{M}^Q$ when we want to emphasize this dependence. We can then introduce the following spaces:

- $\mathscr{L}(\mathbb{R}^p,\mathbb{R}^d)$, the space of all linear operators from $\mathbb{R}^p$ onto $\mathbb{R}^d$.
- $L^\infty = L^\infty(\Omega,\mathcal{F}_T,P;\mathbb{R}^d)$, the space of $\mathbb{R}^d$-valued, bounded and $\mathcal{F}_T$-measurable random variables with the usual ess sup-norm $\|\cdot\|_\infty$.
- $\mathscr{S}^\infty = \mathscr{S}^\infty([0,T],\mathbb{R}^d)$, the space of all continuous and adapted processes $Y : \Omega \times [0,T] \to \mathbb{R}^d$ such that

$$\|Y\|_{\mathscr{S}^\infty} := \sup_{\substack{\tau:\tau \text{ stopping}\\ \text{time } \le T}} \|Y_\tau\|_\infty = \sup_{t\in[0,T]} \|Y_t\|_\infty < \infty.$$

- $\mathscr{H}^Q = \mathscr{H}^Q([0,T],\mathbb{R}^p)$, the space of all predictable processes $H : \Omega \times [0,T] \to \mathbb{R}^p$ such that

$$\|H\|^2_{\mathscr{H}^Q} := \sup_{\substack{\tau:\tau \text{ stopping} \\ \text{time } \leq T}} \left\| E^Q\left[ \int_\tau^T |H_s|^2 ds \middle| \mathcal{F}_\tau \right] \right\|_\infty < \infty.$$

- $BMO^Q = BMO^Q([0,T],\mathbb{R}^d)$, the space of $\mathbb{R}^d$-valued $BMO$-martingales with respect to $Q$, endowed with the norm

$$\|N\|_{BMO_1^Q} := \sup_{\substack{\tau:\tau \text{ stopping} \\ \text{time } \leq T}} \left\| E^Q\left[ |N_T - N_\tau| \middle| \mathcal{F}_\tau \right] \right\|_\infty.$$

We also have to introduce an appropriate space for the solution process $V$. Motivated by the deep connection between bounded solutions of classical quadratic BSDEs and continuous $BMO$-martingales, we set

$$\|V\|_{BMO_1^Q} := \sup_{\substack{\tau:\tau \text{ stopping} \\ \text{time } \leq T}} \left\| E^Q\left[ |V_T - V_\tau| \middle| \mathcal{F}_\tau \right] \right\|_\infty,$$

$$\mathscr{S}^Q_{BMO_1} = \mathscr{S}^Q_{BMO_1}([0,T],\mathbb{R}^d) := \{V : \Omega \times [0,T] \to \mathbb{R}^d \,|\, V \text{ continuous and adapted such that } V_0 = 0 \text{ and } \|V\|_{BMO_1^Q} < \infty\}.$$

It is not difficult to check that $\left(\mathscr{S}^Q_{BMO_1}, \|\cdot\|_{BMO_1^Q}\right)$ is a Banach space. However, for $p > 1$, the norm $\|\cdot\|_{BMO_1^Q}$ is not equivalent to $\|\cdot\|_{BMO_p^Q}$ on $\mathscr{S}_{BMO_1^Q}$ (contrarily to the case of $BMO$-martingales). This can be easily proved by taking a deterministic process $V$ such that $\int_0^1 V_t dt < \infty$ and $\int_0^1 V_t^p dt = \infty$ for $p > 1$.

From now on, the dependence of all operators, spaces and norms on the measure $Q$ will be dropped when $Q = P$. By the definition of $\|\cdot\|_{BMO_1}$, we can easily verify the following result, which will be essential in the sequel and in particular shows the connection of $\mathscr{S}_{BMO_1}$ to the usual $BMO$-framework.

**Lemma 2.1.1.** *Let $\mathcal{Y}$ and $\mathcal{M}$ denote the operators in* (2.1.1)*. Then,*

$$\mathcal{Y} : L^\infty \times \mathscr{S}_{BMO_1} \to \mathscr{S}^\infty, \quad \mathcal{M} : L^\infty \times \mathscr{S}_{BMO_1} \to BMO,$$

*and we have the following estimates:*

$$\|\mathcal{Y}(\xi,V)\|_{\mathscr{S}^\infty} \le \|\xi\|_\infty + \|V\|_{BMO_1},$$
$$\|\mathcal{M}(\xi,V)\|_{BMO_1} \le 2\big(\|\xi\|_\infty + \|V\|_{BMO_1}\big), \quad \xi \in L^\infty,\ V \in \mathscr{S}_{BMO_1}.$$

*Proof.* A direct computation gives that

$$\begin{aligned}\|\mathcal{Y}(\xi,V)\|_{\mathscr{S}^\infty} &= \sup_\tau \Big\| E\big[\xi\big|\mathcal{F}_\tau\big] + E\big[V_T - V_\tau\big|\mathcal{F}_\tau\big]\Big\|_\infty \\ &\le \sup_\tau \big\|E\big[\xi\big|\mathcal{F}_\tau\big]\big\|_\infty + \sup_\tau \Big\|E\big[V_T - V_\tau\big|\mathcal{F}_\tau\big]\Big\|_\infty \\ &\le \|\xi\|_\infty + \sup_\tau \Big\|E\big[|V_T - V_\tau|\big|\mathcal{F}_\tau\big]\Big\|_\infty = \|\xi\|_\infty + \|V\|_{BMO_1},\end{aligned}$$

and similarly

$$\begin{aligned}&\|\mathcal{M}(\xi,V)\|_{BMO_1} = \sup_\tau \Big\|E\big[|\mathcal{M}(\xi,V)_T - \mathcal{M}(\xi,V)_\tau|\big|\mathcal{F}_\tau\big]\Big\|_\infty \\ &\le \sup_\tau \big\|E\big[|\xi - E[\xi|\mathcal{F}_\tau]|\big|\mathcal{F}_\tau\big]\big\|_\infty + \sup_\tau \big\|E\big[|V_T - E[V_T|\mathcal{F}_\tau]|\big|\mathcal{F}_\tau\big]\big\|_\infty \\ &\le 2\|\xi\|_\infty + \sup_\tau\Big(\big\|E\big[|V_T - V_\tau|\big|\mathcal{F}_\tau\big]\big\|_\infty + \Big\|E\Big[\big|E[V_T - V_\tau|\mathcal{F}_\tau]\big|\Big|\mathcal{F}_\tau\Big]\Big\|_\infty\Big) \\ &\le 2\big(\|\xi\|_\infty + \|V\|_{BMO_1}\big).\end{aligned}$$

□

## 2.2 Solutions for small terminal conditions

In this section, we will prove that, under the condition that the $L^\infty$-bound on the terminal value is sufficiently small, the quadratic functional differential equation (2.1.2) has a unique solution. For the moment, we will assume that the driver $f$ satisfies the following (more general) quadratic condition:

**Assumption (A2):** *The function $f : \Omega \times [0,T] \times \mathbb{R}^d \times \mathbb{R}^p \to \mathbb{R}^d$ satisfies Assumption* **(A2)** *with $(C, \theta, \alpha, \Gamma)$ if the measurability condition* (A1.1) *holds, and there are predictable processes $\alpha : \Omega \times [0,T] \to \mathbb{R}^{d\times d}$, $\Gamma : \Omega \times [0,T] \to \mathscr{L}(\mathbb{R}^p, \mathbb{R}^d)$ and constant $C, \theta > 0$ such that $|\alpha_t| \le C$ for $t \in [0,T]$, $\Gamma(\mathrm{Id}) \in \mathscr{H}$*

*and, for any* $(y,z),(y',z') \in \mathbb{R}^d \times \mathbb{R}^p$ *and* $t \in [0,T]$,

$$\begin{aligned}&|f(t,y,z) - f(t,y',z') - \alpha_t(y-y') - \Gamma_t(z-z')|\\ &\qquad \le \Big(C|y-y'| + \theta|z-z'|\Big)\Big(C\big(|y|+|y'|\big) + \theta\big(|z|+|z'|\big)\Big) \quad P\text{-a.s.}.\end{aligned}$$

**Remark 2.2.1.** As noted by Tevzadze [69], by applying the mean value theorem we can verify that, if $f$ satisfies Assumption **(A1)**, then it satisfies Assumption **(A2)** with $(C, \theta, \nabla_y f(t,0,0), \nabla_z f(t,0,0))$ (see also Lemma 2.3.1).

On the other hand, we also have to impose some conditions on the functional $\mathcal{L}$. As for Lipschitz backward dynamics, it is enough to impose some Lipschitz and boundedness assumptions with respect to appropriate norms. More exactly, we have:

**Assumption (L1'):** *The functional* $\mathcal{L}$ *satisfies Assumption* **(L1')** *if:*

(L1.1') $\mathcal{L}$ *maps the space* $BMO([0,T],\mathbb{R}^d)$ *into* $\mathscr{H}([0,T],\mathbb{R}^p)$.

(L1.2') $\mathcal{L}$ *is a bounded and Lipschitz continuous functional, i.e. there exists a constant* $K = K(T) > 0$*, which may depend on* $T$*, such that*

$$\begin{aligned}\|\mathcal{L}(M)\|_{\mathscr{H}} &\le K\|M\|_{BMO_1},\\ \|\mathcal{L}(M) - \mathcal{L}(M')\|_{\mathscr{H}} &\le K\|M - M'\|_{BMO_1},\ M, M' \in BMO([0,T],\mathbb{R}^d).\end{aligned}$$

**Examples 2.2.2.** We give below two examples of functionals $\mathcal{L}$ satisfying Assumption **(L1')**.

($i$) Assume that an $m$-dimensional Brownian motion $W$ on $(\Omega, \mathcal{F}, \mathbb{F}, P)$ is given, and define the mapping $\mathcal{L} : BMO([0,T],\mathbb{R}^d) \to \mathscr{H}([0,T],\mathbb{R}^{d\times m})$ by $\mathcal{L}(M) := Z$ for $M \in BMO([0,T],\mathbb{R}^d)$, where $Z$ is the martingale integrand in the orthogonal decomposition of $M$ w.r.t. $W$. It is easy to verify that $\mathcal{L}$ satisfies Assumption **(L1')**: namely, we have that

$$\begin{aligned}\|\mathcal{L}(M)\|_{\mathscr{H}} &= \|\mathcal{L}(M)\cdot W\|_{BMO_2}\\ &\le \|\mathcal{L}(M)\cdot W\|_{BMO_2} + \|M - \mathcal{L}(M)\cdot W\|_{BMO_2}\\ &= \|M\|_{BMO_2} \le K\|M\|_{BMO_1},\end{aligned}$$

and similarly for $\|\mathcal{L}(M) - \mathcal{L}(M')\|_{\mathscr{H}}$.

$(ii)$ Another example is given by $\mathcal{L} : BMO([0,T],\mathbb{R}) \to \mathscr{H}([0,T],\mathbb{R})$,

$$\mathcal{L}(M)_t := \sqrt{E[\langle M\rangle_{t,T}|\mathcal{F}_t]}, \quad M \in BMO([0,T],\mathbb{R}),\ t \in [0,T],$$

where for notational simplicity $\langle M\rangle_{t,T} := \langle M\rangle_T - \langle M\rangle_t$. Then, by the Kunita-Watanabe and the conditional Cauchy-Schwarz inequalities, we have that

$$\begin{aligned} E\big[\langle M, M'\rangle_{t,T}\big|\mathcal{F}_t\big] \le E\big[|\langle M, M'\rangle_{t,T}|\big|\mathcal{F}_t\big] &\le E\Big[\sqrt{\langle M\rangle_{t,T}\langle M'\rangle_{t,T}}\Big|\mathcal{F}_t\Big] \\ &\le \sqrt{E[\langle M\rangle_{t,T}|\mathcal{F}_t]}\sqrt{E[\langle M'\rangle_{t,T}|\mathcal{F}_t]}, \end{aligned}$$

and therefore, by the bilinearity of $\langle\cdot\rangle_{t,T}$,

$$\begin{aligned} \big|\mathcal{L}(M)_t - \mathcal{L}(M')_t\big|^2 &= E[\langle M\rangle_{t,T}|\mathcal{F}_t] + E[\langle M'\rangle_{t,T}|\mathcal{F}_t] \\ &\quad - 2\sqrt{E[\langle M\rangle_{t,T}|\mathcal{F}_t]}\sqrt{E[\langle M'\rangle_{t,T}|\mathcal{F}_t]} \\ &\le E[\langle M - M'\rangle_{t,T}|\mathcal{F}_t]. \end{aligned}$$

We can then compute that

$$\begin{aligned} \|\mathcal{L}(M) - \mathcal{L}(M')\|_{\mathscr{H}} &= \sup_\tau \left\| E\left[\int_\tau^T \big|\mathcal{L}(M)_s - \mathcal{L}(M')_s\big|^2 ds \middle| \mathcal{F}_\tau\right]\right\|_\infty \\ &\le \sup_\tau \left\| E\left[\int_\tau^T E[\langle M - M'\rangle_{s,T}|\mathcal{F}_s] ds \middle| \mathcal{F}_\tau\right]\right\|_\infty \\ &\le \sup_\tau \left\| E\left[\int_\tau^T \langle M - M'\rangle_{s,T} ds \middle| \mathcal{F}_\tau\right]\right\|_\infty \\ &\le T \sup_\tau \big\| E\big[\langle M - M'\rangle_{\tau,T}\big|\mathcal{F}_\tau\big]\big\|_\infty \\ &= T\|M - M'\|_{BMO_2} \le KT\|M - M'\|_{BMO_1}, \end{aligned}$$

and similarly for $\|\mathcal{L}(M)\|_{\mathscr{H}}$. Therefore, **(L1')** is satisfied.

While the case of martingale integrand processes is certainly the most important (as it leads to classical quadratic BSDEs on general filtrations), the second example shows that it is also possible to construct non-local operators satisfying **(L1')**. We are now ready to prove the main result of this section: the existence and uniqueness of solutions to (2.1.2) when $\|\xi\|_\infty$ is sufficiently small. This is essentially accomplished in two steps: first, the existence and uniqueness of solutions is obtained via a contraction argument for a particular class of drivers $f$.

**Proposition 2.2.3.** *Assume that $f$ satisfies Assumption* **(A2)** *with $(C,\theta,0,0)$ and $\mathcal{L}$ satisfies Assumption* **(L1')**. *Furthermore, define $\beta := 4TC^2 + 8K^2\theta^2$ and assume that $\|\xi\|_\infty < \frac{1}{4\beta}$. Then there is a unique solution to the quadratic functional differential equation*

$$V_t = \int_0^t \big(f(s,\mathcal{Y}(\xi,V)_s,\mathcal{L}(\mathcal{M}(\xi,V))_s) - f(s,0,0)\big)ds, \tag{2.2.1}$$

*satisfying $\|V\|_{BMO_1} \le \|\xi\|_\infty$.*

*Proof.* To simplify our notation, we will omit the dependence of $\mathcal{Y}$ and $\mathcal{M}$ on $\xi$. For $v \in \mathscr{S}_{BMO_1}$, we define $\phi(v) := V$ by

$$V_t = \int_0^t \big(f(s,\mathcal{Y}(v)_s,\mathcal{L}(\mathcal{M}(v))_s) - f(s,0,0)\big)ds,$$

First of all, we check that $\phi$ maps $\mathscr{S}_{BMO_1}$ into itself. Clearly, $V = \phi(v)$ is a continuous and adapted process, and by Assumption (L1.2') we have that

$$\begin{aligned}
&\|V\|_{BMO_1} = \sup_\tau \left\| E\left[ \left| \int_\tau^T \big(f(s,\mathcal{Y}(v)_s,\mathcal{L}(\mathcal{M}(v))_s) - f(s,0,0)\big)ds \right| \middle| \mathcal{F}_\tau \right] \right\|_\infty \\
&\le \sup_\tau \left\| E\left[ \int_\tau^T \big|f(s,\mathcal{Y}(v)_s,\mathcal{L}(\mathcal{M}(v))_s) - f(s,0,0)\big|ds \middle| \mathcal{F}_\tau \right] \right\|_\infty \\
&\le \sup_\tau \left\| E\left[ \int_\tau^T 2\big(C^2|\mathcal{Y}(v)_s|^2 + \theta^2|\mathcal{L}(\mathcal{M}(v))_s|^2\big)ds \middle| \mathcal{F}_\tau \right] \right\|_\infty \\
&\le 2\left( C^2 \sup_\tau \left\| E\left[ \int_\tau^T |\mathcal{Y}(v)_s|^2 ds \middle| \mathcal{F}_\tau \right] \right\|_\infty + \theta^2 \sup_\tau \left\| E\left[ \int_\tau^T |\mathcal{L}(\mathcal{M}(v))_s|^2 ds \middle| \mathcal{F}_\tau \right] \right\|_\infty \right) \\
&\le 2\big(TC^2\|\mathcal{Y}(v)\|^2_{\mathscr{S}^\infty} + K^2\theta^2\|\mathcal{M}(v)\|^2_{BMO_1}\big) < \infty.
\end{aligned}$$

Therefore, the mapping $\phi : \mathscr{S}_{BMO_1} \to \mathscr{S}_{BMO_1}$ is well defined. In order to apply the fixed point theorem to $\phi$ we first need to prove that, for sufficiently small balls $B$ in $\mathscr{S}_{BMO_1}$, $\phi$ maps $B$ into itself.

By the previous computations and Lemma 2.1.1, we have that

$$\begin{aligned}\|V\|_{BMO_1} &\leq 2\big(TC^2\|\mathcal{Y}(v)\|^2_{\mathscr{S}^\infty} + K^2\theta^2\|\mathcal{M}(v)\|^2_{BMO_1}\big)\\ &\leq \beta\big(\|\xi\|^2_\infty + \|v\|^2_{BMO_1}\big).\end{aligned}$$

Consider now the quadratic equation

$$\beta x^2 - x + \beta\|\xi\|^2_\infty \leq 0. \tag{2.2.2}$$

Since $\beta > 0$, there are positive solutions to this inequality if and only if its discriminant is non-negative: this happens if and only if $\|\xi\|_\infty \leq \frac{1}{2\beta}$, and in this case $x = \|\xi\|_\infty$ solves (2.2.2).

By setting $R := \|\xi\|_\infty$ and $B_R := \{V \in \mathscr{S}_{BMO_1} \mid \|V\|_{BMO_1} \leq R\}$, we then have by the above computations that, for $v \in B_R$,

$$\|V\|_{BMO_1} \leq \beta\big(\|\xi\|^2_\infty + \|v\|^2_{BMO_1}\big) \leq 2\beta\|\xi\|^2_\infty \leq \|\xi\|_\infty = R$$

whenever $\|\xi\|_\infty \leq \frac{1}{2\beta}$, and therefore $\phi$ maps $B_R$ into itself.

The next step consists in proving that, due to the particular choice of $R$, $\phi$ is a contraction on $B_R$ whenever $\|\xi\|_\infty < \frac{1}{4\beta}$. Let $v^1$, $v^2 \in B_R$, $V^i := \phi(v^i)$ for $i = 1, 2$, and set $\delta F := F^1 - F^2$ for processes $F^1$, $F^2$. Since $f$ satisfies Assumption **(A2)**, by applying the classical and conditional Cauchy-Schwarz inequalities, we can easily verify that

$$\begin{aligned}\|\delta V\|^2_{BMO_1} &\leq \sup_\tau \Bigg\| E\bigg[\int_\tau^T \big|f\big(s, \mathcal{Y}(v^1)_s, \mathcal{L}(\mathcal{M}(v^1))_s\big)\\ &\qquad\qquad - f\big(s, \mathcal{Y}(v^2)_s, \mathcal{L}(\mathcal{M}(v^2))_s\big)\big| ds \bigg| \mathcal{F}_\tau\bigg]^2 \Bigg\|_\infty\\ &\leq \sup_\tau \Bigg(\bigg\| E\bigg[\int_\tau^T \big(C|\delta\mathcal{Y}(v)_s| + \theta|\delta\mathcal{L}(\mathcal{M}(v))_s|\big)^2 ds \bigg| \mathcal{F}_\tau\bigg]\bigg\|_\infty \times\\ &\bigg\| E\bigg[\int_\tau^T \Big[C\big(|\mathcal{Y}(v^1)_s| + |\mathcal{Y}(v^2)_s|\big) + \theta\big(|\mathcal{L}(\mathcal{M}(v^1))_s| + |\mathcal{L}(\mathcal{M}(v^2))_s|\big)\Big]^2 ds \bigg| \mathcal{F}_\tau\bigg]\bigg\|_\infty,\end{aligned}$$

and therefore

$$\begin{aligned}
\|\delta V\|^2_{BMO_1} &\le 2\big(TC^2\|\delta\mathcal{Y}(v)\|^2_{\mathscr{S}^\infty} + K^2\theta^2\|\delta\mathcal{M}(v)\|^2_{BMO_1}\big)\times \\
&\qquad 4\sup_\tau\Bigg(C^2\bigg\|E\bigg[\int_\tau^T \big(|\mathcal{Y}(v^1)_s|^2 + |\mathcal{Y}(v^2)_s|^2\big)ds\bigg|\mathcal{F}_\tau\bigg]\bigg\|_\infty \\
&\qquad\quad + \theta^2\bigg\|E\bigg[\int_\tau^T \big(|\mathcal{L}(\mathcal{M}(v^1))_s|^2 + |\mathcal{L}(\mathcal{M}(v^2))_s|^2\big)ds\bigg|\mathcal{F}_\tau\bigg]\bigg\|_\infty\Bigg) \\
&\le 8\big(TC^2\|\delta\mathcal{Y}(v)\|^2_{\mathscr{S}^\infty} + K^2\theta^2\|\delta\mathcal{M}(v)\|^2_{BMO_1}\big)\times \qquad (2.2.3)\\
&\quad \Big(TC^2\big(\|\mathcal{Y}(v^1)\|^2_{\mathscr{S}^\infty} + \|\mathcal{Y}(v^2)\|^2_{\mathscr{S}^\infty}\big) + K^2\theta^2\big(\|\mathcal{M}(v^1)\|^2_{BMO_1} + \|\mathcal{M}(v^2)\|^2_{BMO_1}\big)\Big).
\end{aligned}$$

Moreover, since $v^1$, $v^2 \in B_R$ we have that, for $i = 1, 2$,

$$\|\mathcal{Y}(v^i)\|^2_{\mathscr{S}^\infty} \le 2\big(\|\xi\|^2_\infty + \|v^i\|^2_{BMO_1}\big) \le 4\|\xi\|^2_\infty,$$
$$\|\mathcal{M}(v^i)\|^2_{BMO_1} \le 8\big(\|\xi\|^2_\infty + \|v^i\|^2_{BMO_1}\big) \le 16\|\xi\|^2_\infty.$$

Therefore, by the inequality (2.2.3), we have that

$$\begin{aligned}
\|\delta V\|^2_{BMO_1} &\le 8\big(TC^2\|\delta\mathcal{Y}(v)\|^2_{\mathscr{S}^\infty} + K^2\theta^2\|\delta\mathcal{M}(v)\|^2_{BMO_1}\big)(8TC^2 + 32K^2\theta^2)\|\xi\|^2_\infty \\
&\le 64(TC^2 + 4K^2\theta^2)^2\|\xi\|^2_\infty\|\delta v\|^2_{BMO_1} \\
&\le 16\beta^2\|\xi\|^2_\infty\|\delta v\|^2_{BMO_1},
\end{aligned}$$

and hence $\phi$ is a contraction on $B_R$ if $\|\xi\|_\infty < \frac{1}{4\beta}$. The proof is completed by applying the fixed point theorem. □

Observe that the above result is valid for any dimension $d \ge 1$, and that it does not depend on the probability measure $P$ and thus holds on any probability space $(\Omega, \mathcal{F}, Q)$ with $Q \approx P$.

The second step consists in extending Proposition 2.2.3 to general drivers $f$ via an appropriate transformation. Unfortunately, our transformation argument requires that we introduce two major restrictions. *For the rest of the chapter, we assume that $d = 1$, and that the functional $\mathcal{L}$ is explicitly given by the orthogonal decomposition with respect to $W$, as in Example 2.2.2 $(i)$.* This explicit choice of $\mathcal{L}$ has two consequences: first, $\mathcal{L}$ is linear, and second, for $\varepsilon : \Omega \times [0,T] \to \mathbb{R}$ predictable, we have the transformation property

$$\varepsilon_t \mathcal{L}(M)_t = \mathcal{L}\Big(\int_0^\cdot \varepsilon_s dM_s\Big)_t \tag{2.2.4}$$

for any $t \geq 0$ and $M \in BMO$: namely, we have that

$$d\Big(\int_0^\cdot \varepsilon_s dM_s\Big)_t = \varepsilon_t dM_t = \varepsilon_t \mathcal{L}(M)_t dW_t + \underbrace{\varepsilon_t\big(dM_t - \mathcal{L}(M)_t dW_t\big)}_{\text{orthogonal w.r.t. } W}$$

and the claim follows by uniqueness of orthogonal decompositions.

For any $Q \approx P$, the functional $\mathcal{L}^Q : BMO^Q \to \mathscr{H}^Q$ can be defined similarly to $\mathcal{L}$ by taking the orthogonal decomposition with respect to $W^Q$, the Girsanov's transformation of $W$. Via measure changes, we can then derive the following transformation result:

**Lemma 2.2.4.** *Let $d = 1$, and assume that $f$ satisfies Assumption* **(A2)** *with $(C, \theta, \alpha, \Gamma)$. Define $\hat{f}$ by*

$$\hat{f}(t, \hat{y}, \hat{z}) := e_t\big(f(t, e_t^{-1}\hat{y}, e_t^{-1}\hat{z}) - f(t, 0, 0)\big) - \alpha_t \hat{y} - \Gamma_t \hat{z}$$

*for $(t, \hat{y}, \hat{z}) \in [0,T] \times \mathbb{R} \times \mathbb{R}^p$, where $e_t := \exp\big(\int_0^t \alpha_s ds\big)$ for $t \in [0,T]$. Then, $\hat{f}$ satisfies Assumption* **(A2)** *with $(\widehat{C}, \hat{\theta}, 0, 0) := (Ce^{TC/2}, \theta e^{TC/2}, 0, 0)$.*

*Moreover, the functional differential equation* (2.2.1) *has a solution if and only if there is a solution $\widehat{V}$ to*

$$\widehat{V}_t = \int_0^t \hat{f}(s, \mathcal{Y}^Q(\hat{\xi}, \widehat{V})_s, \mathcal{L}^Q(\mathcal{M}^Q(\hat{\xi}, \widehat{V}))_s) ds, \tag{2.2.5}$$

*where $\hat{\xi} := e_T \xi$, and $Q \approx P$ is given by $\frac{dQ}{dP} = \mathcal{E}(\Gamma \cdot W)_T$.*

*Proof.* We first prove that $\hat{f}$ satisfies **(A2)** with $(\widehat{C}, \hat{\theta}, 0, 0)$ via a direct computation: let $\hat{y}, \hat{y}' \in \mathbb{R}$, $\hat{z}, \hat{z}' \in \mathbb{R}^p$ and $\delta\hat{y} = \hat{y} - \hat{y}'$, $\delta\hat{z} = \hat{z} - \hat{z}'$. Then,

$$\begin{aligned}|\hat{f}(t,\hat{y},\hat{z}) - \hat{f}(t,\hat{y}',\hat{z}')| &= \big|e_t\big(f(t, e_t^{-1}\hat{y}, e_t^{-1}\hat{z}) - f(t, e_t^{-1}\hat{y}', e_t^{-1}\hat{z}')\big) - \alpha_t\delta\hat{y} - \Gamma_t\delta\hat{z}\big| \\ &\leq e_t\big|f(t, e_t^{-1}\hat{y}, e_t^{-1}\hat{z}) - f(t, e_t^{-1}\hat{y}', e_t^{-1}\hat{z}') - e_t^{-1}\alpha_t\delta\hat{y} - e_t^{-1}\Gamma_t\delta\hat{z}\big| \\ &\leq e_t\big(Ce_t^{-1}|\delta\hat{y}| + \theta e_t^{-1}|\delta\hat{z}|\big)\big(Ce_t^{-1}(|\hat{y}| + |\hat{y}'|) + \theta e_t^{-1}(|\hat{z}| + |\hat{z}'|)\big)\end{aligned}$$

$$
\begin{aligned}
&= e_t^{-1}\big(C|\delta\hat{y}| + \theta|\delta\hat{z}|\big)\big(C(|\hat{y}| + |\hat{y}'|) + \theta(|\hat{z}| + |\hat{z}'|)\big) \\
&\le \big(\widehat{C}|\delta\hat{y}| + \hat{\theta}|\delta\hat{z}|\big)\big(\widehat{C}(|\hat{y}| + |\hat{y}'|) + \hat{\theta}(|\hat{z}| + |\hat{z}'|)\big),
\end{aligned}
$$

where the last inequality follows from $e_t^{-1} = \exp\big(-\int_0^t \alpha_s ds\big) \le e^{TC}$. Assume now that we are given a solution $V$ of (2.2.1), and set

$$
\widehat{V}_t := e_t V_t - \int_0^t \big(e_s \alpha_s \mathcal{M}^P(\xi, V)_s + e_s \Gamma_s^T \mathcal{L}^P(\mathcal{M}^P(\xi, V))_s\big) ds, \quad t \ge 0.
$$

We claim that $\widehat{V}$ defines a solution of (2.2.5). Indeed, it is easy to verify that $-\widehat{V}$ is the predictable process of finite variation appearing in the canonical decomposition with respect to $Q$ of the special $Q$-semimartingale $e_t \mathcal{Y}^P(\xi, V)_t$, since

$$
\begin{aligned}
e_t \mathcal{Y}^P(\xi, V)_t &= e_t\big(\mathcal{M}^P(\xi, V)_t - V_t\big) \\
&= \int_0^t e_s d\mathcal{M}^P(\xi, V)_s + \int_0^t e_s \alpha_s \mathcal{M}^P(\xi, V)_s ds - e_t V_t \\
&= \int_0^t e_s d\big(\mathcal{M}^P(\xi, V) - \langle \mathcal{M}^P(\xi, V), \Gamma \cdot W\rangle\big)_s - \widehat{V}_t,
\end{aligned}
$$

and because $\int_0^\cdot e_s d\big(\mathcal{M}^P(\xi, V) - \langle \mathcal{M}^P(\xi, V), \Gamma \cdot W\rangle\big)_s$ is a $Q$-martingale by Girsanov's theorem. Therefore, since $e_T \mathcal{Y}^P(\xi, V)_T = \hat{\xi}$, we obtain that

$$
\mathcal{Y}^Q(\hat{\xi}, \widehat{V}) = e\mathcal{Y}^P(\xi, V), \quad \mathcal{M}^Q(\hat{\xi}, \widehat{V}) = \int_0^\cdot e_s d\big(\mathcal{M}^P(\xi, V) - \langle \mathcal{M}^P(\xi, V), \Gamma \cdot W\rangle\big)_s.
$$

On the other hand, by property (2.2.4) we have that

$$
\begin{aligned}
e\mathcal{L}^P(\mathcal{M}^P(\xi, V)) &= \mathcal{L}^P\Big(\int_0^\cdot e_s d\mathcal{M}^P(\xi, V)_s\Big) \\
&= \mathcal{L}^P\Big(\mathcal{M}^Q(\hat{\xi}, \widehat{V}) + \int_0^\cdot e_s d\langle \mathcal{M}^P(\xi, V), \Gamma \cdot W\rangle_s\Big) \\
&= \mathcal{L}^Q\big(\mathcal{M}^Q(\hat{\xi}, \widehat{V})\big),
\end{aligned}
$$

where the last equality follows by the definition of $\mathcal{L}^P$, $\mathcal{L}^Q$. This finally leads us to

$$\begin{aligned} d\widehat{V}_t &= e_t dV_t + e_t\alpha_t V_t dt - e_t\alpha_t \mathcal{M}^P(\xi,V)_t dt - e_t\Gamma_t \mathcal{L}^P(\mathcal{M}^P(\xi,V))_t dt \\ &= \Big[ e_t \Big( f\big(t, \mathcal{Y}^P(\xi,V)_t, \mathcal{L}^P(\mathcal{M}^P(\xi,V))_t\big) - f(t,0,0)\Big) \\ &\qquad - e_t\alpha_t \mathcal{Y}^P(\xi,V)_t - e_t\Gamma_t\mathcal{L}^P(\mathcal{M}^P(\xi,V))_t\Big] dt \\ &= \hat{f}\big(t, e_t\mathcal{Y}^P(\xi,V)_t, e_t\mathcal{L}^P(\mathcal{M}^P(\xi,V))_t\big) dt \\ &= \hat{f}\big(t, \mathcal{Y}^Q(\hat{\xi},\widehat{V})_t, \mathcal{L}^Q(\mathcal{M}^Q(\hat{\xi},\widehat{V}))_t\big) dt. \end{aligned}$$

□

As a consequence of this transformation, we can relax the conditions on $f$ in Proposition 2.2.3, obtaining the following result:

**Lemma 2.2.5.** *Let $d = 1$, and let $f$ satisfy Assumption* **(A2)** *with $(C, \theta, \alpha, \Gamma)$. Denote by $\beta$ the same constant as in Proposition 2.2.3, and assume that $\|\xi\|_\infty \leq \frac{e^{-2TC}}{4\beta}$. Then, the functional differential equation* (2.2.1) *has a unique solution $V \in \mathscr{S}_{BMO_1}$.*

*Proof.* For $e_t := \exp\big(\int_0^t \alpha_s ds\big)$, set $\hat{\xi} := e_T\xi$, and $\hat{f}$ as in Lemma 2.2.4. Then, Lemma 2.2.4 implies that $\hat{f}$ satisfies Assumption **(A2)** with $(\widehat{C}, \hat{\theta}, 0, 0)$, and we have that

$$\|\hat{\xi}\|_\infty \leq e^{TC}\|\xi\|_\infty \leq \frac{e^{-TC}}{4\beta} = \frac{1}{4\hat{\beta}},$$

where $\hat{\beta} := 4T\widehat{C}^2 + 8K^2\hat{\theta}^2$. Hence, we can apply Proposition 2.2.3, obtaining that the functional differential equation with generators $(\hat{f}, \hat{\xi})$ has a unique solution on $(\Omega, \mathcal{F}, Q)$. By Lemma 2.2.4, we obtain the claim. □

## 2.3 Extension to arbitrary bounded terminal conditions

In this section we extend the existence and uniqueness result to any bounded terminal condition. The main idea is to rewrite $\xi$ as $\xi = \sum_{i=1}^n \xi^i$ for some $n \in \mathbb{N}$, where $\|\xi^i\|_\infty$ is sufficiently small, and to construct for each $i$ a functional differential equation with terminal condition $\xi^i$ and on which Lemma 2.2.5 is applicable. The role of the Assumption **(A1)** is to guarantee that these functional differential equations all satisfy Assumption **(A2)** with the same constants $C$, $\theta$.

The first step is to apply the result of Lemma 2.2.5 to a transformed functional differential equation. To obtain this, we first need to derive the following technical result:

**Lemma 2.3.1.** *Let $d = 1$, and assume that $f$ satisfies Assumption* **(A1)**. *For $\tilde{\xi} \in L^\infty$ such that $\|\tilde{\xi}\|_\infty \leq \frac{e^{-2TC}}{4\beta}$, let $\widetilde{V}$ be the solution of*

$$\widetilde{V}_t = \int_0^t \big(f(s, \mathcal{Y}(\tilde{\xi}, \widetilde{V})_s, \mathcal{L}(\mathcal{M}(\tilde{\xi}, \widetilde{V}))_s) - f(s, 0, 0)\big) ds. \tag{2.3.1}$$

*Moreover, for $(t, y, z) \in [0, T] \times \mathbb{R} \times \mathbb{R}^p$ define $\breve{f}$ by*

$$\breve{f}(t, y, z) := f\big(t, y + \mathcal{Y}(\tilde{\xi}, \widetilde{V})_t, z + \mathcal{L}(\mathcal{M}(\tilde{\xi}, \widetilde{V}))_t\big) - f\big(t, \mathcal{Y}(\tilde{\xi}, \widetilde{V})_t, \mathcal{L}(\mathcal{M}(\tilde{\xi}, \widetilde{V}))_t\big).$$

*Then. $\breve{f}$ satisfies Assumption* **(A2)** *with $(C, \theta, \alpha, \Gamma)$, where $\alpha$, $\Gamma$ are given by $\alpha_t = \nabla_y f\big(t, \mathcal{Y}(\tilde{\xi}, \widetilde{V})_t, \mathcal{L}(\mathcal{M}(\tilde{\xi}, \widetilde{V}))_t\big)$, $\Gamma_t = \nabla_z f\big(t, \mathcal{Y}(\tilde{\xi}, \widetilde{V})_t, \mathcal{L}(\mathcal{M}(\tilde{\xi}, \widetilde{V}))_t\big)$.*

*Proof.* This is essentially proved by applying the mean value theorem twice as in [69]: for notational simplicity, we write $\widetilde{Y} = \mathcal{Y}(\tilde{\xi}, \widetilde{V})$ and $\widetilde{Z} = \mathcal{L}(\mathcal{M}(\tilde{\xi}, \widetilde{V}))$. For $y, y' \in \mathbb{R}$, $z, z' \in \mathbb{R}^p$, we set $\delta y = y - y'$, $\delta z = z - z'$. Then,

$$\begin{aligned}
&\big|\breve{f}(t, y, z) - \breve{f}(t, y', z') - \nabla_y f\big(t, \widetilde{Y}_t, \widetilde{Z}_t\big)\delta y - \nabla_z f\big(t, \widetilde{Y}_t, \widetilde{Z}_t\big)\delta z\big| \\
&= \big|f\big(t, y + \widetilde{Y}_t, z + \widetilde{Z}_t\big) - f\big(t, y' + \widetilde{Y}_t, z' + \widetilde{Z}_t\big) \\
&\quad - \nabla_y f\big(t, \widetilde{Y}_t, \widetilde{Z}_t\big)\delta y - \nabla_z f\big(t, \widetilde{Y}_t, \widetilde{Z}_t\big)\delta z\big| \\
&\leq \big|\nabla_y f\big(\Lambda\big)\delta y + \nabla_z f\big(\Lambda\big)\delta z - \nabla_y f\big(t, \widetilde{Y}_t, \widetilde{Z}_t\big)\delta y - \nabla_z f\big(t, \widetilde{Y}_t, \widetilde{Z}_t\big)\delta z\big|,
\end{aligned}$$

where, for some $\lambda \in [0, 1]$,

$$\begin{aligned}
\Lambda &= \Big(t, \lambda(y + \widetilde{Y}_t) + (1 - \lambda)(y' + \widetilde{Y}_t), \lambda(z + \widetilde{Z}_t) + (1 - \lambda)(z' + \widetilde{Z}_t)\Big) \\
&= \Big(t, \lambda y + (1 - \lambda)y' + \widetilde{Y}_t, \lambda z + (1 - \lambda)z' + \widetilde{Z}_t\Big).
\end{aligned}$$

We then apply the mean value theorem to the functions $\nabla_y f$ and $\nabla_z f$. Since

$f$ satisfies Assumption **(A1)**, we conclude that

$$\begin{aligned}
&\big|\breve{f}(t,y,z)-\breve{f}(t,y',z')-\nabla_y f\big(t,\widetilde{Y}_t,\widetilde{Z}_t\big)\delta y-\nabla_z f\big(t,\widetilde{Y}_t,\widetilde{Z}_t\big)\delta z\big|\\
&\le \big|\nabla_y f\big(\Lambda\big)-\nabla_y f\big(t,\widetilde{Y}_t,\widetilde{Z}_t\big)\big||\delta y|+\big|\nabla_z f\big(\Lambda\big)-\nabla_z f\big(t,\widetilde{Y}_t,\widetilde{Z}_t\big)\big||\delta z|\\
&\le C^2\Big|\big(\lambda y+(1-\lambda)y'+\widetilde{Y}_t\big)-\widetilde{Y}_t\Big||\delta y|+C\theta\Big|\big(\lambda z+(1-\lambda)z'+\widetilde{Z}_t\big)-\widetilde{Z}_t\Big||\delta y|\\
&\quad+C\theta\Big|\big(\lambda y+(1-\lambda)y'+\widetilde{Y}_t\big)-\widetilde{Y}_t\Big||\delta z|+\theta^2\Big|\big(\lambda z+(1-\lambda)z'+\widetilde{Z}_t\big)-\widetilde{Z}_t\Big||\delta z|\\
&\le \Big(C|y-y'|+\theta|z-z'|\Big)\Big(C\big(|y|+|y'|\big)+\theta\big(|z|+|z'|\big)\Big),
\end{aligned}$$

as $\big|\lambda y+(1-\lambda)y'\big|\le|y|+|y'|$, $\big|\lambda z+(1-\lambda)z'\big|\le|z|+|z'|$. □

The above assertion allows us to apply Lemma 2.2.5 to the functional differential equation with driver $\breve{f}$. Thus:

**Corollary 2.3.2.** *Assume that the conditions of Lemma 2.3.1 hold true, and let $\breve{\xi}\in L^\infty$ such that $\|\breve{\xi}\|_\infty\le\frac{e^{-2TC}}{4\beta}$. Then, the functional differential equation*

$$\breve{V}_t=\int_0^t\breve{f}\big(s,\mathcal{Y}(\breve{\xi},\breve{V})_s,\mathcal{L}(\mathcal{M}(\breve{\xi},\breve{V}))_s\big)ds \tag{2.3.2}$$

*has a unique solution in $\mathscr{S}_{BMO_1}$.*

Corollary 2.3.2 is an important tool in deriving the desired extension to arbitrary terminal conditions. Namely, it allows us to combine the solutions of (2.3.1) and (2.3.2), obtaining a process which solves a functional differential equation whose terminal condition is the sum of the previous ones:

**Lemma 2.3.3.** *For $d=1$, let $f$ satisfy Assumption* **(A1)**, *and assume that $\|\tilde{\xi}\|_\infty,\|\breve{\xi}\|_\infty\le\frac{e^{-2TC}}{4\beta}$. Let $\widetilde{V}$, $\breve{V}$ denote the solutions to* (2.3.1), *respectively* (2.3.2). *Then, $V=\widetilde{V}+\breve{V}$ is a solution of*

$$V_t=\int_0^t\big(f(s,\mathcal{Y}(\xi,V)_s,\mathcal{L}(\mathcal{M}(\xi,V))_s)-f(s,0,0)\big)ds,\quad \textit{where } \xi=\tilde{\xi}+\breve{\xi}.$$

*Proof.* This follows by simply summing the equations (2.3.1) and (2.3.2), remembering that the operators $\mathcal{Y}$, $\mathcal{M}$ and the functional $\mathcal{L}$ are linear. □

By iterating this argument, we can finally construct a solution to our quadratic functional differential equation for any terminal condition, leading us to the following result:

**Theorem 2.3.4.** *Let $d = 1$, $\xi \in L^\infty$ , and assume that $f$ satisfies Assumption* **(A1)**. *Then, the functional differential equation*

$$dV_t = \big(f\big(t, \mathcal{Y}(\xi, V)_t, \mathcal{L}(\mathcal{M}(\xi, V))_t\big) - f(t, 0, 0)\big)dt \tag{2.3.3}$$

*has a unique solution in $\mathscr{S}_{BMO_1}$.*

*Proof.* We first prove the existence. First of all, we rewrite $\xi$ as $\xi = \sum_{i=1}^n \xi^i$ for some $n \in \mathbb{N}$, where $\|\xi^i\|_\infty \leq \frac{e^{-2TC}}{4\beta}$, $i = 1, \cdots, n$. For $i = 1, \cdots, n$, we can thus construct $V^i$ as the solution of the functional differential equation

$$V_t^i = \int_0^t f^i\big(s, \mathcal{Y}(\xi^i, V^i)_s, \mathcal{L}(\mathcal{M}(\xi^i, V^i))_s\big)ds,$$

where the drivers $f^i$ are defined recursively by $f^1(t, x, z) = f(t, y, z) - f(t, 0, 0)$,

$$\begin{aligned} f^i(t, x, z) = f\Big(t, y + \mathcal{Y}\big(\textstyle\sum_{j=1}^{i-1} \xi^i, \sum_{j=1}^{i-1} V^j\big)_t, z + \mathcal{L}\big(\mathcal{M}\big(\sum_{j=1}^{i-1} \xi^i, \sum_{j=1}^{i-1} V^j\big)\big)_t\Big) \\ - f\Big(t, \mathcal{Y}\big(\textstyle\sum_{j=1}^{i-1} \xi^i, \sum_{j=1}^{i-1} V^j\big)_t, \mathcal{L}\big(\mathcal{M}\big(\sum_{j=1}^{i-1} \xi^i, \sum_{j=1}^{i-1} V^j\big)\big)_t\Big), \end{aligned}$$

for $(t, y, z) \in [0, T] \times \mathbb{R} \times \mathbb{R}^p$. Note that all solutions $V^i$ exist and are unique by Corollary 2.3.2. By applying Lemma 2.3.3, we get that $V := \sum_{i=1}^n V^i$ is a solution of

$$V_t = \int_0^t \big(f\big(s, \mathcal{Y}(\xi, V)_s, \mathcal{L}(\mathcal{M}(\xi, V))_s\big) - f(s, 0, 0)\big)ds.$$

The uniqueness is shown in two steps: we first obtain the uniqueness of $\mathcal{Y}(\xi, V)$ by applying the comparison theorem for quadratic BSDEs proved by Morlais [55] (observe that Morlais' result can be applied since $d = 1$ and because of the choice of $\mathcal{L}$). Once we have the uniqueness of $\mathcal{Y}(\xi, V)$, the claim follows by the uniqueness (up to $\mathcal{F}_0$-measurable random variables) of the canonical decomposition for special semimartingales. □

We end this chapter by observing that Theorem 2.3.4 also implies the existence of a unique solution to the backward stochastic dynamics associated to the functional differential equation (2.3.3):

**Corollary 2.3.5.** *Under the conditions of Theorem 2.3.4, the quadratic backward stochastic dynamics*

$$dY_t = -\big(f(t, Y_t, \mathcal{L}(M)_t) - f(t, 0, 0)\big)dt + dM_t, \quad Y_T = \xi,$$

*have a unique solution such that* $(Y, M) \in \mathscr{S}^\infty \times BMO$ *and* $Y_0 = M_0$.

# Chapter 3

# Fully coupled forward-backward dynamics

In this chapter, we introduce a broad class of fully coupled forward-backward stochastic dynamics on a general filtered probability space, which contains classical FBSDEs as a special case. As in the previous chapters, these forward-backward dynamics are then associated to a system of forward functional differential equations, by introducing appropriate functionals $\mathcal{Y}$, $\mathcal{M}$ and $\mathcal{L}^i$. This way, the problem becomes more homogeneous: both functional differential equations share a common structure, and the conflicting nature between the forward and backward components is thus partly avoided.

When the time interval is sufficiently small, we obtain the main result of this chapter, the existence of a unique solution to the system of functional differential equations under weak assumptions on the functionals $\mathcal{L}^i$ and under Lipschitz and monotonicity conditions on the coefficients. We observe that our approach is purely probabilistic, and that the generality of $\mathcal{L}^i$ allows to locally treat many other types of forward-backward equations that do not fit in the classical FBSDE framework. The extension to arbitrarily time interval is however more problematic, and has to be treated separately for each choice of $\mathcal{L}^i$: we present a study of the case where the filtration is Brownian and the functionals $\mathcal{L}^i$ are given by Itô's representation.

Fully coupled FBSDEs have been studied extensively during the last two decades, and have found many applications, especially in stochastic control theory and mathematical finance. For an overview of the literature, we refer the reader to the book of Ma and Yong [51]. It is well known that, contrarily to the case of decoupled FBSDEs, the standard Lipschitz conditions are not sufficient to obtain the solution for arbitrarily large time intervals. This led to the development of mainly three approaches, each having its constraints and which do not cover each other.

The first is the method of contraction mapping (our approach falls within this category). This methodology has been initiated by Antonelli [1] and later developed by Pardoux and Tang [60]. It works very well when the time horizon is sufficiently small (or alternatively, under some particular monotonicity conditions [60]): the drawback is that it requires particular attention when extending the solution to arbitrarily large time intervals, in order to avoid possible explosions. This method seems to be the more convenient in our case, since it allows us to leave the functionals $\mathcal{L}^i$ unspecified by only imposing some Lipschitz and boundedness conditions.

The other two methods allow, on the other hand, to directly obtain the solution for arbitrarily large time horizons. The four-step scheme has been introduced by Ma et al. [50], by relying on the connection between parabolic PDEs and FBSDEs. Unfortunately, in order to solve the PDE, the coefficients have to be deterministic and have to satisfy strong regularity assumptions. The last approach is the method of continuation, initiated by Hu and Peng [40], and later considered by Peng and Wu [61] and Yong [71], who introduced the concept of bridge. While the coefficients are allowed to be random, this method requires that they satisfy a set of monotonicity conditions.

The chapter is organized as follows. In Section 3.1, after introducing an appropriate framework, we define our class of forward-backward dynamics and the associated system of functional differential equations. Section 3.2 is then dedicated to the existence and uniqueness of solutions to such a system for sufficiently small time horizons. Finally, in Section 3.3 we discuss the extension of the solution to arbitrarily large time intervals, especially in the case of classical Brownian FBSDEs.

## 3.1 Fully coupled systems of functional differential equations

We begin our discussion with some intuitive argumentations in the Brownian setting, which should help us understand how to possibly formulate a functional differential approach to general fully coupled forward-backward equations. Let $(\Omega, \mathcal{F}, P)$ be for the moment a complete probability space with an $m$-dimensional Brownian motion $W = (W_t)_{t\in[0,T]}$ and the corresponding augmented filtration $(\mathcal{F}_t)_{t\in[0,T]}$. Consider a classical fully coupled FBSDE of the form

$$\begin{cases} dX_t = \mu(t, X_t, Y_t, Z_t)dt + \sigma(t, X_t, Y_t, Z_t)dW_t, \\ dY_t = -f(t, X_t, Y_t, Z_t)dt + Z_t dW_t, \\ X_0 = x, \quad Y_T = \phi(X_T), \end{cases} \tag{3.1.1}$$

where the functions $\mu : \Omega \times [0,T] \times \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^{d\times m} \to \mathbb{R}^n$, $\sigma : \Omega \times [0,T] \times \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^{d\times m} \to \mathbb{R}^{n\times m}$, $f : \Omega \times [0,T] \times \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^{d\times m} \to \mathbb{R}^d$, $\phi : \Omega \times \mathbb{R}^n \to \mathbb{R}^d$ satisfy the usual measurability and integrability conditions.

We denote by $\mathscr{C}_X^\phi$ the class of $\mathbb{R}^n$-valued adapted processes $X$ on $[0,T]$ such that $\phi(X_T) \in L^1(\mathcal{F}_T)$, and by $\mathscr{C}_V$ the class of $\mathbb{R}^d$-valued adapted processes $V$ on $[0,T]$ such that $V_T \in L^1(\mathcal{F}_T)$. Then, the methodology considered in the previous chapters induces us to associate such a FBSDE to the system of functional differential equations given by

$$\begin{cases} dX_t = \mu\big(t, X_t, \mathcal{Y}^\phi(X,V)_t, \mathcal{Z}^\phi(X,V)_t\big)dt + \sigma(t, X_t, \mathcal{Y}^\phi(X,V)_t, \mathcal{Z}^\phi(X,V)_t)dW_t, \\ dV_t = f\big(t, X_t, \mathcal{Y}^\phi(X,V)_t, \mathcal{Z}^\phi(X,V)_t\big)dt, \\ X_0 = x, \quad V_0 = 0, \end{cases}$$

where the operators $\mathcal{M}^\phi$ and $\mathcal{Y}^\phi$ are defined on $\mathscr{C}_X^\phi \times \mathscr{C}_V$ by

$$\begin{aligned} \mathcal{M}^\phi(X,V)_t &:= E[\phi(X_T) + V_T | \mathcal{F}_t], \\ \mathcal{Y}^\phi(X,V)_t &:= \mathcal{M}^\phi(X,V)_t - V_t, \quad t \in [0,T], \end{aligned} \tag{3.1.2}$$

whereas $\mathcal{Z}^\phi$ is given implicitly via Itô's representation theorem by

$$\mathcal{M}^\phi(X,V)_T = E[\mathcal{M}^\phi(X,V)_T] + \int_0^T \mathcal{Z}^\phi(X,V)_s dW_s,\ (X,V) \in \mathscr{C}_X^\phi \times \mathscr{C}_V.$$

Therefore, the solution of the FBSDE can be obtained by solving a system of coupled functional equations which are both running forward: this allows us to partly avoid the conflicting nature between the forward and backward components.

For the rest of this chapter, we will drop the dependence of $\mathcal{M}^\phi$, $\mathcal{Y}^\phi$ and $\mathcal{Z}^\phi$ on $\phi$ by writing $\mathcal{M}$, $\mathcal{Y}$ and $\mathcal{Z}$. As in the case of simple BSDEs, the above approach can be made rigorous and extended to a much more general framework. To this end, we first need to introduce some notation: in the following, we fix $T > 0$, and assume that we are given a complete probability space $(\Omega, \mathcal{F}, P)$ together with a general filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ satisfying the usual assumptions. We recall that every $\mathbb{F}$-martingale has under these conditions a càdlàg version, which we will always choose. Moreover, we denote by $\mathcal{P}$ the predictable $\sigma$-field with respect to $\mathbb{F}$ and assume that an $m$-dimensional Brownian motion $W = (W_t)_{t\in[0,T]}$ is defined on $(\Omega, \mathcal{F}, \mathbb{F}, P)$.

As in the previous chapter, for $k$, $l \in \mathbb{N}$, $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^k$, respectively the Hilbert-Schmidt norm on $\mathbb{R}^{k\times l}$, and $\mathbb{R}^{k\times l}$ will often be identified with $\mathbb{R}^{k\cdot l}$. We define $\mathscr{S}^2([0,T],\mathbb{R}^d)$ as the space of all processes $V : \Omega \times [0,T] \to \mathbb{R}^d$ continuous and adapted such that $V_0 = 0$ and $E[\sup_{t\in[0,T]} |V_t|^2] < \infty$, while $\mathscr{M}^2([0,T],\mathbb{R}^d)$ denotes the space of all square integrable $\mathbb{R}^d$-valued martingales on $[0,T]$. Both $\mathscr{S}^2([0,T],\mathbb{R}^d)$ and $\mathscr{M}^2([0,T],\mathbb{R}^d)$ are endowed with the norm

$$\|V\|_{\mathscr{S}^2[0,T]} := \sqrt{E\Big[\sup_{t\in[0,T]} |V_t|^2\Big]},$$

and $\big(\mathscr{S}^2([0,T],\mathbb{R}^d), \|\cdot\|_{\mathscr{S}^2[0,T]}\big)$ then becomes a Banach space. Sometimes, we will also need the direct sum space $\mathscr{S}^2([0,T],\mathbb{R}^d) \oplus \mathscr{M}^2([0,T],\mathbb{R}^d)$, endowed with the same norm $\|\cdot\|_{\mathscr{S}^2[0,T]}$.

Finally, similarly to simple Lipschitz backward dynamics, we will assume that the coefficients $\mu$, $\sigma$ and $f$ depend on more general processes $\mathcal{L}^i(M)$ than just $Z$. In such a case, $\mathcal{L}^i$ are as usual operators mapping $\mathscr{M}^2([0,T],\mathbb{R}^d)$ into spaces of $p_i$-dimensional adapted processes: this will allow both to take into account the generality of the filtration $\mathbb{F}$ and to treat locally other types of forward-backward equations not fitting in the classical framework. We then consider the following forward-backward stochastic dynamics:

$$\begin{cases} dX_t = \mu(t, X_t, Y_t, \mathcal{L}^1(M)_t)dt + \sigma(t, X_t, Y_t, \mathcal{L}^2(M)_t)dW_t, \\ dY_t = -f(t, X_t, Y_t, \mathcal{L}^3(M)_t)dt + dM_t, \\ X_0 = x, \quad Y_T = \phi(X_T), \end{cases} \tag{3.1.3}$$

where $\mu : \Omega\times[0,T]\times\mathbb{R}^n\times\mathbb{R}^d\times\mathbb{R}^{p_1} \to \mathbb{R}^n$, $\sigma : \Omega\times[0,T]\times\mathbb{R}^n\times\mathbb{R}^d\times\mathbb{R}^{p_2} \to \mathbb{R}^{n\times m}$, $f : \Omega \times [0,T] \times \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^{p_3} \to \mathbb{R}^d$, $\phi : \Omega \times \mathbb{R}^n \to \mathbb{R}^d$ have the necessary measurability and integrability properties.

A solution to (3.1.3) is a triplet of processes $(X, Y, M)$ such that $X \in \mathscr{S}^2([0,T],\mathbb{R}^n)$, $Y \in \mathscr{S}^2([0,T],\mathbb{R}^d) \oplus \mathscr{M}^2([0,T],\mathbb{R}^d)$, $M \in \mathscr{M}^2([0,T],\mathbb{R}^d)$, and satisfying the integral formulation of (3.1.3). The previous Brownian example suggests us a viable approach to the solution of (3.1.3): namely, with the help of the operators $\mathcal{Y}$ and $\mathcal{M}$ (whose definition can obviously be extended to the general space $(\Omega, \mathcal{F}, \mathbb{F}, P)$), the problem can be reformulated as the following system of functional differential equations:

$$\begin{cases} dX_t = \mu(t, X_t, \mathcal{Y}(X,V)_t, \mathcal{L}^1(\mathcal{M}(X,V))_t)dt + \sigma(t, X_t, \mathcal{Y}(X,V)_t, \mathcal{L}^2(\mathcal{M}(X,V))_t)dW_t, \\ dV_t = f(t, X_t, \mathcal{Y}(X,V)_t, \mathcal{L}^3(\mathcal{M}(X,V))_t)dt, \\ X_0 = x, \quad V_0 = 0. \end{cases} \tag{3.1.4}$$

The system is then completely determined by $(\mu, \sigma, f, \phi, \mathcal{L}^1, \mathcal{L}^2, \mathcal{L}^3)$. Moreover, it is easy to check that the problems (3.1.3) and (3.1.4) are equivalent: indeed, if $(X, Y, M)$ solves (3.1.3), then we obtain a solution of (3.1.4) via the canonical decomposition of $Y$, and conversely, if $(X, V)$ is a solution of the functional differential equation, then $\big(X, \mathcal{Y}(X,V), \mathcal{M}(X,V)\big)$ solves (3.1.3).

As mentioned previously, this approach has the important advantage of increasing the homogeneity of the problem. Indeed, while the system (3.1.3) consists of a forward and a backward equation that don't share a common structure, both the functional equations in (3.1.4) are running forward in time and show a similar dependence of the coefficients on both the present and the terminal values of the solution processes. In particular, this homogeneity allows to rewrite the problem more compactly as

$$\begin{cases} d\mathcal{U}_t = \Psi(t, \pi_1(\mathcal{U}_t), \mathcal{Y}(\mathcal{U})_t, \mathcal{L}^1(\mathcal{M}(\mathcal{U}))_t, \mathcal{L}^3(\mathcal{M}(\mathcal{U}))_t)dt \\ \qquad\qquad\qquad +\Sigma(t, \pi_1(\mathcal{U}_t), \mathcal{Y}(\mathcal{U})_t, \mathcal{L}^2(\mathcal{M}(\mathcal{U}))_t)dW_t, \\ \mathcal{U}_0 = (x, 0)^{\mathrm{T}}, \end{cases}$$

where $\mathcal{U} = (X, V)^{\mathrm{T}}$, $\pi_1 : \mathbb{R}^{n+d} \to \mathbb{R}^n$ is the projection on the first $n$ components, $\Sigma = (\sigma, 0)^{\mathrm{T}}$, and for $t \in [0, T]$, $(x, y, z_1, z_3) \in \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^{p_1} \times \mathbb{R}^{p_3}$, $\Psi(t, x, y, z_1, z_3) = (\mu(t, x, y, z_1), f(t, x, y, z_3))^{\mathrm{T}}$. For the remainder of the chapter, we prefer however to consider the system as formulated in (3.1.4), since it will be more convenient to study in such a framework the coupling between $X$ and $V$.

## 3.2 Existence and uniqueness of local solutions

The purpose of this section is to prove that, under some monotonicity and Lipschitz assumptions on the coefficients $\mu$, $\sigma$, $f$, $\phi$ and the functionals $\mathcal{L}^1$, $\mathcal{L}^2$ and $\mathcal{L}^3$, the system of functional differential equations introduced in the previous section has a unique solution, provided that the time horizon $T$ is sufficiently small.

We point out that, for our main result to hold, it is sufficient that $\mathcal{L}^1$, $\mathcal{L}^2$ and $\mathcal{L}^3$ satisfy the same Lipschitz and boundedness assumptions as for simple backward stochastic dynamics. As already anticipated in the Introduction, these conditions are rather mild and allow to study many types of operators different from the usual martingale integrand processes considered in classical FBSDEs: we emphasize the generality of these assumptions by giving several examples of functionals fitting within our framework, and we present possible financial applications.

In the following, we denote by $\mathscr{H}^2([0,T],\mathbb{R}^l)$ the space of $\mathcal{P}$-measurable processes $H:\Omega\times[0,T]\to\mathbb{R}^l$ such that $\|H\|^2_{\mathscr{H}^2[0,T]} := E[\int_0^T |H_t|^2 dt] < \infty$, where $l\in\mathbb{N}$. We shall assume that the coefficients of (3.1.4) satisfy the following conditions:

**Assumption (B1):** *The functions* $\mu:\Omega\times[0,T]\times\mathbb{R}^n\times\mathbb{R}^d\times\mathbb{R}^{p_1}\to\mathbb{R}^n$, $\sigma:\Omega\times[0,T]\times\mathbb{R}^n\times\mathbb{R}^d\times\mathbb{R}^{p_2}\to\mathbb{R}^{n\times m}$, $f:\Omega\times[0,T]\times\mathbb{R}^n\times\mathbb{R}^d\times\mathbb{R}^{p_3}\to\mathbb{R}^d$ *and* $\phi:\Omega\times\mathbb{R}^n\to\mathbb{R}^d$ *satisfy Assumption* **(B1)** *if there exists a constant* $C>0$ *such that:*

(B1.1) *For any* $(x,y,z_1,z_2,z_3)\in\mathbb{R}^n\times\mathbb{R}^d\times\mathbb{R}^{p_1}\times\mathbb{R}^{p_2}\times\mathbb{R}^{p_3}$, *the processes* $\mu(\cdot,x,y,z_1)$, $\sigma(\cdot,x,y,z_2)$ *and* $f(\cdot,x,y,z_3)$ *are* $\mathcal{P}$*-measurable and* $\phi(x)$ *is* $\mathcal{F}_T$*-measurable.*

(B1.2) *For every* $(x,y,z_1),(x',y',z_1')\in\mathbb{R}^n\times\mathbb{R}^d\times\mathbb{R}^{p_1}$,

$$\begin{aligned}(x-x')^{\mathrm{T}}\big(\mu(\cdot,x,y,z_1)-\mu(\cdot,x',y,z_1)\big) &\le C|x-x'|^2,\\ |\mu(\cdot,x,y,z_1)-\mu(\cdot,x,y',z_1')| &\le C(|y-y'|+|z_1-z_1'|),\\ |\mu(\cdot,x,0,0)| &\le C(1+|x|) \quad dP\otimes dt\text{-a.s.},\end{aligned}$$

*and the function* $u\mapsto\mu(\cdot,u,y,z_1)$ *is* $dP\otimes dt$*-a.s. continuous.*

(B1.3) $f(\cdot,0,0,0)\in\mathscr{H}^2([0,T],\mathbb{R}^d)$, $\sigma(\cdot,0,0,0)\in\mathscr{H}^2([0,T],\mathbb{R}^{n\times m})$ *and* $\phi(0)\in L^2(\Omega,\mathbb{R}^d)$.

(B1.4) *For every* $(x,y,z_2,z_3),(x',y',z_2',z_3')\in\mathbb{R}^n\times\mathbb{R}^d\times\mathbb{R}^{p_2}\times\mathbb{R}^{p_3}$,

$$\begin{aligned}|\sigma(\cdot,x,y,z_2)-\sigma(\cdot,x',y',z_2')|^2 &\le C(|x-x'|^2+|y-y'|^2+|z_2-z_2'|^2),\\ |f(\cdot,x,y,z_3)-f(\cdot,x',y',z_3')| &\le C(|x-x'|+|y-y'|+|z_3-z_3'|),\\ |\phi(x)-\phi(x')| &\le C|x-x'| \quad dP\otimes dt\text{-a.s.}.\end{aligned}$$

Observe that these conditions are quite standard in the theory of FBSDEs (see for instance [51]). The reader can easily verify that the condition (B1.2) could also be replaced by the following stronger assumption:

(B1.2') $\mu(\cdot,0,0,0)\in\mathscr{H}^2([0,T],\mathbb{R}^n)$ *and, for* $(x,y,z_1),(x',y',z_1')\in\mathbb{R}^n\times\mathbb{R}^d\times\mathbb{R}^{p_1}$,

$$|\mu(t,x,y,z_1)-\mu(t,x',y',z_1')|\le C(|x-x'|+|y-y'|+|z_1-z_1'|)\quad dP\otimes dt\text{-}a.s..$$

Moreover, as a consequence of the assumptions on $\phi$, we obtain without much effort the following estimates:

**Lemma 3.2.1.** *Assume that $\phi$ satisfies the conditions in* **(B1)**. *Then we have that, for the functionals introduced in* (3.1.2)*,*

$$\mathcal{Y}:\mathscr{S}^2([0,T],\mathbb{R}^n)\times\mathscr{S}^2([0,T],\mathbb{R}^d)\to\mathscr{S}^2([0,T],\mathbb{R}^d)\oplus\mathscr{M}^2([0,T],\mathbb{R}^d),$$
$$\mathcal{M}:\mathscr{S}^2([0,T],\mathbb{R}^n)\times\mathscr{S}^2([0,T],\mathbb{R}^d)\to\mathscr{M}^2([0,T],\mathbb{R}^d).$$

*Moreover, for any* $X,X'\in\mathscr{S}^2([0,T],\mathbb{R}^n)$ *and* $V,V'\in\mathscr{S}^2([0,T],\mathbb{R}^d)$,

$$\|\mathcal{Y}(X,V)-\mathcal{Y}(X',V')\|_{\mathscr{S}^2[0,T]}\le 2C\|X-X'\|_{\mathscr{S}^2[0,T]}+3\|V-V'\|_{\mathscr{S}^2[0,T]},$$
$$\|\mathcal{M}(X,V)-\mathcal{M}(X',V')\|_{\mathscr{S}^2[0,T]}\le 2C\|X-X'\|_{\mathscr{S}^2[0,T]}+2\|V-V'\|_{\mathscr{S}^2[0,T]}.$$

*Proof.* We first prove the second assertion. By the triangle inequality,

$$\begin{aligned}&\|\mathcal{M}(X,V)-\mathcal{M}(X',V')\|_{\mathscr{S}^2[0,T]}\\&\le\big\|E[\phi(X_T)-\phi(X_T')|\mathcal{F}_\cdot]\big\|_{\mathscr{S}^2[0,T]}+\big\|E[V_T-V_T'|\mathcal{F}_\cdot]\big\|_{\mathscr{S}^2[0,T]}\\&=E\Big[\sup_{t\in[0,T]}\big|E[\phi(X_T)-\phi(X_T')|\mathcal{F}_t]\big|^2\Big]^{1/2}+E\Big[\sup_{t\in[0,T]}\big|E[V_T-V_T'|\mathcal{F}_t]\big|^2\Big]^{1/2}\end{aligned}$$

and therefore, by Doob's inequality and the assumption on $\phi$,

$$\begin{aligned}\|\mathcal{M}(X,V)-\mathcal{M}(X',V')\|_{\mathscr{S}^2[0,T]}&\le 2\Big(E\big[|\phi(X_T)-\phi(X_T')|^2\big]^{1/2}+E\big[|V_T-V_T'|^2\big]^{1/2}\Big)\\&\le 2C\|X-X'\|_{\mathscr{S}^2[0,T]}+2\|V-V'\|_{\mathscr{S}^2[0,T]}.\end{aligned}$$

The estimate for $\|\mathcal{Y}(X,V)-\mathcal{Y}(X',V')\|_{\mathscr{S}^2[0,T]}$ then follows by applying again the triangle inequality. □

These Lipschitz estimates for $\mathcal{Y}$ and $\mathcal{M}$ will be very useful in the sequel. Clearly, the above assumptions on the coefficients $\mu$, $\sigma$, $f$ and $\phi$ aren't enough to hope for general results on existence and uniqueness of solutions, without further specification for the abstract functionals $\mathcal{L}^1$, $\mathcal{L}^2$ and $\mathcal{L}^3$. As already anticipated, the same Lipschitz and boundedness assumptions as the ones introduced in [49] will be sufficient: we report them here for the reader's convenience.

**Assumption (L1):** *The functional $\mathcal{L}$ satisfies Assumption* **(L1)** *if:*

(L1.1) *$\mathcal{L}$ maps $\mathscr{M}^2([0,T],\mathbb{R}^d)$ into $\mathscr{O}^2([0,T],\mathbb{R}^p)$, where $\mathscr{O}^2([0,T],\mathbb{R}^p)$ is either the space $\mathscr{H}^2([0,T],\mathbb{R}^p)$ or $\mathscr{S}^2([0,T],\mathbb{R}^p)$.*

(L1.2) *$\mathcal{L}$ is bounded and Lipschitz continuous, i.e. there exists a constant $K>0$ independent of $T$ such that, for all $M, M' \in \mathscr{M}^2([0,T],\mathbb{R}^d)$,*

$$\begin{aligned} \|\mathcal{L}(M)\|_{\mathscr{O}^2[0,T]} &\leq K\|M\|_{\mathscr{S}^2[0,T]}, \\ \|\mathcal{L}(M)-\mathcal{L}(M')\|_{\mathscr{O}^2[0,T]} &\leq K\|M-M'\|_{\mathscr{S}^2[0,T]}. \end{aligned}$$

**Examples 3.2.2.** To convince the reader of the generality of Assumption **(L1)**, we give here some examples of possible operators $\mathcal{L}$. We begin with the classical case of integrand processes generated by martingale representations.

$(i)$ Assume that $(\mathcal{F}_t)_{t\in[0,T]}$ is the augmented filtration generated by the Brownian motion $W$, and take $\mathscr{O}^2([0,T],\mathbb{R}^p) = \mathscr{H}^2([0,T],\mathbb{R}^{d\times m})$. Then, we define $\mathcal{L}: \mathscr{M}^2([0,T],\mathbb{R}^d) \to \mathscr{H}^2([0,T],\mathbb{R}^{d\times m})$ implicitly via Itô's representation theorem by

$$M_t = M_0 + \int_0^t \mathcal{L}(M)_s dW_s, \quad t \in [0,T].$$

By Itô's isometry we have that

$$\begin{aligned} \|\mathcal{L}(M)\|^2_{\mathscr{H}^2[0,T]} &= E\Big[\int_0^T |\mathcal{L}(M)_t|^2 dt\Big] \\ &= E\big[\big(M_T - M_0\big)^2\big] = E[M_T^2] - E[M_0^2] \leq \|M\|^2_{\mathscr{S}^2[0,T]}, \end{aligned}$$

and the Lipschitz property follows by the linearity of $\mathcal{L}$. Note that in the case where $\mathcal{L}^1 = \mathcal{L}^2 = \mathcal{L}^3 = \mathcal{L}$, the system (3.1.3) is reduced to a classical FBSDE. The applications of FBSDEs to financial markets are countless: for an overview, we refer the reader in particular to [32, 51].

(*ii*) Let now $(\mathcal{F}_t)_{t\in[0,T]}$ be a general filtration with just the usual assumptions. As in the previous case, we can take $\mathscr{O}^2([0,T],\mathbb{R}^p) = \mathscr{H}^2([0,T],\mathbb{R}^{d\times m})$, and define $\mathcal{L} : \mathscr{M}^2([0,T],\mathbb{R}^d) \to \mathscr{H}^2([0,T],\mathbb{R}^{d\times m})$ implicitly via the orthogonal decomposition with respect to $W$, i.e.

$$M_t = \int_0^t \mathcal{L}(M)_s dW_s + N_t, \quad t \in [0,T],$$

where $N$ is some martingale orthogonal with respect to $W$. The reader may easily notice the connection between this functional and generalized BSDEs (see [31]). Moreover, because of the orthogonality, we can prove similarly to (*i*) that $\mathcal{L}$ satisfies **(L1)**, by applying the Burkholder-Davis-Gundy inequality instead of Itô's isometry. Hence, generalized fully coupled FBSDEs can be treated within our framework, and by the generality of the filtration, we can extend the financial applications of FBSDEs to the case of incomplete markets. Consider for instance a large investor trading in an incomplete market. Since this investor buys and sells large amounts of assets, it is reasonable to assume that his trading strategy affects the prices of the stocks: by considering the hedging problem for this investor, we therefore obtain a fully coupled system, and the incompleteness of the market leads to a generalized fully coupled FBSDE. For more details, we refer the reader to [51].

While martingale integrand processes are the case most studied in the literature, they are not the only type of functional treatable within our framework: there are indeed several other classes of non-local operators satisfying Assumption **(L1)**. Let us give some examples:

(*iii*) Assume that $(\mathcal{F}_t)_{t\in[0,T]}$ just satisfies the usual assumptions. We take $\mathscr{O}^2([0,T],\mathbb{R}^p) = \mathscr{S}^2([0,T],\mathbb{R}^d)$, and $\mathcal{L} : \mathscr{M}^2([0,T],\mathbb{R}^d) \to \mathscr{S}^2([0,T],\mathbb{R}^d)$ is simply defined by $\mathcal{L}(M) := M$; in this case, Assumption **(L1)** becomes trivial. In a financial context, $\mathcal{L}(M)$ may represent the diffusion part of the claim $Y = M - V$.

(*iv*) Let for simplicity $d = 1$. Fix $\widetilde{T} > 0$ and assume that, for $T \leq \widetilde{T}$, $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is such that all martingales with respect to $\mathbb{F}$ are continuous. Let $\mathscr{O}^2([0,T],\mathbb{R}^p) = \mathscr{H}^2([0,T],\mathbb{R})$, and define $\mathcal{L} : \mathscr{M}^2([0,T],\mathbb{R}) \to \mathscr{H}^2([0,T],\mathbb{R})$ by

$$\mathcal{L}(M)_t := \sqrt{E\big[\langle M\rangle_{t,T}\big|\mathcal{F}_t\big]}, \quad M \in \mathscr{M}^2([0,T],\mathbb{R}),\ t \in [0,T],$$

where $\langle M\rangle_{t,T} := \langle M\rangle_T - \langle M\rangle_t$. We recall that this functional has already been considered in Chapter 2 (see Example 2.2.2 (*ii*)): there, by applying the Kunita-Watanabe inequality we obtained that, for $t \geq 0$,

$$\big|\mathcal{L}(M)_t - \mathcal{L}(M')_t\big|^2 \leq E[\langle M - M'\rangle_{t,T}|\mathcal{F}_t],$$

and therefore, by Fubini's theorem,

$$\begin{aligned}\|\mathcal{L}(M) - \mathcal{L}(M')\|^2_{\mathscr{H}^2[0,T]} &= E\left[\int_0^T \big|\mathcal{L}(M)_t - \mathcal{L}(M')_t\big|^2 dt\right]\\ &\leq E\left[\int_0^T E\big[\langle M - M'\rangle_{t,T}\big|\mathcal{F}_t\big]dt\right]\\ &= E\left[\int_0^T \langle M - M'\rangle_{t,T} dt\right] \leq \widetilde{T} E[\langle M - M'\rangle_T].\end{aligned}$$

By applying the Burkholder-Davis-Gundy inequality, we finally get the desired Lipschitz property. The boundedness condition is obtained via similar computations.

(*v*) We choose again for simplicity $d = 1$, and we assume that $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is as in the example (*iv*). We can define $\mathcal{L}$ by taking $\mathscr{O}^2([0,T],\mathbb{R}^p) = \mathscr{S}^2([0,T],\mathbb{R})$, and

$$\mathcal{L} : \mathscr{M}^2([0,T],\mathbb{R}) \to \mathscr{S}^2([0,T],\mathbb{R}), \quad \mathcal{L}(M)_t := \sqrt{\langle M\rangle_t},\ t \in [0,T].$$

By the Burkholder-Davis-Gundy inequality, we have that

$$\|\mathcal{L}(M)\|^2_{\mathscr{S}^2[0,T]} = E\Big[\sup_{t\in[0,T]} \langle M\rangle_t\Big] = E[\langle M\rangle_T] \leq K\|M\|^2_{\mathscr{S}^2[0,T]}.$$

On the other hand, by the Kunita-Watanabe inequality,

$$\begin{aligned}\left|\sqrt{\langle M\rangle_t}-\sqrt{\langle M'\rangle_t}\right|^2 &= \langle M\rangle_t+\langle M'\rangle_t-2\sqrt{\langle M\rangle_t\langle M'\rangle_t}\\ &\leq \langle M\rangle_t+\langle M'\rangle_t-2\big|\langle M,M'\rangle_t\big|\\ &\leq \langle M\rangle_t+\langle M'\rangle_t-2\langle M,M'\rangle_t=\langle M-M'\rangle_t,\end{aligned}$$

and the Lipschitz property then follows by applying the Burkholder-Davis-Gundy inequality as in the previous example. We restrict for a moment to the Brownian setting to give a financial interpretation: in this case, $\mathcal{L}(M)$ may be explicitly rewritten as $\mathcal{L}(M)_t=\sqrt{\int_0^t|Z_s|^2ds}$, where $Z$ is the martingale integrand in the Itô representation of $M$. In the usual BSDE framework for hedging (see for instance [32]), $\mathcal{L}(M)$ is then closely connected to the accumulated cost of the portfolio strategy: this could allow us, for instance, to model some types of storage problems.

$(vi)$ We modify the previous example by combining it with orthogonal decompositions. Let $d=1$ and $\mathbb{F}=(\mathcal{F}_t)_{t\in[0,T]}$ as in the example $(iv)$. Fix a martingale $\widetilde{M}$, and define $\mathcal{R}:\mathscr{M}^2([0,T],\mathbb{R})\to\mathscr{M}^2([0,T],\mathbb{R})$ as the orthogonal term in the orthogonal decomposition with respect to $\widetilde{M}$, i.e.

$$M_t=\big(M_t-\mathcal{R}(M)_t\big)+\mathcal{R}(M)_t,\quad t\in[0,T],$$

where $\mathcal{R}(M)$ is orthogonal with respect to $\widetilde{M}$, and $M_t-\mathcal{R}(M)_t=\int_0^t Z_s d\widetilde{M}_s$ for some process $Z$. Then, we define the functional $\mathcal{L}$ by

$$\mathcal{L}:\mathscr{M}^2([0,T],\mathbb{R})\to\mathscr{S}^2([0,T],\mathbb{R}),\quad \mathcal{L}(M)_t:=\sqrt{\langle\mathcal{R}(M)\rangle_t},\ t\in[0,T].$$

Because of the orthogonality, it is easy to check that, for all $t\geq 0$,

$$\begin{aligned}\langle M\rangle_t &= \langle M-\mathcal{R}(M)+\mathcal{R}(M)\rangle_t\\ &= \langle M-\mathcal{R}(M)\rangle_t+\langle\mathcal{R}(M)\rangle_t\geq\langle\mathcal{R}(M)\rangle_t,\end{aligned}$$

and similarly for $M-M'$. Therefore, $\mathcal{L}$ satisfies Assumption **(L1)** because of example $(v)$. This functional may have interesting applications in mathematical finance: namely, in the typical BSDE framework for hedging in incomplete markets, the operator $\mathcal{R}(M)$ represents the non-hedgeable part of the claim. While we cannot hedge this risk, we can partly incorporate its effect on the price process since the coefficients $\mu$, $\sigma$ and $f$ are allowed to depend on $\sqrt{\langle \mathcal{R}(M) \rangle_t}$.

(*vii*) As a final example, we introduce a functional intimately related to backward equations with time delayed generators (see [26]). Let $\widetilde{T} > 0$ be fixed. For $T \leq \widetilde{T}$, let $(\mathcal{F}_t)_{t\in[0,T]}$ satisfy the usual assumptions, and $\mathscr{O}^2([0,T],\mathbb{R}^p) = \mathscr{H}^2([0,T],\mathbb{R}^{d\times m})$. Motivated by the framework introduced by Dos Reis et al. [27], we can then define $\mathcal{L} : \mathscr{M}^2([0,T],\mathbb{R}^d) \to \mathscr{H}^2([0,T],\mathbb{R}^{d\times m})$ by

$$\mathcal{L}(M)_t := \int_{-t}^{0} \widehat{\mathcal{L}}(M)_{t+s}\alpha_Z(ds),$$

where $\widehat{\mathcal{L}}$ is the functional introduced in (*ii*), and $\alpha_Z$ is a non-random finite measure with support in $[-\widetilde{T},0]$. Observe that $\mathcal{L}$ is closely related to the functional in (*v*) when $\alpha_Z$ is the Lebesgue measure restricted to $[-\widetilde{T},0]$. Moreover, by applying the change of integration order proved in [27], we can show that

$$\begin{aligned} \|\mathcal{L}(M)\|_{\mathscr{H}^2[0,T]} &= E\left[\int_0^T \left|\int_{-t}^0 \widehat{\mathcal{L}}(M)_{t+s}\alpha_Z(ds)\right|^2 dt\right]^{1/2} \\ &\leq \alpha_Z([-\widetilde{T},0])\|\widehat{\mathcal{L}}(M)\|_{\mathscr{H}^2[0,T]} \\ &\leq K\|M\|_{\mathscr{S}^2[0,T]} \end{aligned} \tag{3.2.1}$$

for some constant $K = K(\widetilde{T})$, where the last inequality follows from the boundedness of $\widehat{\mathcal{L}}$. Assumption **(L1)** then follows by linearity.

This functional $\mathcal{L}$ leads us to a particular class of fully coupled forward-backward equations with delayed generators. Let us note however that, as already mentioned in Remark 1.2.3 in the Introduction, the study of backward delayed equations in their full generality requires to replace the operator $\mathcal{Y}(X,V)$ in (3.1.4) by a new functional $\mathbb{Y}(X,V)_t :=$

$\int_{-t}^{0} \mathcal{Y}(X,V)_{t+s}\alpha_Y(ds)$, where $\alpha_Y$ is a non-random finite measure: as pointed out in [27], this kind of equation is very difficult to study even in the simple decoupled case, and it will not be treated here. For an overview of the several applications of backward equations with time delay to mathematical finance, the reader may consult for instance the article by Delong [25].

These examples should help us understand the importance of the next theorem: this is the main result of this section, and states the existence of a unique square-integrable solution to the system (3.1.4) on sufficiently small intervals $[0,T]$, provided that the functionals $\mathcal{L}^i$ satisfy Assumption **(L1)**. The generality of **(L1)** therefore gives us a great flexibility in the choice of $\mathcal{L}^i$, allowing to study many different types of coupled forward-backward stochastic systems not considered in the classical literature.

**Theorem 3.2.3.** *Let $\mu$, $\sigma$, $f$ and $\phi$ satisfy Assumption* **(B1)** *with respect to the constant $C$. Furthermore, assume that $\mathcal{L}^1$, $\mathcal{L}^3$ satisfy Assumption* **(L1)** *and that $\mathcal{L}^2$ satisfies* **(L1)** *with respect to $\mathscr{O}^2([0,T],\mathbb{R}^{p_2}) = \mathscr{S}^2([0,T],\mathbb{R}^{p_2})$, denoting by $K$ the common Lipschitz constant of $\mathcal{L}^1$, $\mathcal{L}^2$, $\mathcal{L}^3$. Then there is a constant $\ell = \ell(C,K)$ depending only on $C$ and $K$ so that, for $T < \ell$,* (3.1.4) *admits a unique solution $(X,V)$ in $\mathscr{S}^2([0,T],\mathbb{R}^n) \times \mathscr{S}^2([0,T],\mathbb{R}^d)$.*

*Proof.* From now on, we will write for notational simplicity $\|\cdot\|^2_{\mathscr{O}}$ for the norm $\|\cdot\|_{\mathscr{O}^2[0,T]}$. Moreover, we denote the product space $\mathscr{S}^2([0,T],\mathbb{R}^n) \times \mathscr{S}^2([0,T],\mathbb{R}^d)$ by $\mathscr{S}^2_X \times \mathscr{S}^2_V$, and endow it with the norm

$$\|(X,V)\|_{\mathscr{S}^2_X \times \mathscr{S}^2_V} := \sqrt{\|X\|^2_{\mathscr{S}^2} + \|V\|^2_{\mathscr{S}^2}}, \quad (X,V) \in \mathscr{S}^2_X \times \mathscr{S}^2_V.$$

The mapping $\mathbb{L} : \mathscr{S}^2_X \times \mathscr{S}^2_V \to \mathscr{S}^2_X \times \mathscr{S}^2_V$, $\mathbb{L}(X,V) := (\widetilde{X},\widetilde{V})$, is defined as follows: first, $\widetilde{X}$ is constructed as the unique solution in $\mathscr{S}^2([0,T],\mathbb{R}^n)$ to the forward stochastic differential equation

$$\begin{cases} d\widetilde{X}_t = \mu(t,\widetilde{X}_t,\mathcal{Y}(X,V)_t,\mathcal{L}^1(\mathcal{M}(X,V))_t)dt + \sigma(t,\widetilde{X}_t,\mathcal{Y}(X,V)_t,\mathcal{L}^2(\mathcal{M}(X,V))_t)dW_t, \\ \widetilde{X}_0 = x. \end{cases} \tag{3.2.2}$$

Then, once $\widetilde{X}$ has been obtained, $\widetilde{V}$ is given explicitly by the expression

$$\widetilde{V}_t = \int_0^t f(s, \widetilde{X}_s, \mathcal{Y}(X,V)_s, \mathcal{L}^3(\mathcal{M}(X,V))_s) ds, \tag{3.2.3}$$

We show that the mapping $\mathbb{L}$ is well defined and maps $\mathscr{S}^2_X \times \mathscr{S}^2_V$ into itself. First of all, we note that the existence of a unique solution in $\mathscr{S}^2([0,T], \mathbb{R}^n)$ to (3.2.2) follows by Assumption **(B1)** and the results on stochastic differential equations with monotonous coefficients obtained, for instance, by Rozovsky [66]: indeed, by setting $\bar{\mu}(t,x) := \mu(t,x,\mathcal{Y}(X,V)_t, \mathcal{L}^1(\mathcal{M}(X,V))_t)$ and $\bar{\sigma}(t,x) := \sigma(t,x,\mathcal{Y}(X,V)_t, \mathcal{L}^2(\mathcal{M}(X,V))_t)$, the reader can easily check that all the conditions of [66] are satisfied, as $\|\mathcal{Y}(X,V)\|_{\mathscr{S}^2}$ and $\|\mathcal{M}(X,V)\|_{\mathscr{S}^2}$ are finite by Lemma 3.2.1. On the other hand, it is not difficult to verify that $\widetilde{V} \in \mathscr{S}^2([0,T], \mathbb{R}^d)$, as for $f_0 := f(\cdot, 0, 0, 0)$,

$$\begin{aligned}
\|\widetilde{V}\|_{\mathscr{S}^2} &\leq \sqrt{T}\Big(\|f_0\|_{\mathscr{H}^2} + \|f\big(\cdot, \widetilde{X}_\cdot, \mathcal{Y}(X,V)_\cdot, \mathcal{L}^3(\mathcal{M}(X,V))_\cdot\big) - f_0\|_{\mathscr{H}^2}\Big) \\
&\leq \sqrt{T}\Big(\|f_0\|_{\mathscr{H}^2} + C(\sqrt{T} \vee 1)\big(\|\widetilde{X}\|_{\mathscr{S}^2} + \|\mathcal{Y}(X,V)\|_{\mathscr{S}^2} + \|\mathcal{L}^3(\mathcal{M}(X,V))\|_{\mathscr{O}^2}\big)\Big) \\
&\leq \sqrt{T}\Big(\|f_0\|_{\mathscr{H}^2} + C(\sqrt{T} \vee 1)\big(\|\widetilde{X}\|_{\mathscr{S}^2} + \|\mathcal{Y}(X,V)\|_{\mathscr{S}^2} + K\|\mathcal{M}(X,V)\|_{\mathscr{S}^2}\big)\Big) < \infty.
\end{aligned}$$

Since the pair $(X,V)$ is a solution of (3.1.4) if and only if it is a fixed point of $\mathbb{L}$, it suffices to prove that $\mathbb{L}$ is a contraction on $\mathscr{S}^2_X \times \mathscr{S}^2_V$ for small enough $T > 0$. Let $\mathbb{L}(X^1, V^1) = (\widetilde{X}^1, \widetilde{V}^1)$, $\mathbb{L}(X^2, V^2) = (\widetilde{X}^2, \widetilde{V}^2)$, and assume without loss of generality that $T \leq 1$. By Itô's formula, we can compute that

$$\begin{aligned}
d|\widetilde{X}^1_t - \widetilde{X}^2_t|^2 &= 2(\widetilde{X}^1_t - \widetilde{X}^2_t)^{\mathrm{T}} d(\widetilde{X}^1 - \widetilde{X}^2)_t + d\langle \widetilde{X}^1 - \widetilde{X}^2\rangle_t \\
&= 2\big(\widetilde{X}^1_t - \widetilde{X}^2_t\big)^{\mathrm{T}} \Big(\mu\big(t, \widetilde{X}^1_t, \mathcal{Y}(X^1,V^1)_t, \mathcal{L}^1(\mathcal{M}(X^1,V^1))_t\big) \\
&\qquad\qquad - \mu\big(t, \widetilde{X}^2_t, \mathcal{Y}(X^2,V^2)_t, \mathcal{L}^1(\mathcal{M}(X^2,V^2))_t\big)\Big) dt \\
&\quad + 2\big(\widetilde{X}^1_t - \widetilde{X}^2_t\big)^{\mathrm{T}} \Big(\sigma\big(t, \widetilde{X}^1_t, \mathcal{Y}(X^1,V^1)_t, \mathcal{L}^2(\mathcal{M}(X^1,V^1))_t\big) \\
&\qquad\qquad - \sigma\big(t, \widetilde{X}^2_t, \mathcal{Y}(X^2,V^2)_t, \mathcal{L}^2(\mathcal{M}(X^2,V^2))_t\big) dW_t\Big) \\
&\quad + \big|\sigma\big(t, \widetilde{X}^1_t, \mathcal{Y}(X^1,V^1)_t, \mathcal{L}^2(\mathcal{M}(X^1,V^1))_t\big) \\
&\qquad - \sigma\big(t, \widetilde{X}^2_t, \mathcal{Y}(X^2,V^2)_t, \mathcal{L}^2(\mathcal{M}(X^2,V^2))_t\big)\big|^2 dt
\end{aligned}$$

for any $t \geq 0$. Thus, by applying Assumption (B1.2), (B1.4) and classical inequalities we obtain that, for a constant $\theta_1 = \theta_1(C)$ depending only on $C$,

$$\begin{aligned}
&\|\widetilde{X}^1 - \widetilde{X}^2\|^2_{\mathscr{S}^2} = E\Big[\sup_{t\in[0,T]} |\widetilde{X}^1_t - \widetilde{X}^2_t|^2\Big] \\
&\leq \theta_1 \Bigg( E\Big[\int_0^T |\widetilde{X}^1_s - \widetilde{X}^2_s|^2 ds\Big] + E\Big[\int_0^T |\widetilde{X}^1_s - \widetilde{X}^2_s||\mathcal{Y}(X^1,V^1)_s - \mathcal{Y}(X^2,V^2)_s| ds\Big] \\
&\quad + E\Big[\int_0^T |\widetilde{X}^1_s - \widetilde{X}^2_s||\mathcal{L}^1(\mathcal{M}(X^1,V^1))_s - \mathcal{L}^1(\mathcal{M}(X^2,V^2))_s| ds\Big] \\
&\quad + E\Big[\int_0^T |\mathcal{Y}(X^1,V^1)_s - \mathcal{Y}(X^2,V^2)_s|^2 ds\Big] \\
&\quad + E\Big[\int_0^T |\mathcal{L}^2(\mathcal{M}(X^1,V^1))_s - \mathcal{L}^2(\mathcal{M}(X^2,V^2))_s|^2 ds\Big] \\
&\quad + E\Big[\Big(\int_0^T |\widetilde{X}^1_s - \widetilde{X}^2_s|^2 \big(|\widetilde{X}^1_s - \widetilde{X}^2_s|^2 + |\mathcal{Y}(X^1,V^1)_s - \mathcal{Y}(X^2,V^2)_s|^2 \\
&\qquad + |\mathcal{L}^2(\mathcal{M}(X^1,V^1))_s - \mathcal{L}^2(\mathcal{M}(X^2,V^2))_s|^2\big) ds\Big)^{1/2}\Big]\Bigg).
\end{aligned} \tag{3.2.4}$$

The next step consists in deriving estimates for all the terms on the right hand side of this inequality. The first term can simply be estimated by

$$E\Big[\int_0^T |\widetilde{X}^1_s - \widetilde{X}^2_s|^2 ds\Big] \leq T\|\widetilde{X}^1 - \widetilde{X}^2\|^2_{\mathscr{S}^2}. \tag{3.2.5}$$

For the second term, we obtain by the Cauchy-Schwarz inequality that

$$\begin{aligned}
&E\Big[\int_0^T |\widetilde{X}^1_s - \widetilde{X}^2_s||\mathcal{Y}(X^1,V^1)_s - \mathcal{Y}(X^2,V^2)_s| ds\Big] \\
&\qquad \leq \frac{T}{2}\Big(\|\widetilde{X}^1 - \widetilde{X}^2\|^2_{\mathscr{S}^2} + \|\mathcal{Y}(X^1,V^1) - \mathcal{Y}(X^2,V^2)\|^2_{\mathscr{S}^2}\Big).
\end{aligned}$$

Because of assumption (B1.4) on $\phi$, we can apply Lemma 3.2.1 and get that

$$\begin{aligned}
&E\Big[\int_0^T |\widetilde{X}^1_s - \widetilde{X}^2_s||\mathcal{Y}(X^1,V^1)_s - \mathcal{Y}(X^2,V^2)_s| ds\Big] \\
&\qquad \leq \theta_2 T\Big(\|\widetilde{X}^1 - \widetilde{X}^2\|^2_{\mathscr{S}^2} + \|V^1 - V^2\|^2_{\mathscr{S}^2} + \|X^1 - X^2\|^2_{\mathscr{S}^2}\Big),
\end{aligned} \tag{3.2.6}$$

for some constant $\theta_2 = \theta_2(C)$ depending on $C$. To estimate the third term, one can apply again the Cauchy-Schwarz inequality to obtain that

$$\begin{aligned}
&E\Big[\int_0^T |\widetilde{X}_s^1 - \widetilde{X}_s^2||\mathcal{L}^1(\mathcal{M}(X^1,V^1))_s - \mathcal{L}^1(\mathcal{M}(X^2,V^2))_s|ds\Big] \\
&\quad\le \frac{T}{2\varepsilon}\|\widetilde{X}^1 - \widetilde{X}^2\|_{\mathscr{S}^2}^2 + \frac{\varepsilon}{2}E\Big[\int_0^T |\mathcal{L}^1(\mathcal{M}(X^1,V^1))_s - \mathcal{L}^1(\mathcal{M}(X^2,V^2))_s|^2 ds\Big] \\
&\quad\le \frac{\sqrt{T}}{2}\Big(\|\widetilde{X}^1 - \widetilde{X}^2\|_{\mathscr{S}^2}^2 + \|\mathcal{L}^1(\mathcal{M}(X^1,V^1)) - \mathcal{L}^1(\mathcal{M}(X^2,V^2))\|_{\mathscr{O}^2}^2\Big),
\end{aligned}$$

where the last inequality is obtained by taking $\varepsilon = \sqrt{T}$ if $\mathscr{O}^2 = \mathscr{H}^2$ and $\varepsilon = 1$ if $\mathscr{O}^2 = \mathscr{S}^2$ (remember that $T \le 1$). On the other hand, by Assumption **(L1)** we know that

$$\|\mathcal{L}^1(\mathcal{M}(X^1,V^1)) - \mathcal{L}^1(\mathcal{M}(X^2,V^2))\|_{\mathscr{O}^2}^2 \le K^2\|\mathcal{M}(X^1,V^1) - \mathcal{M}(X^2,V^2)\|_{\mathscr{S}^2}^2,$$

and by applying Lemma 3.2.1,

$$\begin{aligned}
&E\Big[\int_0^T |\widetilde{X}_s^1 - \widetilde{X}_s^2||\mathcal{L}^1(\mathcal{M}(X^1,V^1))_s - \mathcal{L}^1(\mathcal{M}(X^2,V^2))_s|ds\Big] \\
&\qquad\qquad \le \theta_3\sqrt{T}\Big(\|\widetilde{X}^1 - \widetilde{X}^2\|_{\mathscr{S}^2}^2 + \|V^1 - V^2\|_{\mathscr{S}^2}^2 + \|X^1 - X^2\|_{\mathscr{S}^2}^2\Big), \quad (3.2.7)
\end{aligned}$$

for a constant $\theta_3 = \theta_3(C, K)$ depending only on $C$ and $K$. The fourth term is estimated by using the same argument as for (3.2.6), obtaining that

$$E\Big[\int_0^T |\mathcal{Y}(X^1,V^1)_s - \mathcal{Y}(X^2,V^2)_s|^2 ds\Big] \le \theta_4 T\Big(\|V^1 - V^2\|_{\mathscr{S}^2}^2 + \|X^1 - X^2\|_{\mathscr{S}^2}^2\Big), \quad (3.2.8)$$

for some constant $\theta_4 = \theta_4(C)$. For the fifth term, we have that

$$\begin{aligned}
&E\Big[\int_0^T |\mathcal{L}^2(\mathcal{M}(X^1,V^1))_s - \mathcal{L}^2(\mathcal{M}(X^2,V^2))_s|^2 ds\Big] \\
&\qquad\qquad\le T\|\mathcal{L}^2(\mathcal{M}(X^1,V^1)) - \mathcal{L}^2(\mathcal{M}(X^2,V^2))\|_{\mathscr{S}^2}^2 \\
&\qquad\qquad\le K^2 T\|\mathcal{M}(X^1,V^1) - \mathcal{M}(X^2,V^2)\|_{\mathscr{S}^2}^2,
\end{aligned}$$

and hence, by Lemma 3.2.1,

$$E\Big[\int_0^T |\mathcal{L}^2(\mathcal{M}(X^1,V^1))_s - \mathcal{L}^2(\mathcal{M}(X^2,V^2))_s|^2 ds\Big] \\ \leq \theta_5 T\Big(\|V^1-V^2\|^2_{\mathscr{S}^2} + \|X^1-X^2\|^2_{\mathscr{S}^2}\Big), \quad (3.2.9)$$

for $\theta_5 = \theta_5(K)$. It only remains to estimate the last term: for notational simplicity, we introduce the process $A^i := \Big(\mathcal{Y}(X^i,V^i), \mathcal{L}^2(\mathcal{M}(X^i,V^i))\Big)$ for $i=1,2$. By applying the Cauchy-Schwarz inequality, we can verify that

$$E\Big[\Big(\int_0^T |\widetilde{X}^1_s - \widetilde{X}^2_s|^2\big(|\widetilde{X}^1_s - \widetilde{X}^2_s|^2 + |A^1_s - A^2_s|^2\big)ds\Big)^{1/2}\Big] \\ \leq E\Big[\Big(\sup_{t\in[0,T]} |\widetilde{X}^1_t - \widetilde{X}^2_t|^2\Big)^{1/2}\Big(\int_0^T \big(|\widetilde{X}^1_s - \widetilde{X}^2_s|^2 + |A^1_s - A^2_s|^2\big)ds\Big)^{1/2}\Big] \\ \leq \frac{\sqrt{T}}{2} E\Big[\sup_{t\in[0,T]} |\widetilde{X}^1_t - \widetilde{X}^2_t|^2\Big] + \frac{1}{2\sqrt{T}} E\Big[\int_0^T \big(|\widetilde{X}^1_s - \widetilde{X}^2_s|^2 + |A^1_s - A^2_s|^2\big)ds\Big],$$

and it is not difficult to check that

$$E\Big[\int_0^T \big(|\widetilde{X}^1_s - \widetilde{X}^2_s|^2 + |A^1_s - A^2_s|^2\big)ds\Big] \leq T\Big(\|\widetilde{X}^1 - \widetilde{X}^2\|^2_{\mathscr{S}^2} + \|A^1 - A^2\|^2_{\mathscr{S}^2}\Big),$$

which finally leads us to

$$E\Big[\Big(\int_0^T |\widetilde{X}^1_s - \widetilde{X}^2_s|^2\big(|\widetilde{X}^1_s - \widetilde{X}^2_s|^2 + |A^1_s - A^2_s|^2\big)ds\Big)^{1/2}\Big] \\ \leq \theta_6\sqrt{T}\Big(\|\widetilde{X}^1 - \widetilde{X}^2\|^2_{\mathscr{S}^2} + \|V^1-V^2\|^2_{\mathscr{S}^2} + \|X^1-X^2\|^2_{\mathscr{S}^2}\Big), \quad (3.2.10)$$

for some constant $\theta_6 = \theta_6(C)$. Therefore, we can plug the estimates (3.2.5)–(3.2.10) back into (3.2.4), obtaining the inequality

$$\|\widetilde{X}^1 - \widetilde{X}^2\|^2_{\mathscr{S}^2} \leq \theta_7\sqrt{T}\Big(\|\widetilde{X}^1 - \widetilde{X}^2\|^2_{\mathscr{S}^2} + \|V^1-V^2\|^2_{\mathscr{S}^2} + \|X^1-X^2\|^2_{\mathscr{S}^2}\Big), \quad (3.2.11)$$

for some constant $\theta_7 = \theta_7(C,K)$. We should now consider the difference $\widetilde{V}^1 - \widetilde{V}^2$. Thanks to the explicit nature of the functional differential equation

(3.2.3), we can easily derive the following estimate:

$$\begin{aligned}
\|\widetilde{V}^1 - \widetilde{V}^2\|^2_{\mathscr{S}^2} &\le E\Bigg[\bigg(\int_0^T \Big| f\big(s, \widetilde{X}^1_s, \mathcal{Y}(X^1,V^1)_s, \mathcal{L}^3(\mathcal{M}(X^1,V^1))_s\big) \\
&\qquad - f\big(s, \widetilde{X}^2_s, \mathcal{Y}(X^2,V^2)_s, \mathcal{L}^3(\mathcal{M}(X^2,V^2))_s\big)\Big| ds\bigg)^2\Bigg] \\
&\le TE\Bigg[\int_0^T \Big| f\big(s, \widetilde{X}^1_s, \mathcal{Y}(X^1,V^1)_s, \mathcal{L}^3(\mathcal{M}(X^1,V^1))_s\big) \\
&\qquad - f\big(s, \widetilde{X}^2_s, \mathcal{Y}(X^2,V^2)_s, \mathcal{L}^3(\mathcal{M}(X^2,V^2))_s\big)\Big|^2 ds\Bigg] \\
&\le 3C^2T\Bigg(E\bigg[\int_0^T |\widetilde{X}^1_s - \widetilde{X}^2_s|^2 ds\bigg] + E\bigg[\int_0^T |\mathcal{Y}(X^1,V^1)_s - \mathcal{Y}(X^2,V^2)_s|^2 ds\bigg] \\
&\quad + E\bigg[\int_0^T |\mathcal{L}^3(\mathcal{M}(X^1,V^1))_s - \mathcal{L}^3(\mathcal{M}(X^2,V^2))_s|^2 ds\bigg]\Bigg) \\
&\le 3C^2T\Big(\|\widetilde{X}^1 - \widetilde{X}^2\|^2_{\mathscr{S}^2} + \|\mathcal{Y}(X^1,V^1) - \mathcal{Y}(X^2,V^2)\|^2_{\mathscr{S}^2} \\
&\qquad + \|\mathcal{L}^3(\mathcal{M}(X^1,V^1)) - \mathcal{L}^3(\mathcal{M}(X^2,V^2))\|^2_{\mathscr{O}^2}\Big).
\end{aligned}$$

By using the same arguments as for the estimates (3.2.5)–(3.2.10) this yields that, for a constant $\theta_8 = \theta_8(C,K)$,

$$\|\widetilde{V}^1 - \widetilde{V}^2\|^2_{\mathscr{S}^2} \le \theta_8\sqrt{T}\Big(\|\widetilde{X}^1 - \widetilde{X}^2\|^2_{\mathscr{S}^2} + \|V^1 - V^2\|^2_{\mathscr{S}^2} + \|X^1_t - X^2_t\|^2_{\mathscr{S}^2}\Big). \quad (3.2.12)$$

Hence, we can sum the inequalities (3.2.11) and (3.2.12), obtaining a constant $\theta_9 = \theta_9(C,K)$ depending only on $C$ and $K$ such that

$$\begin{aligned}
&\|(\widetilde{X}^1, \widetilde{V}^1) - (\widetilde{X}^2, \widetilde{V}^2)\|^2_{\mathscr{S}^2_X \times \mathscr{S}^2_V} \le \theta_9\sqrt{T}\Big(\|\widetilde{X}^1 - \widetilde{X}^2\|^2_{\mathscr{S}^2} + \|V^1 - V^2\|^2_{\mathscr{S}^2} + \|X^1 - X^2\|^2_{\mathscr{S}^2}\Big) \\
&\le \theta_9\sqrt{T}\Big(\|(\widetilde{X}^1, \widetilde{V}^1) - (\widetilde{X}^2, \widetilde{V}^2)\|^2_{\mathscr{S}^2_X \times \mathscr{S}^2_V} + \|(X^1, V^1) - (X^2, V^2)\|^2_{\mathscr{S}^2_X \times \mathscr{S}^2_V}\Big),
\end{aligned}$$

which implies that, for $T > 0$ such that $\theta_9\sqrt{T} < 1/2$,

$$\|(\widetilde{X}^1, \widetilde{V}^1) - (\widetilde{X}^2, \widetilde{V}^2)\|^2_{\mathscr{S}^2_X \times \mathscr{S}^2_V} \le \underbrace{\left(\frac{1}{\theta_9\sqrt{T}} - 1\right)^{-1}}_{<1} \|(X^1, V^1) - (X^2, V^2)\|^2_{\mathscr{S}^2_X \times \mathscr{S}^2_V}.$$

This shows that $\mathbb{L}$ is a contraction if $T < \ell(C,K) := \frac{1}{4\theta_3^2(C,K)} \wedge 1$, and thus admits a unique fixed point $(X,V)$. □

It is not difficult to see that Theorem 3.2.3 also holds for functionals $\mathcal{L}^1$ and $\mathcal{L}^3$ of the form $\mathcal{L}^i = (\mathcal{L}^i_1, \cdots, \mathcal{L}^i_{k_i}, \mathcal{L}^i_{k_i+1}, \ldots, \mathcal{L}^i_{l_i})$, $i = 1,3$, where $\mathcal{L}^i_j$ satisfies Assumption **(L1)** with respect to $\mathscr{S}^2\big([0,T],\mathbb{R}^{p^i_j}\big)$ for $1 \le j \le k_i$, and with respect to $\mathscr{H}^2\big([0,T],\mathbb{R}^{p^i_j}\big)$ for $k_i+1 \le j \le l_i$: indeed, it suffices to take the norm on $\mathscr{O}^2\big([0,T],\mathbb{R}^{p^i}\big) := \prod_{j=1}^{k_i} \mathscr{S}^2\big([0,T],\mathbb{R}^{p^i_j}\big) \times \prod_{j=k_i+1}^{l_i} \mathscr{H}^2\big([0,T],\mathbb{R}^{p^i_j}\big)$ given by

$$\|L\|^2_{\mathscr{O}^2([0,T],\mathbb{R}^{p^i})} := \sum_{j=1}^{k_i} \|L\|^2_{\mathscr{S}^2([0,T],\mathbb{R}^{p^i_j})} + \sum_{j=k_i+1}^{l_i} \|L\|^2_{\mathscr{H}^2([0,T],\mathbb{R}^{p^i_j})}.$$

**Remarks 3.2.4.** $(i)$ Without going into more detail, we point out that the flexibility in the choice of the functionals $\mathcal{L}^i$ opens the door to probabilistic interpretations for many classes of integro-partial differential equations, similarly to the well known non-linear Feynman-Kac formula for BSDEs. We do not attempt to investigate this problem here, leaving it for future research.

$(ii)$ The result of Theorem 3.2.3 can also be extended to other types of terminal condition. Indeed, assume that $C([0,T],\mathbb{R}^n)$ denotes the space of all $\mathbb{R}^n$-valued continuous functions on $[0,T]$, and let $\Phi : \Omega \times C([0,T],\mathbb{R}^n) \to \mathbb{R}^d$ satisfy the $L^\infty$-Lipschitz condition

$$|\Phi(\mathbf{x}) - \Phi(\mathbf{x}')| \le C \sup_{t\in[0,T]} |\mathbf{x}_t - \mathbf{x}'_t| \quad P\text{-a.s.} \quad \forall\, \mathbf{x}, \mathbf{x}' \in C([0,T],\mathbb{R}^n),$$

where $C > 0$. Then, the operators

$$\overline{\mathcal{M}}(X,V)_t := E[\Phi(X) + V_T|\mathcal{F}_t], \quad \overline{\mathcal{Y}}(X,V)_t := \overline{\mathcal{M}}(X,V)_t - V_t,$$

satisfy estimates similar to those of Lemma 3.2.1. Therefore, the reader can easily verify that Theorem 3.2.3 remains valid if we substitute the operators $\mathcal{M}$ and $\mathcal{Y}$ in the system (3.1.4) by $\overline{\mathcal{M}}$ and $\overline{\mathcal{Y}}$. Moreover, the same conclusion remains true if $\Phi$ satisfies the $L^1$-Lipschitz condition

$$|\Phi(\mathbf{x}) - \Phi(\mathbf{x}')| \leq C \int_0^T |\mathbf{x}_t - \mathbf{x}'_t| dt \quad P\text{-a.s.} \quad \forall \mathbf{x}, \mathbf{x}' \in C([0,T], \mathbb{R}^n),$$

instead of the above $L^\infty$-Lipschitz condition. Two typical examples are the functionals $\Phi^1(\mathbf{x}) = \sup_{t\in[0,T]} |\mathbf{x}_t|$ and $\Phi^2(\mathbf{x}) = \int_0^T \mathbf{x}_t dt$, which are related to lookback and Asian options.

(*iii*) By applying a change of integration order similar to the one discussed in (3.2.1), it is possible to generalize the system (3.1.4) by substituting the equation for the component $V$ by

$$dV_t = f(t, X_t, \mathcal{Y}(X,V)_t, \mathcal{L}^3(\mathcal{M}(X,V))_t)\alpha_V(dt), \quad V_0 = 0,$$

where, for some $\widetilde{T} > 0$, $\alpha_V$ is a non-random, finite Borel measure on $[0, \widetilde{T}]$ such that $\alpha_V(\{0\}) = 0$.

(*iv*) The extension of the above result to the case where $\mathcal{L}^2$ satisfies **(L1)** with respect to $\mathscr{O}^2([0,T], \mathbb{R}^{p_2}) = \mathscr{H}^2([0,T], \mathbb{R}^{p_2})$ is problematic, as shown by the following counterexample borrowed from the theory of FBSDEs. Assume we have an augmented Brownian filtration, and consider the functional differential equation

$$\begin{cases} dX_t = \mathcal{L}^2(\mathcal{M}^\phi(X,V))_t dW_t, \\ dV_t = 0, \\ X_0 = V_0 = 0, \end{cases}$$

where $\mathcal{L}^2$ is the functional given by Itô's representation, and $\phi(x) := x + W_T$. Then, we obtain that $V \equiv 0$, and the first equation can be rewritten as

$$dX_t = d\mathcal{M}^\phi(X,0)_t, \quad X_0 = 0.$$

Assume it has an adapted solution $X$. Then, we would have that $X_t = E[X_T + W_T | \mathcal{F}_t]$ for all $t \geq 0$, which would lead in particular, for $t = T$, to the contradiction $W_T = 0$.

We end this section by observing that the result of Theorem 3.2.3 can be extended to any initial time $\tau > 0$ without additional difficulties: indeed, consider systems of functional differential equations on $[\tau, T]$ of the form

$$\begin{cases} dX_t = \mu(t, X_t, \mathcal{Y}(X,V)_t, \mathcal{L}^1(\mathcal{M}(X,V))_t)dt + \sigma(t, X_t, \mathcal{Y}(X,V)_t, \mathcal{L}^2(\mathcal{M}(X,V))_t)dW_t, \\ dV_t = f(t, X_t, \mathcal{Y}(X,V)_t, \mathcal{L}^3(\mathcal{M}(X,V))_t)dt, \\ X_\tau = \eta, \quad V_\tau = \zeta, \end{cases}$$

where $\eta, \zeta \in L^2(\mathcal{F}_\tau)$. Then it is possible to prove that, under the same conditions of Theorem 3.2.3 and provided that $T - \tau < \ell$, there is a unique solution in $\mathscr{S}^2([\tau, T], \mathbb{R}^n) \times \mathscr{S}^2([\tau, T], \mathbb{R}^d)$.

## 3.3 Extension to global solutions

The purpose of this section is to extend the results of Theorem 3.2.3 to arbitrarily large time intervals: as discussed in the Introduction, a similar extension has been showed in [49] for general Lipschitz backward dynamics by imposing additional conditions on $\mathcal{L}$. While these assumptions are quite restrictive, the approach of [49] has the advantage of leaving the assumptions on the driver $f$ and the terminal condition $\xi$ unchanged.

The situation is however quite different for coupled forward-backward dynamics: as it is now well known in the theory of classical FBSDEs, an extension of Theorem 3.2.3 to arbitrary intervals is possible only if we impose additional assumptions on the coefficients of the forward-backward system (3.1.4), since the solution could explode for large time horizons (without going into more detail, we refer the reader to classical counterexamples for Markovian FBSDEs and related PDEs which can be found for instance in [51]). Several techniques have been adopted for classical FBSDEs to overcome this difficulty: let us cite for instance the works [1, 40, 50, 60, 61, 71]. However, we prefer to apply another approach which appears to be the most natural in the case of functional differential equations.

First of all, we briefly discuss the intuition. Similarly to [49], the first step consists in dividing the interval $[0, T]$ into a finite number of subintervals

$I_j := [T_{j-1}, T_j]$, $0 = T_0 < \cdots < T_N = T$: then, we solve the system separately on any subinterval, starting from the last one and going backward. There is, however, an important additional difficulty with respect to [49], which clarifies why additional assumptions on the coefficients are needed. For such an approach to work, we need that the length of the subintervals $I_j$, on which the system has to be solvable, can be bounded by below by a constant independent of $j$. We have seen that such a length depends on the Lipschitz constants $C$ and $K$ of the system, and the only potential complication could arise from the terminal condition on each $I_j$: indeed, while the other coefficients $\mu$, $\sigma$, $f$ and $\mathcal{L}^i$ remain the same on all intervals $I_j$, the terminal condition for the subsystem on $I_j$ is given by $\Xi^j = \mathcal{Y}(X^{j+1}, V^{j+1})_{T_j}$, where $(X^{j+1}, V^{j+1})$ is the local solution on $I_{j+1}$. In order to obtain the desired lower bound for the length of all $I_j$, it is therefore sufficient that $\Xi^j = \theta_j(X^{j+1}_{T_j})$, where $\theta_j(\omega, x)$ is for each $j$ Lipschitz continuous in $x$ with respect to some constant $C$ independent of $j$.

In order to obtain such a result, the need to study the interplay between the two components $X$ and $V$ appears to be unavoidable, due to the strong coupling between the two functional differential equations. This is however very difficult without a concrete expression for the functionals $\mathcal{L}^i$, and is especially true in the case of non-local operators: namely, $\mathcal{L}(M)$ could still depend on the whole path of $(M_t)_{t\in[0,T]}$, and the non-locality may easily lead to explosions. Contrarily to the case of Lipschitz backward dynamics, it seems therefore inevitable that the extension to global intervals of the fully coupled system (3.1.4) has to be treated on a case by case basis. We will thus assume for the rest of this section that the filtration is generated by an $m$-dimensional Brownian motion $(W_t)_{t\in[0,T]}$ on $(\Omega, \mathcal{F}, P)$, and that $\mathcal{L}^1 = \mathcal{L}^2 = \mathcal{L}^3 = \mathcal{L}$, where $\mathcal{L}$ is the functional given by Itô's representation theorem. As usual, we will write $\mathcal{Z}(X, V) = \mathcal{L}(\mathcal{M}(X, V))$.

The choice of $\mathcal{L}^i$ allows us to exploit the well known connection between FBSDEs and parabolic PDEs: in the deterministic case, we have that

$$\mathcal{Y}(X, V)_t = \theta(t, X_t),$$

where $\theta$ is the solution of the corresponding PDE. Therefore, the desired uniform property for the terminal conditions $\Xi^j$ can be obtained by proving that $\theta$ is Lipschitz in $x$ uniformly with respect to $t$. This is the approach that has been adopted, for instance, by Delarue [18]. His results can be reformulated in our setting as follows:

**Proposition 3.3.1.** *Assume that the coefficients $\mu$, $\sigma$, $f$ and $\phi$ satisfy Assumption* **(B1)***, that they are deterministic, and that $\sigma$ does not depend on $z_2$. Moreover, let $\mu(t,\cdot,y,z)$, $\sigma(t,\cdot,y)$, $f(t,\cdot,y,z)$ and $\phi(\cdot)$ be bounded for all $t \in [0,T]$, $(y,z) \in \mathbb{R}^d \times \mathbb{R}^{d\times m}$. Finally, assume that $\sigma$ is continuous on its definition set, and that there is a constant $\lambda > 0$ such that, for all $(t,x,y) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^d$,*

$$\langle \zeta, \sigma\sigma^{\mathrm{T}}(t,x,y)\zeta \rangle \geq \lambda |\zeta|^2, \quad \zeta \in \mathbb{R}^n.$$

*Then, for any $T > 0$, the fully coupled system* (3.1.4) *has a unique solution in $\mathscr{S}^2([0,T],\mathbb{R}^n) \times \mathscr{S}^2([0,T],\mathbb{R}^d)$.*

However, the choice of $\mathcal{L}^i$ allows us to extend Theorem 3.2.3 to any time interval even in the case of random coefficients. This has been shown in Zhang's work [74, 75], by deriving uniform Lipschitz estimates with respect to the initial condition of the component $X$. In the following, we reformulate in our setting Zhang's results on uniform Lipschitz continuity, and we briefly expose how these imply the extension of the solution to arbitrary intervals. *For the rest of this section, we will assume that $n = 1$, i.e. the component $X$ is 1-dimensional, and that $\sigma$ does not depend on $z_2$, i.e. $\sigma = \sigma(t,x,y)$.* Moreover, we assume that the coefficients of (3.1.4) satisfy the following condition:

**Assumption (B2):** *The functions $\mu : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^{d\times m} \to \mathbb{R}$, $\sigma : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^m$, $f : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^{d\times m} \to \mathbb{R}^d$ and $\phi : \Omega \times \mathbb{R} \to \mathbb{R}^d$ satisfy Assumption* **(B2)** *if they satisfy Assumption* **(B1)** *with* (B1.2) *replaced by* (B1.2'), *and:*

(B2.1) *$\phi$ is uniformly Lipschitz in $x$ with constant $C'$.*

(B2.2) *There is a constant* $\gamma > 0$ *such that*

$$\Lambda^1_t(y) \leq -\gamma|\Lambda^2_t(y)|,$$

*for all* $y \in \mathbb{R}^d$ *with* $|y| = 1$, *where*

$$\begin{aligned}\Lambda^1_t(y) &:= \sum_{i=1}^d y_i \Big( \operatorname{Tr}\left(\partial_z f^i (\partial_z \mu)^{\mathrm{T}}\right) - y^{\mathrm{T}} \partial_z \mu (\partial_z f^i)^{\mathrm{T}} y + y^{\mathrm{T}} \partial_y \sigma (\partial_z f^i)^{\mathrm{T}} y \Big) \\ &\quad + \partial_x \sigma (\partial_z \mu)^{\mathrm{T}} y + (\partial_y \mu)^{\mathrm{T}} y, \\ \Lambda^2_t(y) &:= |\partial_z \mu|^2 - |(\partial_z \mu)^{\mathrm{T}} y|^2 + 2 y^{\mathrm{T}} \partial_z \mu (\partial_y \sigma)^{\mathrm{T}} y,\end{aligned}$$

*and where we assumed that all corresponding derivatives exist.*

The assumption on $\phi$ is simply needed in order to differentiate between the Lipschitz constant of $\phi$ and those of the other coefficients. Observe that we require stronger regularity conditions than those of Delarue, since the coefficients of the system (3.1.4) are random (however, $\sigma$ is allowed to be degenerate). Under these conditions, by deriving some clever estimates for linear FBSDEs, Zhang obtained the following result:

**Lemma 3.3.2.** *Assume that the coefficients* $\mu$, $\sigma$, $f$ *and* $\phi$ *satisfy Assumption* **(B2)** *with* $\gamma = \frac{1}{C}$, *and let* $\ell$ *denote the constant in Theorem 3.2.3. Let* $T < \ell$, *and for* $x^i \in \mathbb{R}$, $i = 1, 2$, *let* $(X^i, V^i)$ *denote the solution of*

$$\begin{cases} dX^i_t = \mu(t, X^i_t, \mathcal{Y}(X^i, V^i)_t, \mathcal{Z}(X^i, V^i)_t)dt + \sigma(t, X^i_t, \mathcal{Y}(X^i, V^i)_t)dW_t, \\ dV^i_t = f(t, X^i_t, \mathcal{Y}(X^i, V^i)_t, \mathcal{Z}(X^i, V^i)_t)dt, \\ X^i_0 = x^i, \quad V^i_0 = 0. \end{cases}$$

*Then, there is a constant* $\varrho_C$, *depending only on* $C$, *such that*

$$\big|\mathcal{Y}(X^1, V^1)_0 - \mathcal{Y}(X^2, V^2)_0\big| \leq \overline{C}|x^1 - x^2|,$$

*where* $\overline{C} := \sqrt{\big(|C'|^2 + 1\big)e^{\varrho_C T} - 1} > 0.$

The proof of this result can be found in [75]. With the help of Lemma 3.3.2, we can extend Theorem 3.2.3 to arbitrarily large time intervals.

**Theorem 3.3.3.** *Assume that the coefficients $\mu$, $\sigma$, $f$ and $\phi$ satisfy Assumption* **(B2)**. *Then, for any $T > 0$, the fully coupled system* (3.1.4) *has a unique solution $(X, V)$ in $\mathscr{S}^2([0,T], \mathbb{R}^n) \times \mathscr{S}^2([0,T], \mathbb{R}^d)$.*

*Proof.* We assume without loss of generality that $\gamma = \frac{1}{C}$ in Assumption **(B2)**, by changing $C$ or $\gamma$ if necessary. Let $\overline{C}$ denote the constant in Lemma 3.3.2, and let $\ell = \ell(\overline{C})$ be the constant in Theorem 3.2.3. We consider a partition $(T_0, \cdots, T_N)$ of $[0, T]$ such that $0 = T_0 < \cdots < T_N = T$ and $0 < T_i - T_{i-1} < \ell$ for $i = 1, \cdots, N$, and set $I_i := [T_{i-1}, T_i]$. Moreover, to emphasize the dependence of the functionals $\mathcal{Y}$, $\mathcal{M}$ on the terminal time and condition, we introduce a slightly different notation and write, for all $t \in [0, T_i]$,

$$\breve{\mathcal{M}}^{T_i}(\xi, V)_t = E[\xi + V_{T_i} | \mathcal{F}_t],$$
$$\breve{\mathcal{Y}}^{T_i}(\xi, V)_t = E[\xi + V_{T_i} | \mathcal{F}_t] - V_t.$$

$\breve{\mathcal{Z}}^{T_i}(\xi, V)_t$ is then defined via the Itô representation of $(\breve{\mathcal{M}}^{T_i}(\xi, V)_t)_{t \in [0, T_i]}$.

The first step consists in constructing appropriate terminal conditions for all subintervals $I_i$. This is accomplished via a backward procedure. We set $\theta_N := \phi$, $C_N := C'$, and for all $x \in \mathbb{R}$, we consider the following system on $I_N$:

$$\begin{cases} dX_t^{N,x} = \mu(t, X_t^{N,x}, \breve{\mathcal{Y}}^{T_N}(\theta_N(X_{T_N}^{N,x}), V^{N,x})_t, \breve{\mathcal{Z}}^{T_N}(\theta_N(X_{T_N}^{N,x}), V^{N,x})_t)dt \\ \qquad\qquad + \sigma(t, X_t^{N,x}, \breve{\mathcal{Y}}^{T_N}(\theta_N(X_{T_N}^{N,x}), V^{N,x})_t)dW_t, \\ dV_t^{N,x} = f(t, X_t^{N,x}, \breve{\mathcal{Y}}^{T_N}(\theta_N(X_{T_N}^{N,x}), V^{N,x})_t, \breve{\mathcal{Z}}^{T_N}(\theta_N(X_{T_N}^{N,x}), V^{N,x})_t)dt, \\ X_{T_{N-1}}^{N,x} = x, \quad V_{T_{N-1}}^{N,x} = 0. \end{cases}$$

Since $\theta_N$ has Lipschitz constant $C_N \leq \overline{C}$, by Theorem 3.2.3 the system has a unique solution $(X^{N,x}, V^{N,x})$ for all $x \in \mathbb{R}$. We can thus define $\theta_{N-1}$ by

$$\theta_{N-1}(x) := \breve{\mathcal{Y}}^{T_N}(\theta_N(X_{T_N}^{N,x}), V^{N,x})_{T_{N-1}}.$$

It is then easy to check that $\theta_{N-1}(x)$ is $\mathcal{F}_{T_{N-1}}$-measurable for all $x \in \mathbb{R}$. Moreover, by Lemma 3.3.2, $\theta_{N-1}$ is uniformly Lipschitz in $x$ with constant

$$C_{N-1} := \sqrt{\left(|C_N|^2 + 1\right)e^{\varrho_C (T_N - T_{N-1})} - 1}.$$

Since $C_{N-1} \leq \overline{C}$, we can iterate the same argument: for $i = N-1, \cdots, 2$ and for all $x \in \mathbb{R}$, we consider the solution $(X^{i,x}, V^{i,x})$ on $I_i$ of the system

$$\begin{cases} dX_t^{i,x} = \mu(t, X_t^{i,x}, \breve{\mathcal{Y}}^{T_i}(\theta_i(X_{T_i}^{i,x}), V^{i,x})_t, \breve{\mathcal{Z}}^{T_i}(\theta_i(X_{T_i}^{i,x}), V^{i,x})_t)dt \\ \qquad\qquad + \sigma(t, X_t^{i,x}, \breve{\mathcal{Y}}^{T_i}(\theta_i(X_{T_i}^{i,x}), V^{i,x})_t)dW_t, \\ dV_t^{i,x} = f(t, X_t^{i,x}, \breve{\mathcal{Y}}^{T_i}(\theta_i(X_{T_i}^{i,x}), V^{i,x})_t, \breve{\mathcal{Z}}^{T_i}(\theta_i(X_{T_i}^{i,x}), V^{i,x})_t)dt, \\ X_{T_{i-1}}^{i,x} = x, \quad V_{T_{i-1}}^{i,x} = 0, \end{cases}$$

which exists by Theorem 3.2.3. We then define $\theta_{i-1}$ by

$$\theta_{i-1}(x) := \breve{\mathcal{Y}}^{T_i}(\theta_i(X_{T_i}^{i,x}), V^{i,x})_{T_{i-1}}.$$

$\theta_{i-1}(x)$ is thus $\mathcal{F}_{T_{i-1}}$-measurable for all $x \in \mathbb{R}$, and by Lemma 3.3.2, $\theta_{i-1}$ is uniformly Lipschitz in $x$ with constant $C_{i-1} := \sqrt{\left(|C_i|^2 + 1\right)e^{\varrho C(T_i - T_{i-1})} - 1}$. Moreover, we can easily verify by induction that

$$C_{i-1} = \sqrt{\left(|C_N|^2 + 1\right)e^{\varrho C(T_N - T_{i-1})} - 1} \leq \overline{C}.$$

Now that we have derived appropriate terminal conditions $\theta_i$, we can construct the solution on the whole interval $[0, T]$ by a forward procedure. We set $X_{T_0}^0 := x$, $V_{T_0}^0 := 0$. Then, for $i = 1, \cdots, N$, we denote by $(X^i, V^i)$ the solution on $I_i$ of the system

$$\begin{cases} dX_t^{i} = \mu(t, X_t^{i}, \breve{\mathcal{Y}}^{T_i}(\theta_i(X_{T_i}^{i}), V^{i})_t, \breve{\mathcal{Z}}^{T_i}(\theta_i(X_{T_i}^{i}), V^{i})_t)dt \\ \qquad\qquad + \sigma(t, X_t^{i}, \breve{\mathcal{Y}}^{T_i}(\theta_i(X_{T_i}^{i}), V^{i})_t)dW_t, \\ dV_t^{i} = f(t, X_t^{i}, \breve{\mathcal{Y}}^{T_i}(\theta_i(X_{T_i}^{i}), V^{i})_t, \breve{\mathcal{Z}}^{T_i}(\theta_i(X_{T_i}^{i}), V^{i})_t)dt, \\ X_{T_{i-1}}^{i} = X_{T_{i-1}}^{i-1}, \quad V_{T_{i-1}}^{i} = V_{T_{i-1}}^{i-1}. \end{cases}$$

which exists because of Theorem 3.2.3. For $t \in I_i$, we set

$$X_t := X_t^i, \quad V_t := V_t^i.$$

To prove that $(X, V)$ solves (3.1.4) on $[0, T]$, it suffices to check that, for $t \in I_i$,

$$\breve{\mathcal{Y}}^{T_i}(\theta_i(X_{T_i}^i), V^i)_t = \mathcal{Y}(\phi(X_{T_N}), V)_t, \quad \breve{\mathcal{Z}}^{T_i}(\theta_i(X_{T_i}^i), V^i)_t = \mathcal{Z}(\phi(X_{T_N}), V)_t.$$

However, for $i = 1, \cdots, N-1$ and $t \in [0, T_i]$, we have that

$$\begin{aligned}\breve{\mathcal{Y}}^{T_i}(\theta_i(X_{T_i}), V)_t &= \breve{\mathcal{Y}}^{T_i}(\theta_i(X^i_{T_i}), V^i)_t \\ &= E\big[\mathcal{Y}^{T_{i+1}}\big(\theta_{i+1}(X^{i+1}_{T_{i+1}}), V^{i+1}\big)_{T_i} + V^i_{T_i}\big|\mathcal{F}_t\big] - V^i_t \\ &= E\big[\theta_{i+1}(X_{T_{i+1}}) + V_{T_{i+1}} - V_{T_i} + V_{T_i}\big|\mathcal{F}_t\big] - V_t \\ &= \breve{\mathcal{Y}}^{T_{i+1}}(\theta_{i+1}(X_{T_{i+1}}), V)_t\end{aligned}$$

by the construction of $\theta_i$. This gives by induction that $\breve{\mathcal{Y}}^{T_i}(\theta_i(X^i_{T_i}), V^i)_t = \mathcal{Y}(\theta_N(X_{T_N}), V)_t = \mathcal{Y}(\phi(X_{T_N}), V)_t$ on $I_i$ for $i = 1, \cdots, N$. On the other hand, this implies that $\breve{\mathcal{M}}^{T_i}(\theta_i(X^i_{T_i}), V^i)_t = \mathcal{M}(\phi(X_{T_N}), V)_t$ on $I_i$ for all $i$. In other words, $(\breve{\mathcal{M}}^{T_i}(\theta_i(X^i_{T_i}), V^i)_t)_{t\in I_i}$ is the restriction to $I_i$ of the martingale $(\mathcal{M}(\phi(X_{T_N}), V)_t)_{t\in[0,T]}$ and, due to the locality of the operator $\mathcal{Z}$, this gives us that $\breve{\mathcal{Z}}^{T_i}(\theta_i(X^i_{T_i}), V^i)_t = \mathcal{Z}(\phi(X_{T_N}), V)_t$ on $I_i$.

This shows that $(X, V)$ is a solution of (3.1.4) on $[0, T]$, and the proof is concluded by observing that the uniqueness is a consequence of the uniqueness of $(X^i, V^i)$ on $I_i$. □

We end this section by briefly mentioning some interesting results recently derived by Zhang (private communication). Motivated by the well known connection between FBSDEs and PDEs in the deterministic case, Zhang suggests that, for random coefficients, the solution can be extended to any interval by relying on the existence of a random field $\theta$ (called decoupling field) such that $\mathcal{Y}(X, V)_t = \theta(t, X_t)$, and by showing that $\theta$ is uniformly Lipschitz continuous. This can be obtained via the introduction of a backward stochastic Riccati equation of quadratic growth, under much weaker assumptions than **(B2)** (however, all processes have to be one-dimensional). Since the results of Zhang are still work in progress and thus susceptible to changes, we do not go into more details here.

# Chapter 4

# Numerical analysis of functional differential equations

The purpose of this chapter is to introduce an approximation scheme for systems of functional differential equations associated to classical decoupled FBSDEs, obtaining a new interpretation for existing Euler-type schemes for Lipschitz FBSDEs. Our numerical scheme is based on a time discretization combined with a local Picard iteration approach, which is motivated by the contraction results for fully coupled FBSDEs obtained in Chapter 3.

While the numerical simulations of forward SDEs is now well understood (see for instance the book of Kloeden and Platen [45]), the approximation of Lipschitz decoupled FBSDEs has been the subject of several studies in the last decade. One of the main motivations to develop probabilistic numerical schemes for this type of FBSDEs is certainly the non-linear Feynman-Kac representation (introduced by Pardoux and Peng [57–59]), which provides an interpretation of semilinear parabolic PDEs in terms of decoupled FBSDEs.

The works on the simulation of decoupled FBSDEs can be mainly regrouped in two categories. The first class of algorithms is based on the four-step scheme introduced in [50], and concentrates on the solution of the related PDE (see for instance [28, 54]). However, it is well known that the simulation of PDEs via finite-difference and finite-element methods is affected by the so called "curse of dimensionality", and special attention is therefore needed. The second class of methods tackles the decoupled FBSDE directly, and was initiated by Bally [2] and Chevance [15]. Several other approaches have then been developed, but the most notable progress was certainly the work of Zhang [72,73], who derived a new notion of $L^2$-regularity for the martingale integrand process $Z$. His results opened the door to several new types of algorithms, for instance [7,35,73]. We cite in particular the work of Bender and Denk [4], which is based on a global contraction procedure and, as our scheme, avoids the nesting of conditional expectations along the partition, which arises in most other approaches.

The case of fully coupled FBSDEs is, on the other hand, considerably more problematic: namely, the discretized equation is not explicitly solvable, and it is not possible to consider the problem locally. To our knowledge, the only works studying the fully coupled case via a probabilistic approach are those of Delarue and Menozzi [20] and of Bender and Zhang [5].

Finally, we observe that the numerical simulation of quadratic BSDEs has received particular attention in the last couple of years. However, the problem is significantly more difficult to tackle than in the Lipschitz case, mainly due to the lack of a global contraction result for quadratic BSDEs. The delicacy of the problem has been pointed out, for instance, by Cheridito and Stadje [14]. Interesting results have been obtained by Richou [65] and Imkeller et al. [41,42]: however, the main drawback of these approaches resides in their speed of convergence, and an efficiently implementable algorithm is still subject of research.

The chapter is organized as follows. In Section 4.1, we introduce a time discretization and define an appropriate implicit Euler scheme for the approximation of the system of functional differential equations, with the help of discrete versions of the operators $\mathcal{Y}$, $\mathcal{M}$ and $\mathcal{Z}$ considered in Chapter 3. Then, we prove in Section 4.2 that, as the mesh of the partition goes to zero, the solution of our implicit scheme converges to the true solution, and the rate of convergence is the same as for classical Euler schemes (see [45]). Finally, in Section 4.3 we approximate the solution of the implicit Euler scheme via a local Picard iteration procedure, and prove that the rate of convergence remains the same, provided that the number of iterations is sufficiently large.

## 4.1 Implicit Euler scheme for functional differential equations

Let $T > 0$, and assume that $(\Omega, \mathcal{F}, P)$ is a complete probability space endowed with an $m$-dimensional Brownian motion $W = (W_t)_{t\in[0,T]}$ and the corresponding augmented filtration $(\mathcal{F}_t)_{t\in[0,T]}$. In the following, we consider decoupled systems of functional differential equations of the form

$$\begin{cases} dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \\ dV_t = f\big(t, X_t, \mathcal{Y}(X,V)_t, \mathcal{Z}(X,V)_t\big)dt, \\ X_0 = x,\ V_0 = 0, \end{cases} \tag{4.1.1}$$

where $\mathcal{M}$, $\mathcal{Y}$, $\mathcal{Z}$ denote as in the previous chapter the functionals

$$\begin{aligned} \mathcal{M}(X,V)_t &= E\big[\Phi(X_T) + V_T \big| \mathcal{F}_t\big], \\ \mathcal{Y}(X,V)_t &= \mathcal{M}(X,V)_t - V_t, \\ \mathcal{M}(X,V)_t &= \mathcal{M}(X,V)_0 + \int_0^t \mathcal{Z}(X,V)_s dW_s, \quad t \geq 0, \end{aligned}$$

and $\mu : [0,T]\times\mathbb{R}^n \to \mathbb{R}^n$, $\sigma : [0,T]\times\mathbb{R}^n \to \mathbb{R}^{n\times m}$, $f : [0,T]\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^m \to \mathbb{R}$ and $\Phi : \mathbb{R}^n \to \mathbb{R}$ (note in particular that the solution process $V$ is in this case 1-dimensional, and that the equation for $X$ is just a classical forward SDE). As we have seen previously, this system is essentially equivalent to the corresponding decoupled FBSDE, and is therefore closely related to the following semilinear parabolic PDE on $[0,T]\times\mathbb{R}^n$:

$$\begin{cases} u_t + \mathcal{L}u = -f\big(t, x, u, \nabla_x u \cdot \sigma(t,x)\big), \\ u(T,x) = \Phi(x), \end{cases}$$

where $\mathcal{L} := \mu(t,x)\nabla_x + \frac{1}{2}\operatorname{Tr}(\sigma\sigma^T(t,x)H_x)$ , $H_x := \big(\frac{\partial^2}{\partial x^i \partial x^j}\big)_{i,j=1,\cdots,n}$.

To numerically approximate the system (4.1.1), we first introduce an appropriate time discretization. Let $0 = t_0 < t_1 < \cdots < t_N = T$, and consider the partition $\pi := (t_0, \ldots, t_N)$ of $[0,T]$ with mesh $|\pi| := \max_{1\leq i\leq N} |t_i - t_{i-1}|$.

For notational simplicity, we set $I_i := [t_{i-1}, t_i]$, $\Delta t_i := t_i - t_{i-1}$ and $\Delta F_{t_i} := F_{t_i} - F_{t_{i-1}}$ for $1 \le i \le N$ and for any process $F$. We can then introduce an implicit Euler scheme $(X^\pi_{t_i}, V^\pi_{t_i})_{i=0,\cdots,N}$ for our system by setting $X^\pi_0 = x$, $V^\pi_0 = 0$, and for $0 \le i \le N-1$,

$$X^\pi_{t_{i+1}} = X^\pi_{t_i} + \mu(t_i, X^\pi_{t_i})\Delta t_{i+1} + \sigma(t_i, X^\pi_{t_i})\Delta W_{t_{i+1}}, \tag{4.1.2a}$$

$$V^\pi_{t_{i+1}} = V^\pi_{t_i} + f\big(t_i, X^\pi_{t_i}, \mathcal{Y}(X^\pi, V^\pi)_{t_i}, \mathcal{Z}^\pi(X^\pi, V^\pi)_{t_i}\big)\Delta t_{i+1}, \tag{4.1.2b}$$

where $\mathcal{M}(X^\pi, V^\pi)_{t_i} = E\big[\Phi(X^\pi_T) + V^\pi_T \big| \mathcal{F}_{t_i}\big]$, $\mathcal{Y}(X^\pi, V^\pi)_{t_i} = \mathcal{M}(X^\pi, V^\pi)_{t_i} - V_{t_i}$, and

$$\begin{aligned}\mathcal{Z}^\pi(X^\pi, V^\pi)_{t_i} &:= \frac{1}{\Delta t_{i+1}} E\big[\Delta\mathcal{M}(X^\pi, V^\pi)_{t_{i+1}}\Delta W_{t_{i+1}} \big| \mathcal{F}_{t_i}\big] \\ &= \frac{1}{\Delta t_{i+1}} E\big[\big(\Phi(X^\pi_T) + V^\pi_T\big)\Delta W_{t_{i+1}} \big| \mathcal{F}_{t_i}\big].\end{aligned}$$

While the expressions for $\mathcal{M}$ and $\mathcal{Y}$ are inherited directly from their continuous time versions, the operator $\mathcal{Z}^\pi$ is defined via an Euler-type approximation of $\mathcal{Z}$: namely, we have for $0 \le i \le N-1$ that

$$\Delta\mathcal{M}(X^\pi, V^\pi)_{t_{i+1}} = \int_{t_i}^{t_{i+1}} \mathcal{Z}(X^\pi, V^\pi)_s dW_s \approx \mathcal{Z}(X^\pi, V^\pi)_{t_i}\Delta W_{t_{i+1}},$$

and by multiplying by $\Delta W_{t_{i+1}}$ and conditioning with respect to $\mathcal{F}_{t_i}$, we get

$$\mathcal{Z}(X^\pi, V^\pi)_{t_i} \approx \frac{1}{\Delta t_{i+1}} E\big[\Delta\mathcal{M}(X^\pi, V^\pi)_{t_{i+1}}\Delta W_{t_{i+1}} \big| \mathcal{F}_{t_i}\big] = \mathcal{Z}^\pi(X^\pi, V^\pi)_{t_i}.$$

**Remark 4.1.1.** For the rest of this chapter, we will sometimes need to consider local modifications of $\mathcal{M}$, $\mathcal{Y}$ and $\mathcal{Z}^\pi$, while emphasizing their dependence on the interval and on the terminal condition. These are defined, for any $V = (V_{t_i}, V_{t_{i+1}}) \in L^1\big(\mathcal{F}_{t_i}\big) \times L^1\big(\mathcal{F}_{t_{i+1}}\big)$ and $\xi \in L^1\big(\mathcal{F}_{t_{i+1}}\big)$, by

$$\begin{aligned}\widetilde{\mathcal{M}}^{I_i}(\xi, V)_{t_i} &:= E\big[\xi + V_{t_{i+1}} \big| \mathcal{F}_{t_i}\big], \\ \widetilde{\mathcal{Y}}^{I_i}(\xi, V)_{t_i} &:= \widetilde{\mathcal{M}}^{t_{i+1}}(\xi, V)_{t_i} - V_{t_i}, \\ \widetilde{\mathcal{Z}}^{I_i}(\xi, V)_{t_i} &:= \frac{1}{\Delta t_{i+1}} E\big[\big(\xi + V_{t_{i+1}}\big)\Delta W_{t_{i+1}} \big| \mathcal{F}_{t_i}\big].\end{aligned}$$

The first difficulty lies in the solvability of the discretization procedure (4.1.2). Indeed, while the Euler approximation of the classical stochastic differential equation for $X$ does not pose any problem (see for instance [45]), it is a priori not clear if the approximation scheme for $V$ will have a unique solution $V^\pi$, due to the dependence of the driver on the terminal value $V_T^\pi$ (which makes the scheme implicit).

This problem can be solved under essentially the same assumptions as in Chapter 3. As usual, $|\cdot|$ denotes either the Euclidean norm on $\mathbb{R}^k$ or the Hilbert-Schmidt norm on $\mathbb{R}^{k\times l}$. We assume that the coefficients satisfy the following assumption:

**Assumption (C1):** *The functions $\mu$, $\sigma$, $f$ and $\Phi$ satisfy Assumption* **(C1)** *if there exist constants $C_1, C_2 > 0$ such that:*

(C1.1) *$\mu$, $\sigma$ and $f$ are Lipschitz in the space variables with common constant $C_1$, and $\Phi$ is Lipschitz continuous with constant $C_2$.*

(C1.2) $\sup_{t\in[0,T]} \Big(|\mu(t,0)| + |\sigma(t,0)| + |f(t,0,0,0)|\Big) \le C_1, \quad \Phi(0) \le C_2.$

This allows us to prove the following result:

**Theorem 4.1.2.** *Assume that $\mu$, $\sigma$, $f$ and $\Phi$ satisfy Assumption* **(C1)**. *Then, for $|\pi| \le \frac{1}{16C_1^2} \wedge 1$, there is a unique solution $(X_{t_i}^\pi, V_{t_i}^\pi)_{i=0,\cdots,N}$ to the discretization procedure* (4.1.2) *such that*

$$\max_{0\le i\le N} E|X_{t_i}^\pi|^2 < \infty, \quad \max_{0\le i\le N} E|V_{t_i}^\pi|^2 < \infty.$$

*Proof.* Because of the decoupling, the discretization (4.1.2a) of $X$ can be considered separately from that of $V$: however, (4.1.2a) is just an explicit Euler scheme for forward SDEs, and the result for $(X_{t_i}^\pi)_{i=0,\cdots,N}$ then follows by [45]. Once $(X_{t_i}^\pi)_{i=0,\cdots,N}$ is obtained, the existence of $(V_{t_i}^\pi)_{i=0,\cdots,N}$ is derived by applying the same strategy as in the continuous time case: anyway, we prefer to sketch the main arguments since they will be applied later to derive a local iteration scheme.

We start by considering the interval $I_N$ and the discretization problem on $I_N$ given by $V^{I_N}_{t_{N-1}} = 0$,

$$V^{I_N}_{t_N} = f\big(t_{N-1}, X^\pi_{t_{N-1}}, \widetilde{\mathcal{Y}}^{I_N}(\Phi(X^\pi_T), V^{I_N})_{t_{N-1}}, \widetilde{\mathcal{Z}}^{I_N}(\Phi(X^\pi_T), V^{I_N})_{t_{N-1}}\big)\Delta t_N.$$

This can easily be solved by applying a Picard iteration: set $V^{I_N,0}_{t_{N-1}} = V^{I_N,0}_{t_N} = 0$ and, for $p \in \mathbb{N}$, $V^{I_N,p+1}_{t_{N-1}} = 0$ and

$$V^{I_N,p+1}_{t_N} = f\big(t_{N-1}, X^\pi_{t_{N-1}}, \widetilde{\mathcal{Y}}^{I_N}(\Phi(X^\pi_T), V^{I_N,p})_{t_{N-1}}, \widetilde{\mathcal{Z}}^{I_N}(\Phi(X^\pi_T), V^{I_N,p})_{t_{N-1}}\big)\Delta t_N.$$

By applying classical arguments, it is easy to check that this procedure is contractive for $|\Delta t_N| \leq \frac{1}{16C_1^2} \wedge 1$ with respect to the $L^2$-norm (with contraction constant $\frac{1}{2}$). Therefore, the discretization on $I_N$ has a unique solution, denoted by $V^{I_N} = (0, V^{I_N}_{t_N})$.

We then set $\Xi^{N-1} := \widetilde{\mathcal{Y}}^{I_N}(\Phi(X^\pi_T), V^{I_N})_{t_{N-1}}$. By iterating this argument on $I_i$ for $i = N-1, \cdots, 1$, we obtain similarly a unique solution $V^{I_i} = (0, V^{I_i}_{t_i})$ to

$$V^{I_i}_{t_i} = f\big(t_{i-1}, X^\pi_{t_{i-1}}, \widetilde{\mathcal{Y}}^{I_i}(\Xi^i, V^{I_i})_{t_{i-1}}, \widetilde{\mathcal{Z}}^{I_i}(\Xi^i, V^{I_i})_{t_{i-1}}\big)\Delta t_i$$

whenever $|\Delta t_i| \leq \frac{1}{16C_1^2} \wedge 1$, where $\Xi^i$ is defined recursively for $i < N-1$ by

$$\Xi^i := \widetilde{\mathcal{Y}}^{I_{i+1}}(\Xi^{i+1}, V^{I_{i+1}})_{t_i}.$$

In particular, since the length of the contraction interval does not depend on the local terminal condition $\Xi^i$, we can choose a partition $\pi$ with $|\pi| \leq \frac{1}{16C_1^2} \wedge 1$, and we get the local solutions $V^{I_1}, \cdots, V^{I_N}$ recursively on all the subintervals $I_i$ of $\pi$ (starting from $I_N$). Finally, we combine $V^{I_1}, \cdots, V^{I_N}$ by adding them as follows:

$$V^\pi_{t_0} := 0, \quad V^\pi_{t_i} := \sum_{j=1}^{i} V^{I_j}_{t_j}, \quad i = 1, \cdots, N.$$

It remains to check that $V^\pi$ is the solution of (4.1.2b). Because of the linearity of the operator $\widetilde{\mathcal{Y}}^{I_i}$ and the definition of the terminal conditions $\Xi^i$, by iterating conditional expectations similarly to Theorem 3.3.3 we can verify that

$$\widetilde{\mathcal{Y}}^{I_i}(\Xi^i, V^{I_i}) = \widetilde{\mathcal{Y}}^{I_i}(\Xi^i, V^\pi) = \mathcal{Y}(\xi, V^\pi).$$

On the other hand, this implies that, for all $t \in I_i$,

$$\widetilde{\mathcal{M}}^{I_i}(\Xi^i, V^i)_t - \widetilde{\mathcal{M}}^{I_i}(\Xi^i, V^i)_{t_i} = \mathcal{M}(\xi, V^\pi)_t - \mathcal{M}(\xi, V^\pi)_{t_i},$$

due to the relation between $\widetilde{\mathcal{M}}^{I_i}$ and $\widetilde{\mathcal{Y}}^{I_i}$. Therefore, by the locality of $\widetilde{\mathcal{Z}}^{I_i}$, we obtain that $\widetilde{\mathcal{Z}}^{I_i}(\Xi^i, V^i) = \mathcal{Z}^\pi(\xi, V^\pi)$. This concludes the proof. □

## 4.2 Convergence of the Euler discretization

After showing that the discretization procedure (4.1.2) is solvable, the next step consists in proving that $(X^\pi_{t_i}, V^\pi_{t_i})_{i=0,\cdots,N}$ converges to the true solution $(X, V)$ when $|\pi| \to 0$. This will be obtained as a consequence of existing results of Zhang [73] and Bouchard and Touzi [7], recovering the usual convergence rate of $|\pi|^{1/2}$, the best possible one for Euler-types of schemes (see [45]). However, we will need the coefficients to be regular in time. More exactly:

**Assumption (C2):** *The functions $\mu$, $\sigma$, $f$ and $\Phi$ satisfy Assumption* **(C2)** *if they satisfy Assumption* **(C1)**, *and $\mu$, $\sigma$, $f$ are additionally 1/2-Hölder continuous in time with Hölder constant $C_1$.*

To simplify our computations, we extend the approximations $(X^\pi_{t_i})_{i=0,\cdots,N}$, $(V^\pi_{t_i})_{i=0,\cdots,N}$ to continuous time processes $X^\pi = (X^\pi_t)_{t\in[0,T]}$, $V^\pi = (V^\pi_t)_{t\in[0,T]}$, which are naturally induced by setting $X^\pi_0 = x$, $V^\pi_0 = 0$, and

$$\begin{aligned} X^\pi_t &= X^\pi_{t_i} + \mu(t_i, X^\pi_{t_i})(t - t_i) + \sigma(t_i, X^\pi_{t_i})(W_t - W_{t_i}), \\ V^\pi_t &= f\big(t_i, X^\pi_{t_i}, \mathcal{Y}(X^\pi, V^\pi)_{t_i}, \mathcal{Z}^\pi(X^\pi, V^\pi)_{t_i}\big)(t - t_i) \end{aligned} \tag{4.2.1}$$

on $(t_i, t_{i+1}]$ for $i = 0, \cdots, N-1$. We first state without proof a couple of auxiliary results from the literature: the first one is a result on the path regularity of the processes $\mathcal{Y}(X, V)$ and $\mathcal{Z}(X, V)$, which has been first proved by Zhang in his thesis [72] and later published in [73]. We introduce an auxiliary process $\overline{\mathcal{Z}}^\pi$ by defining

$$\overline{\mathcal{Z}}^\pi(X,V)_{t_i} := \frac{1}{\Delta t_{i+1}} E\Big[\langle \mathcal{M}(X,V), W\rangle_{t_i,t_{i+1}} \Big| \mathcal{F}_{t_i}\Big], \quad i = 0,\cdots N-1,$$

for any processes $X$ and $V$, where $\langle\cdot,\cdot\rangle_{t_i,t_{i+1}} := \langle\cdot,\cdot\rangle_{t_{i+1}} - \langle\cdot,\cdot\rangle_{t_i}$. A simple computation allows to verify that $\overline{\mathcal{Z}}^\pi(X^\pi,V^\pi)$ and $\mathcal{Z}^\pi(X^\pi,V^\pi)$ are identical at the partition points, i.e.

$$\overline{\mathcal{Z}}^\pi(X^\pi,V^\pi)_{t_i} = \mathcal{Z}^\pi(X^\pi,V^\pi)_{t_i}, \quad i = 0,\cdots N-1. \tag{4.2.2}$$

With the help of $\overline{\mathcal{Z}}^\pi$, we can then formulate the following $L^2$-regularity result:

**Lemma 4.2.1.** *Under Assumption* **(C2)**, *there is a constant $\theta$ depending only on $C_1$ and $T$ such that*

$$\max_{0\le i\le N-1} \sup_{s\in[t_i,t_{i+1}]} \|\mathcal{Y}(X,V)_s - \mathcal{Y}(X,V)_{t_i}\|_2 + \left(\sum_{i=0}^{N-1}\int_{t_i}^{t_{i+1}} E|\mathcal{Z}(X,V)_s - \overline{\mathcal{Z}}^\pi(X,V)_{t_i}|^2 ds\right)^{1/2} \le \theta\sqrt{|\pi|}.$$

Besides the $L^2$-regularity, we will also apply an approximation result of Bouchard and Touzi [7], which in our setting is reformulated as follows:

**Lemma 4.2.2.** *Assume that Assumption* **(C2)** *holds. Then there is a constant $\theta$ depending only on $C_1$ and $T$ such that, for $|\pi|$ small enough,*

$$\sup_{s\in[0,T]} \|\mathcal{Y}(X,V)_s - \mathcal{Y}(X^\pi,V^\pi)_s\|_2 + \sqrt{E\big[\langle \mathcal{M}(X,V) - \mathcal{M}(X^\pi,V^\pi)\rangle_{0,T}\big]} \le \theta\sqrt{|\pi|}.$$

We can now derive the desired convergence of our implicit Euler scheme to the exact solution, with the same rate as the classical Euler scheme:

**Theorem 4.2.3.** *Assume that $\mu$, $\sigma$, $f$ and $\Phi$ satisfy Assumption* **(C2)**. *Then, there is a constant $\theta$ depending only on $C_1$ and $T$ such that*

$$\sup_{t\in[0,T]} \|X_t^\pi - X_t\|_2 + \|V_t^\pi - V_t\|_2 \le \theta\sqrt{|\pi|},$$

*whenever $|\pi|$ is sufficiently small.*

*Proof.* We concentrate our efforts on $\sup_{t\in[0,T]} \|V_t^\pi - V_t\|_2$, since the result for $\sup_{t\in[0,T]} \|X_t^\pi - X_t\|_2$ is standard in the literature (see [45]). Fix $i \in \{0, \cdots, N-1\}$ and $t \in (t_i, t_{i+1}]$. Then, we have by Itô's lemma that

$$d|V_t - V_t^\pi|^2 = 2(V_t - V_t^\pi)\Big(f\big(t, X_t, \mathcal{Y}(X,V)_t, \mathcal{Z}(X,V)_t\big) \\ - f\big(t_i, X_{t_i}^\pi, \mathcal{Y}(X^\pi, V^\pi)_{t_i}, \mathcal{Z}^\pi(X^\pi, V^\pi)_{t_i}\big)\Big)dt.$$

By the Cauchy-Schwarz inequality and Assumption **(C2)**, we obtain that

$$E|V_t - V_t^\pi|^2 - E|V_{t_i} - V_{t_i}^\pi|^2 \le 2\int_{t_i}^t E\Big[|V_s - V_s^\pi| \cdot \Big|f\big(s, X_s, \mathcal{Y}(X,V)_s, \mathcal{Z}(X,V)_s\big) \\ - f\big(t_i, X_{t_i}^\pi, \mathcal{Y}(X^\pi, V^\pi)_{t_i}, \mathcal{Z}^\pi(X^\pi, V^\pi)_{t_i}\big)\Big|\Big]ds \\ \le \int_{t_i}^t E|V_s - V_s^\pi|^2 ds + C_1\Bigg(\int_{t_i}^t \big(|s - t_i| + E|X_s - X_{t_i}^\pi|^2 \\ + E|\mathcal{Y}(X,V)_s - \mathcal{Y}(X^\pi, V^\pi)_{t_i}|^2\big)ds \\ + \int_{t_i}^t E|\mathcal{Z}(X,V)_s - \mathcal{Z}^\pi(X^\pi, V^\pi)_{t_i}|^2 ds\Bigg). \tag{4.2.3}$$

We estimate the last two terms: by Lemma 4.2.1 and 4.2.2, we have

$$E|\mathcal{Y}(X,V)_s - \mathcal{Y}(X^\pi, V^\pi)_{t_i}|^2 \\ \le 2\Big(E|\mathcal{Y}(X,V)_s - \mathcal{Y}(X,V)_{t_i}|^2 + E|\mathcal{Y}(X,V)_{t_i} - \mathcal{Y}(X^\pi, V^\pi)_{t_i}|^2\Big) \le \theta_1|\pi|$$

for some constant $\theta_1$. This gives us that

$$\int_{t_i}^t \big(|s - t_i| + E|X_s - X_{t_i}^\pi|^2 + E|\mathcal{Y}(X,V)_s - \mathcal{Y}(X^\pi, V^\pi)_{t_i}|^2\big)ds \le \theta_1|\pi|\Delta t_{i+1}.$$

On the other hand, by applying (4.2.2) we get that

$$E|\mathcal{Z}(X,V)_s - \mathcal{Z}^\pi(X^\pi, V^\pi)_{t_i}|^2 \\ \le 2\Big(E|\mathcal{Z}(X,V)_s - \overline{\mathcal{Z}}^\pi(X,V)_{t_i}|^2 + E|\overline{\mathcal{Z}}^\pi(X,V)_{t_i} - \mathcal{Z}^\pi(X^\pi, V^\pi)_{t_i}|^2\Big) \\ \le 2\Big(E|\mathcal{Z}(X,V)_s - \overline{\mathcal{Z}}^\pi(X,V)_{t_i}|^2 + E|\overline{\mathcal{Z}}^\pi(X,V)_{t_i} - \overline{\mathcal{Z}}^\pi(X^\pi, V^\pi)_{t_i}|^2\Big),$$

and the second term can be further estimated as follows:

$$
\begin{aligned}
E|\overline{\mathcal{Z}}^{\pi}(X,V)_{t_i} - \overline{\mathcal{Z}}^{\pi}(X^{\pi},V^{\pi})_{t_i}|^2 \\
&= \frac{1}{\Delta t_{i+1}^2} E\Big|E\Big[\langle \mathcal{M}(X,V) - \mathcal{M}(X^{\pi},V^{\pi}), W\rangle_{t_i,t_{i+1}}\Big|\mathcal{F}_{t_i}\Big]\Big|^2 \\
&\leq \frac{1}{\Delta t_{i+1}^2} E\Big|\langle \mathcal{M}(X,V) - \mathcal{M}(X^{\pi},V^{\pi}), W\rangle_{t_i,t_{i+1}}\Big|^2 \\
&\leq \frac{1}{\Delta t_{i+1}} E\big[\langle \mathcal{M}(X,V) - \mathcal{M}(X^{\pi},V^{\pi})\rangle_{t_i,t_{i+1}}\big],
\end{aligned}
$$

where the last inequality is a consequence of Kunita-Watanabe's inequality. By applying these estimates to (4.2.3), we can find a constant $\theta_2$ such that

$$
\begin{aligned}
E|V_t - V_t^{\pi}|^2 \leq \int_{t_i}^{t} E|V_s - V_s^{\pi}|^2 ds + \Big[E|V_{t_i} - V_{t_i}^{\pi}|^2 + \theta_2\Big(|\pi|\Delta t_{i+1} \\
+ \int_{t_i}^{t_{i+1}} E|\mathcal{Z}(X,V)_s - \overline{\mathcal{Z}}^{\pi}(X,V)_{t_i}|^2 ds + E\big[\langle \mathcal{M}(X,V) - \mathcal{M}(X^{\pi},V^{\pi})\rangle_{t_i,t_{i+1}}\big]\Big)\Big]
\end{aligned}
$$

Thus, we can apply Gronwall's inequality, obtaining that

$$
\begin{aligned}
E|V_t - V_t^{\pi}|^2 \leq e^{t-t_i}\Big[E|V_{t_i} - V_{t_i}^{\pi}|^2 + \theta_2\Big(|\pi|\Delta t_{i+1} \\
+ \int_{t_i}^{t_{i+1}} E|\mathcal{Z}(X,V)_s - \overline{\mathcal{Z}}^{\pi}(X,V)_{t_i}|^2 ds + E\big[\langle \mathcal{M}(X,V) - \mathcal{M}(X^{\pi},V^{\pi})\rangle_{t_i,t_{i+1}}\big]\Big)\Big]
\end{aligned}
\tag{4.2.4}
$$

Now, by taking $k \in \{1, \cdots, N\}$ and iterating this inequality with $t = t_{i+1}$ for $i = k-1, \cdots, 0$, we get that

$$
\begin{aligned}
E|V_{t_k} - V_{t_k}^{\pi}|^2 &\leq \theta_2 e^{\sum_{i=0}^{k-1}\Delta t_{i+1}}\Big(|\pi|\sum_{i=0}^{k-1}\Delta t_{i+1} + \sum_{i=0}^{k-1}\int_{t_i}^{t_{i+1}} E|\mathcal{Z}(X,V)_s - \overline{\mathcal{Z}}^{\pi}(X,V)_{t_i}|^2 ds \\
&\quad + E\big[\langle \mathcal{M}(X,V) - \mathcal{M}(X^{\pi},V^{\pi})\rangle_{0,t_k}\big]\Big) \\
&\leq \theta_2 e^{T}\Big(T|\pi| + \sum_{i=0}^{N-1}\int_{t_i}^{t_{i+1}} E|\mathcal{Z}(X,V)_s - \overline{\mathcal{Z}}^{\pi}(X,V)_{t_i}|^2 ds \\
&\quad + E\big[\langle \mathcal{M}(X,V) - \mathcal{M}(X^{\pi},V^{\pi})\rangle_{0,T}\big]\Big).
\end{aligned}
$$

Finally, Lemma 4.2.1 and 4.2.2 imply that

$$E|V_{t_k} - V^\pi_{t_k}|^2 \le \theta_3 |\pi|$$

for some constant $\theta_3$. It remains to check that the estimate also holds for $t \neq t_k$. However, by taking (4.2.4) and applying the same estimates, we have that, for any $t \in [t_i, t_{i+1}]$,

$$\begin{aligned} E|V_t - V^\pi_t|^2 \le e^{|\pi|} \Big[ E|V_{t_i} - V^\pi_{t_i}|^2 + \theta_2 \Big( |\pi|^2 + \int_{t_i}^{t_{i+1}} E|\mathcal{Z}(X,V)_s - \overline{\mathcal{Z}}^\pi(X,V)_{t_i}|^2 ds \\ + E\big[\langle \mathcal{M}(X,V) - \mathcal{M}(X^\pi, V^\pi)\rangle_{t_i,t_{i+1}}\big] \Big) \Big] \\ \le \theta_4 |\pi| e^{|\pi|} \big(1 + |\pi|\big) \le \theta_5 |\pi|, \end{aligned}$$

for some constants $\theta_4$ and $\theta_5$ depending only on $C_1$ and $T$, provided that $|\pi|$ is small enough. Since $\theta_5$ is independent of $i$, this proves the claim. □

## 4.3 Local Picard iteration

As already anticipated in the first section, while Theorem 4.2.3 shows that the implicit Euler scheme (4.1.2) converges to the solution $(X, V)$ of the functional differential equation with the best possible rate, this is still not sufficient to obtain an implementable algorithm, since the discretization procedure (4.1.2b) is not explicitly solvable.

In this section, we propose a local Picard iteration scheme for the approximation of the solution $V^\pi$ of (4.1.2b): this is based on the strategy for the proof of Theorem 4.1.2, and even includes an estimate of the necessary number of Picard iterations needed. We observe that this approach presents an interesting feature: namely, as first noted by Gobet et al. [35] experimentally, and later studied in depth by Bender and Denk [4], a major drawback of the large majority of numerical schemes for Lipschitz BSDEs [7, 16, 35, 73] is the presence of nested conditional expectations. Indeed, all conditional expectations have to be replaced by some estimators in practical implementations, and the corresponding approximation error grows when $|\pi| \to 0$, since the nesting takes place along $\pi$: thus, higher computational cost are needed to avoid error explosions. On the other hand, by applying approaches based on Picard iterations, the nesting of conditional expectations takes place along the iterations instead that along the partition. As observed by Bender and Denk, this allows to better control the amplification of the error.

In the following, we assume that the solution $(X^\pi_{t_i})_{i=0,\cdots,N}$ of the explicit Euler scheme (4.1.2a) has already been constructed. We start by considering the last interval $I_N$: we set $\breve{\Xi}^N = \Phi(X^\pi_T)$ and $V^{I_N,0}_{t_{N-1}} = V^{I_N,0}_{t_N} = 0$. For $p \in \mathbb{N}$, we then construct the couple $(V^{I_N,p}_{t_{N-1}}, V^{I_N,p}_{t_N})$ iteratively by $V^{I_N,p+1}_{t_{N-1}} = 0$,

$$V^{I_N,p+1}_{t_N} = f\big(t_{N-1}, X^\pi_{t_{N-1}}, \widetilde{\mathfrak{Y}}^{I_N}(\Phi(X^\pi_T), V^{I_N,p})_{t_{N-1}}, \widetilde{\mathcal{Z}}^{I_N}(\Phi(X^\pi_T), V^{I_N,p})_{t_{N-1}}\big)\Delta t_N.$$

The procedure is stopped after a finite number of iterations $p_N$ (yet to be determined): we obtain $(V^{I_N,p_N}_{t_{N-1}}, V^{I_N,p_N}_{t_N}) = (0, V^{I_N,p_N}_{t_N})$ as result, and we set $\breve{\Xi}^{N-1} := \widetilde{\mathfrak{Y}}^{I_N}(\Phi(X^\pi_T), V^{I_N,p_N})_{t_{N-1}}$, which is $\mathcal{F}_{t_{N-1}}$-measurable.

We can then iterate the same procedure on $I_i$ for $i = N-1, \cdots, 1$: for some finite $p_i$ yet to be determined, the couple $(V^{I_i,p_i}_{t_{i-1}}, V^{I_i,p_i}_{t_i}) = (0, V^{I_i,p_i}_{t_i})$ is constructed iteratively by $V^{I_i,0}_{t_{i-1}} = V^{I_i,0}_{t_i} = 0$ and $V^{I_i,p+1}_{t_{i-1}} = 0$,

$$V^{I_i,p+1}_{t_i} = f\big(t_{i-1}, X^\pi_{t_{i-1}}, \widetilde{\mathfrak{Y}}^{I_i}(\breve{\Xi}^i, V^{I_i,p})_{t_{i-1}}, \widetilde{\mathcal{Z}}^{I_i}(\breve{\Xi}^i, V^{I_i,p})_{t_{i-1}}\big)\Delta t_i$$

for $0 \le p < p_i - 1$, where $\breve{\Xi}^i$ is defined recursively for $i < N-1$ by

$$\breve{\Xi}^i := \widetilde{\mathfrak{Y}}^{I_{i+1}}(\breve{\Xi}^{i+1}, V^{I_{i+1},p_{i+1}})_{t_i}.$$

Finally, we construct $V^{\pi,(p_1,\cdots,p_N)}$ by adding $V^{I_1,p_1}, \cdots, V^{I_N,p_N}$ as follows:

$$V^{\pi,(p_1,\cdots,p_N)}_{t_0} := 0, \quad V^{\pi,(p_1,\cdots,p_N)}_{t_i} := \sum_{j=1}^{i} V^{I_j,p_j}_{t_j}, \quad i = 1, \cdots, N.$$

The scope of this section is to prove that $V^{\pi,(p_1,\cdots,p_N)}$ is a good approximation of the solution of the implicit Euler scheme $V^\pi$, provided that the number of iterations is sufficiently large. In the following, we draw on the notation used in the proof of Theorem 4.1.2: $V^{I_i}$ denotes the local process on $I_i$, and

$\Xi^i$ is the corresponding terminal condition. Moreover, for the entire section, we suppose that $f$ satisfies the conditions of Assumption **(C1)**.

Since $V^{\pi,(p_1,\cdots,p_N)}$ and $V^\pi$ are both constructed by adding up the respective localizations, the main step consists in proving that, for $i = 1, \cdots, N$, $V^{I_i,p_i}_{t_i}$ is sufficiently close to $V^{I_i}_{t_i}$ for $p_i$ large enough. Due to the similarity in the construction of $V^{I_i,p_i}_{t_i}$ and $V^{I_i}_{t_i}$, we can prove the following result:

**Proposition 4.3.1.** *Let $|\pi| \leq \frac{1}{16C_1^2} \wedge 1$. Then, for $p_i \in \mathbb{N}$ and $i = 1, \cdots, N$,*

$$\|V^{I_i,p_i}_{t_i} - V^{I_i}_{t_i}\|_2 \leq \|\breve{\Xi}^i - \Xi^i\|_2 + (2C_1\sqrt{|\pi|})^{p_i}\|V^{I_i}_{t_i}\|_2.$$

*Proof.* Let $i \in \{1, \cdots, N\}$ be fixed. Then:

$$\begin{aligned}\|V^{I_i,p_i}_{t_i} - V^{I_i}_{t_i}\|_2 &= \Delta t_i \Big\| f\big(t_{i-1}, X^\pi_{t_{i-1}}, \widetilde{\mathfrak{Y}}^{I_i}(\breve{\Xi}^i, V^{I_i,p_i-1})_{t_{i-1}}, \widetilde{\mathcal{Z}}^{I_i}(\breve{\Xi}^i, V^{I_i,p_i-1})_{t_{i-1}}\big) \\ &\qquad - f\big(t_{i-1}, X^\pi_{t_{i-1}}, \widetilde{\mathfrak{Y}}^{I_i}(\Xi^i, V^{I_i})_{t_{i-1}}, \widetilde{\mathcal{Z}}^{I_i}(\Xi^i, V^{I_i})_{t_{i-1}}\big)\Big\|_2 \\ &\leq C_1\Delta t_i\Big(\|\widetilde{\mathfrak{Y}}^{I_i}(\breve{\Xi}^i, V^{I_i,p_i-1})_{t_{i-1}} - \widetilde{\mathfrak{Y}}^{I_i}(\Xi^i, V^{I_i})_{t_{i-1}}\|_2 \\ &\qquad + \|\widetilde{\mathcal{Z}}^{I_i}(\breve{\Xi}^i, V^{I_i,p_i-1})_{t_{i-1}} - \widetilde{\mathcal{Z}}^{I_i}(\Xi^i, V^{I_i})_{t_{i-1}}\|_2\Big). \end{aligned} \tag{4.3.1}$$

By the triangle and the Jensen inequalities, we have that

$$\|\widetilde{\mathfrak{Y}}^{I_i}(\breve{\Xi}^i, V^{I_i,p_i-1})_{t_{i-1}} - \widetilde{\mathfrak{Y}}^{I_i}(\Xi^i, V^{I_i})_{t_{i-1}}\|_2 \leq \|\breve{\Xi}^i - \Xi^i\|_2 + \|V^{I_i,p_i-1}_{t_i} - V^{I_i}_{t_i}\|_2,$$

while the conditional Cauchy-Schwarz inequality gives that

$$\begin{aligned}&\|\widetilde{\mathcal{Z}}^{I_i}(\breve{\Xi}^i, V^{I_i,p_i-1})_{t_{i-1}} - \widetilde{\mathcal{Z}}^{I_i}(\Xi^i, V^{I_i})_{t_{i-1}}\|_2 \\ &\quad = \frac{1}{\Delta t_i}\Big\| E\Big[\big((\breve{\Xi}^i - \Xi^i) + (V^{I_i,p_i}_{t_i} - V^{I_i}_{t_i})\big)\Delta W_{t_i}\Big|\mathcal{F}_{t_{i-1}}\Big]\Big\|_2 \\ &\quad \leq \frac{1}{\Delta t_i} E\Big[E[|(\breve{\Xi}^i - \Xi^i) + (V^{I_i,p_i}_{t_i} - V^{I_i}_{t_i})|^2|\mathcal{F}_{t_{i-1}}] \cdot E[|\Delta W_{t_i}|^2|\mathcal{F}_{t_{i-1}}]\Big]^{1/2} \\ &\quad \leq \frac{1}{\sqrt{\Delta t_i}}\Big(\|\breve{\Xi}^i - \Xi^i\|_2 + \|V^{I_i,p_i-1}_{t_i} - V^{I_i}_{t_i}\|_2\Big).\end{aligned}$$

We can then plug these estimates back into (4.3.1), and

$$\|V^{I_i,p_i}_{t_i} - V^{I_i}_{t_i}\|_2 \leq 2C_1\sqrt{\Delta t_i}\Big(\|\breve{\Xi}^i - \Xi^i\|_2 + \|V^{I_i,p_i-1}_{t_i} - V^{I_i}_{t_i}\|_2\Big),$$

since $|\pi| \leq 1$. By iterating this inequality, we finally obtain that

$$\|V_{t_i}^{I_i,p_i} - V_{t_i}^{I_i}\|_2 \leq \left(\sum_{j=1}^{p_i}(2C_1\sqrt{|\pi|})^j\right)\|\breve{\Xi}^i - \Xi^i\|_2 + (2C_1\sqrt{|\pi|})^{p_i}\|V_{t_i}^{I_i}\|_2,$$

and the claim follows since $\sum_{j=1}^{p_i}(2C_1\sqrt{|\pi|})^j \leq 1$ by the choice of $\pi$. □

Therefore, it is evident that the desired approximation of $V^{I_i}$ by $V^{I_i,p_i}$ can be obtained by controlling $\|\breve{\Xi}^i - \Xi^i\|_2$ and $\|V_{t_i}^{I_i}\|_2$ appropriately. The first term $\|\breve{\Xi}^i - \Xi^i\|_2$ can be estimated as follows:

**Lemma 4.3.2.** *For $i = 1, \cdots, N-1$, we have that*

$$\|\breve{\Xi}^i - \Xi^i\|_2 \leq \sum_{j=i+1}^{N} \|V_{t_j}^{I_j,p_j} - V_{t_j}^{I_j}\|_2.$$

*Proof.* This is easily shown by backward induction. By the definition of the operator $\widetilde{\mathcal{Y}}^{I_N}$ we have that

$$\begin{aligned}\|\breve{\Xi}^{N-1} - \Xi^{N-1}\|_2 &= \|\widetilde{\mathcal{Y}}^{I_N}(\Phi(X_T^\pi), V^{I_N,p_N})_{t_{N-1}} - \widetilde{\mathcal{Y}}^{I_N}(\Phi(X_T^\pi), V^{I_N})_{t_{N-1}}\|_2 \\ &\leq \|V_{t_N}^{I_N,p_N} - V_{t_N}^{I_N}\|_2.\end{aligned}$$

Assume now that the claim is true for $\|\breve{\Xi}^{i+1} - \Xi^{i+1}\|_2$. Then:

$$\begin{aligned}\|\breve{\Xi}^i - \Xi^i\|_2 &= \|\widetilde{\mathcal{Y}}^{I_{i+1}}(\breve{\Xi}^{i+1}, V^{I_{i+1},p_{i+1}})_{t_i} - \widetilde{\mathcal{Y}}^{I_{i+1}}(\Xi^{i+1}, V^{I_{i+1}})_{t_i}\|_2 \\ &\leq \|\breve{\Xi}^{i+1} - \Xi^{i+1}\|_2 + \|V_{t_{i+1}}^{I_{i+1},p_{i+1}} - V_{t_{i+1}}^{I_{i+1}}\|_2 \leq \sum_{j=i+1}^{N}\|V_{t_j}^{I_j,p_j} - V_{t_j}^{I_j}\|_2.\end{aligned}$$ □

In other words, $\|\breve{\Xi}^i - \Xi^i\|_2$ depends on the errors of all the local approximations already computed. This is a consequence of the fact that we only apply a finite number of iterations $p_i$, and shows that we could have uncontrollable error amplifications if the previous estimates were not sufficiently accurate. We now have to derive an appropriate estimate for $\|V_{t_i}^{I_i}\|_2$: this can be done with the help of the following discrete version of Gronwall's inequality.

**Lemma 4.3.3.** *Assume that we have two constants $a$, $\delta > 0$ and a sequence of positive numbers $(x_i)_{i=0,\cdots,n}$ such that*

$$x_i \le \delta + a \sum_{j=0}^{i-1} x_j, \quad i = 0, \cdots, n.$$

*Then,*

$$x_i \le \delta(1+a)^i, \quad i = 0, \cdots, n.$$

This result can be shown by induction on $i$. Then, we can prove that:

**Proposition 4.3.4.** *Assume that $|\pi| \le \frac{1}{16C_1^2} \wedge 1$. Then there is a constant $\theta$, depending only on $C_1$, such that*

$$\|V_{t_i}^{I_i}\|_2 \le \theta\big(1 + 4C_1\sqrt{|\pi|}\big)^{N-i}, \quad i = 1, \cdots, N.$$

*Proof.* By definition of $V_{t_i}^{I_i}$, we have that

$$\begin{aligned} \|V_{t_i}^{I_i}\|_2 &= \|f\big(t_{i-1}, X_{t_{i-1}}^\pi, \widetilde{\mathcal{Y}}^{I_i}(\Xi^i, V^{I_i})_{t_{i-1}}, \widetilde{\mathcal{Z}}^{I_i}(\Xi^i, V^{I_i})_{t_{i-1}}\big)\Delta t_i\|_2 \\ &\le \Delta t_i \Big( \|f\big(t_{i-1}, X_{t_{i-1}}^\pi, 0, 0\big)\|_2 \\ &\quad + \|f\big(t_{i-1}, X_{t_{i-1}}^\pi, \widetilde{\mathcal{Y}}^{I_i}(\Xi^i, V^{I_i})_{t_{i-1}}, \widetilde{\mathcal{Z}}^{I_i}(\Xi^i, V^{I_i})_{t_{i-1}}\big) - f\big(t_{i-1}, X_{t_{i-1}}^\pi, 0, 0\big)\|_2\Big) \end{aligned} \tag{4.3.2}$$

The first term does not pose any problem: namely, by the assumption on $\mu$ and $\sigma$ as well as classical results on the Euler scheme for SDEs (see [45]), we have that $\max_{1\le j\le N} \|X_{t_j}^\pi\|_2 \le \theta_1(1 + \sqrt{|\pi|})$ for some constant $\theta_1$ depending only on $C_1$. Then, it is easy to compute that

$$\begin{aligned} \|f\big(t_{i-1}, X_{t_{i-1}}^\pi, 0, 0\big)\|_2 &\le C_1 + \|f\big(t_{i-1}, X_{t_{i-1}}^\pi, 0, 0\big) - f(t_{i-1}, 0, 0, 0)\|_2 \\ &\le C_1 + C_1\|X_{t_{i-1}}^\pi\|_2 \le C_1 + C_1\theta_1(1 + \sqrt{|\pi|}) \le \theta_2 \end{aligned}$$

for some constant $\theta_2$, since $|\pi| \le 1$. For the second term, we have that

$$\begin{aligned} \|f\big(t_{i-1}, &X_{t_{i-1}}^\pi, \widetilde{\mathcal{Y}}^{I_i}(\Xi^i, V^{I_i})_{t_{i-1}}, \widetilde{\mathcal{Z}}^{I_i}(\Xi^i, V^{I_i})_{t_{i-1}}\big) - f\big(t_{i-1}, X_{t_{i-1}}^\pi, 0, 0\big)\|_2 \\ &\le C_1\Big(\|\widetilde{\mathcal{Y}}^{I_i}(\Xi^i, V^{I_i})_{t_{i-1}}\|_2 + \|\widetilde{\mathcal{Z}}^{I_i}(\Xi^i, V^{I_i})_{t_{i-1}}\|_2\Big). \end{aligned}$$

On the other hand, we can estimate $\|\widetilde{\mathcal{Y}}^{I_i}(\Xi^i, V^{I_i})_{t_{i-1}}\|_2$ and $\|\widetilde{\mathcal{Z}}^{I_i}(\Xi^i, V^{I_i})_{t_{i-1}}\|_2$ by applying the Jensen and the conditional Cauchy-Schwarz inequalities similarly to Proposition 4.3.1, and hence

$$\begin{aligned}\|f\big(t_{i-1},&X^\pi_{t_{i-1}}, \widetilde{\mathcal{Y}}^{I_i}(\Xi^i, V^{I_i})_{t_{i-1}}, \widetilde{\mathcal{Z}}^{I_i}(\Xi^i, V^{I_i})_{t_{i-1}}\big) - f\big(t_{i-1}, X^\pi_{t_{i-1}}, 0, 0\big)\|_2 \\ &\leq C_1\big(1 + (\Delta t_i)^{-1/2}\big)\big(\|\Xi^i\|_2 + \|V^{I_i}_{T_i}\|_2\big).\end{aligned}$$

By plugging these estimates back into (4.3.2), we obtain that

$$\begin{aligned}\|V^{I_i}_{t_i}\|_2 &\leq \Delta t_i\Big(\theta_2 + C_1\big(1 + (\Delta t_i)^{-1/2}\big)\big(\|\Xi^i\|_2 + \|V^{I_i}_{T_i}\|_2\big)\Big) \\ &\leq \theta_2 \Delta t_i + 2C_1\sqrt{\Delta t_i}\big(\|\Xi^i\|_2 + \|V^{I_i}_{T_i}\|_2\big).\end{aligned}$$

Hence, the fact that $\Delta t_i \leq |\pi| \leq 1$ implies that

$$(1 - 2C_1\sqrt{|\pi|})\|V^{I_i}_{t_i}\|_2 \leq (1 - 2C_1\sqrt{\Delta t_i})\|V^{I_i}_{t_i}\|_2 \leq \theta_2 + 2C_1\sqrt{|\pi|}\|\Xi^i\|_2,$$

and since $|\pi| \leq \frac{1}{16C_1^2}$, this leads us to

$$\|V^{I_i}_{t_i}\|_2 \leq \frac{\theta_2}{1 - 2C_1\sqrt{|\pi|}} + \frac{2C_1\sqrt{|\pi|}}{1 - 2C_1\sqrt{|\pi|}}\|\Xi^i\|_2 \leq 2\theta_2 + 4C_1\sqrt{|\pi|}\,\|\Xi^i\|_2.$$

On the other hand, by the same induction argument as in Lemma 4.3.2 we get that, for $i = 1, \cdots, N-1$,

$$\|\Xi^i\|_2 \leq \|\Xi^N\|_2 + \sum_{j=i+1}^{N} \|V^{I_j}_{t_j}\|_2 \leq \theta_3 + \sum_{j=i+1}^{N} \|V^{I_j}_{t_j}\|_2,$$

for some constant $\theta_3$, and therefore

$$\|V^{I_i}_{t_i}\|_2 \leq \theta_4 + 4C_1\sqrt{|\pi|}\sum_{j=i+1}^{N} \|V^{I_j}_{t_j}\|_2.$$

Since this is true for all $i \in \{1, \cdots, N\}$, we can apply Lemma 4.3.3, and we conclude that

$$\|V_{t_i}^{I_i}\|_2 \leq \theta_4\big(1 + 4C_1\sqrt{|\pi|}\big)^{N-i}. \qquad \square$$

We are finally ready to derive the main result of this section. Namely, by combining all the above conclusions, we can estimate the error $\|V_{t_i}^{I_i,p_i} - V_{t_i}^{I_i}\|_2$ of the local approximations $V^{I_i,p_i}$ when the number of iterations is sufficiently large. More exactly, we have:

**Proposition 4.3.5.** *Assume that $|\pi| \leq \frac{1}{16C_1^2} \wedge 1$, and choose $p_i$ such that*

$$\begin{aligned} p_N &\geq \frac{N\log(\sqrt{|\pi|}) - N\log(1+\sqrt{|\pi|})}{\log(2C_1\sqrt{|\pi|})}, \\ p_i &\geq \frac{(N+1)\log(\sqrt{|\pi|}) - N\log(1+\sqrt{|\pi|}) - (N-i)\log\big(1+4C_1\sqrt{|\pi|}\big)}{\log(2C_1\sqrt{|\pi|})}, \end{aligned} \tag{4.3.3}$$

*for $i = 1, \ldots, N-1$. Then, for some constant $\theta$ depending only on $C_1$,*

$$\|V_{t_i}^{I_i,p_i} - V_{t_i}^{I_i}\|_2 \leq \theta\bigg(\frac{\sqrt{|\pi|}}{1+\sqrt{|\pi|}}\bigg)^i, \quad i = 1, \cdots, N.$$

*Proof.* The result is shown by backward induction on $i$. Since $\Xi^N = \breve{\Xi}^N = \Phi(X_T^\pi)$, it follows by Proposition 4.3.1 and 4.3.4 that

$$\|V_{t_N}^{I_N,p_N} - V_{t_N}^{I_N}\|_2 \leq (2C_1\sqrt{|\pi|})^{p_N}\|V_{t_N}^{I_N}\|_2 \leq \theta(2C_1\sqrt{|\pi|})^{p_N},$$

and the choice of $p_N$ gives the desired estimate. Assume now that, for all $j \geq i+1$ and $p_j$ chosen as in (4.3.3),

$$\|V_{t_j}^{I_j,p_j} - V_{t_j}^{I_j}\|_2 \leq \theta\bigg(\frac{\sqrt{|\pi|}}{1+\sqrt{|\pi|}}\bigg)^j.$$

Then, by applying Lemma 4.3.2 and Proposition 4.3.4 to the estimate of Proposition 4.3.1, we obtain that

$$\|V_{t_i}^{I_i,p_i} - V_{t_i}^{I_i}\|_2 \leq \sum_{j=i+1}^{N} \|V_{t_j}^{I_j,p_j} - V_{t_j}^{I_j}\|_2 + \theta(2C_1\sqrt{|\pi|})^{p_i}\big(1+4C_1\sqrt{|\pi|}\big)^{N-i}.$$

On the other hand, by the choice of $p_i$,

$$(2C_1\sqrt{|\pi|})^{p_i}\big(1+4C_1\sqrt{|\pi|}\big)^{N-i} \le \sqrt{|\pi|}\left(\frac{\sqrt{|\pi|}}{1+\sqrt{|\pi|}}\right)^N,$$

and the claim then follows by the induction assumption, since

$$\begin{aligned}\sum_{j=i+1}^{N} \|V_{t_j}^{I_j,p_j} - V_{t_j}^{I_j}\|_2 &\le \theta \sum_{j=i+1}^{N}\left(\frac{\sqrt{|\pi|}}{1+\sqrt{|\pi|}}\right)^j \\ &\le \theta\left(\frac{\sqrt{|\pi|}}{1+\sqrt{|\pi|}}\right)^i - \theta\sqrt{|\pi|}\left(\frac{\sqrt{|\pi|}}{1+\sqrt{|\pi|}}\right)^N. \qquad \square\end{aligned}$$

**Remarks 4.3.6.** ($i$) We observe that the required accuracy of the local iterations $V_{t_i}^{I_i,p_i}$ is progressively relaxed as $i$ decreases: indeed, the previous approximations $V_{t_j}^{I_j,p_j}$, $j > i$, have to be more precise in order to avoid possible problems of error amplification, because of the result of Lemma 4.3.2 on the difference $\breve{\Xi}^i - \Xi^i$.

($ii$) While Proposition 4.3.5 apparently suggests that $p_i$ have to be chosen quite large, we point out that the estimates (4.3.3) have to be understood only as maximal bounds: when implementing the algorithm in practice, it is more convenient to dynamically choose $p_i$ by stopping the algorithm when the difference between the computed solution $V_{t_i}^{I_i,p_i}$ and the previous iteration is sufficiently small. This works because of the contraction property, and allows to drastically reduce the number of iterations in most practical examples. Such a feature is not found, for instance, in Bender and Denk [4], as their iteration is global and not local in time.

By the construction of $V_{t_i}^{\pi,(p_1,\cdots,p_N)}$ and $V_{t_i}^{\pi}$, Proposition 4.3.5 gives us the following result on global approximation:

**Theorem 4.3.7.** *Assume that* $|\pi| \le \frac{1}{16C_1^2} \wedge 1$*, and choose* $p_i$ *as in Proposition 4.3.5. Then, for some constant* $\theta$ *depending only on* $C_1$*,*

$$\max_{1\le i\le N} \|V_{t_i}^{\pi,(p_1,\cdots,p_N)} - V_{t_i}^{\pi}\|_2 \le \theta\sqrt{|\pi|}.$$

*Proof.* This is a direct consequence of Proposition 4.3.5. Namely:

$$\begin{aligned}
\max_{1\le i\le N} \|V^{\pi,(p_1,\cdots,p_N)}_{t_i} - V^{\pi}_{t_i}\|_2 &= \max_{1\le i\le N} \left\| \sum_{j=1}^{i} (V^{I_j,p_j}_{t_j} - V^{I_j}_{t_j}) \right\|_2 \\
&\le \max_{1\le i\le N} \sum_{j=1}^{i} \left\| V^{I_j,p_j}_{t_j} - V^{I_j}_{t_j} \right\|_2 \\
&\le \theta \sum_{j=1}^{N} \left( \frac{\sqrt{|\pi|}}{1+\sqrt{|\pi|}} \right)^j \le \theta\sqrt{|\pi|}. \qquad \square
\end{aligned}$$

We end this section by combining the conclusions of Theorem 4.2.3 and Theorem 4.3.7, which shows that our local Picard iteration converges to the true solution of the functional differential equation with the same convergence rate of the classical Euler scheme, provided that the number of iterations is sufficiently large.

**Corollary 4.3.8.** *Assume that $\mu$, $\sigma$, $f$ and $\Phi$ satisfy Assumption* **(C2)***, and choose $p_i$ as in Proposition 4.3.5. Then, there is a constant $\theta$ depending only on $C_1$ such that*

$$\max_{1\le i\le N} \|X^{\pi}_{t_i} - X_{t_i}\|_2 + \|V^{\pi,(p_1,\cdots,p_N)}_{t_i} - V_{t_i}\|_2 \le \theta\sqrt{|\pi|},$$

*whenever $|\pi|$ is sufficiently small.*

# Chapter 5

# Predictable projections of conformal stochastic integrals

The purpose of this chapter is to introduce some complexification techniques for stochastic processes, that allow to consider real-valued processes as appropriate projections of corresponding complex-valued, conformal stochastic processes. While this chapter is not related to the functional differential equation approach discussed in the previous chapters, we believe that it provides a useful tool that might have applications, for instance, in the study of forward and backward SDEs. As an application of our complexification techniques, we derive a characterization of Widder's integral representation for Brownian martingales, which is obtained by adapting to the probabilistic setting a classical result for the heat equation [70].

We start the chapter by studying predictable projections in a conformal Brownian setting. Conformal martingales have been first introduced by Getoor and Sharpe [34] to prove the duality between the Hardy space $H^1$ and the space $BMO$ in the martingale setting: conformal martingales later played an important role in the probabilistic study of analytic functions as well as in the derivation of the conformal invariance of Brownian motion (see for instance the survey article [17]).

While stochastic integration with respect to conformal martingales is particularly interesting because of the properties of the complex plane, to our knowledge there has not been any attempt to introduce a notion of projection of such integrals on the real line. As a first step in this direction, we consider the predictable projection on the real component of a conformal Brownian motion. It turns out that such a projection behaves well under integration, and in particular powers of the conformal Brownian motion project onto the corresponding Hermite polynomials. Such a remarkable property stresses once more the importance of Hermite polynomials in stochastic analysis (which is due especially to their close relation with iterated stochastic integrals and the Wiener chaos decomposition, see for instance Nualart [56]), and it motivates the subsequent study of series of Hermite polynomials, allowing us to obtain, in a stochastic setting, interesting connections to analytic functions.

In the second part of the chapter, the techniques derived in the first part are applied to a wide class of Brownian martingales, obtaining a further characterization of Widder's representation. We recall that, by the results of Widder [70], any positive solution of the heat equation can be rewritten in terms of a Laplace-Stieltjes integral with respect to some measure $\mu$, which however remains undetermined. We will show that the quadratic exponential moments of $\mu$ can be characterized by applying our results on series of Hermite polynomials and related power series of conformal Brownian motion. Moreover, we observe that our results hold for any dimension $d \in \mathbb{N}$, and that we also obtain a relation between Widder's representation and a particular class of analytic functions.

The chapter is organized as follows. In Section 5.1, we recall the notion of predictable projections of stochastic processes and show how stochastic integrals with respect to the conformal Brownian motion are projected on the real line. Then, in Section 5.2 we derive $L^p$-convergence properties for series of Hermite polynomials from well known $L^p$-estimates on the Wiener chaos. Section 5.3 is dedicated to the presentation in a purely probabilistic setting of Widder's representation result as well as its extension to $L^1$-bounded martingales. Finally, we derive in Section 5.4 the characterization of the moments of Widder's measure $\mu$, as well as the aforementioned connection to analytic functions.

## 5.1 Predictable projections of stochastic integrals

We begin by introducing some notation. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space, and assume that $X, Y$ are two independent, $d$-dimensional Brownian motions on $(\Omega, \mathcal{F}, P)$. We denote by $Z$ the conformal $d$-dimensional Brownian motion given by $Z = X + i\,Y$. Furthermore, let $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ be the augmented filtration generated by $Z$, and let $\mathbb{F}^X = (\mathcal{F}^X_t)_{t\geq 0}$, $\mathbb{F}^Y = (\mathcal{F}^Y_t)_{t\geq 0}$ denote the filtrations generated by $X$, respectively $Y$, and augmented by the $P$-nullsets from $\mathbb{F}$. Unless otherwise stated, we will always define stochastic integrals with respect to the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$. We denote by $b\mathbb{L}$ the space of all adapted processes with bounded càglàd paths, and by $b\mathcal{P}$ the space of all bounded predictable processes. Moreover, we define as usual

$$\mathscr{H}^2(Z) := \left\{ H : \Omega \times [0,T] \to \mathbb{C}^d \;\middle|\; H \text{ predictable with respect to } \mathcal{P}^{\mathbb{F}}, \text{ and } \|H\|_{\mathscr{H}^2(Z)} := E\left[\int_0^\infty |H_t|^2 dt\right] < \infty \right\},$$

where $|\cdot|$ is the Euclidean norm, and similarly for $\mathscr{H}^2(X)$. Finally, we denote by $\Pi^X$ the orthogonal projection from $\mathscr{H}^2(Z)$ onto the space $\mathscr{H}^2(X)$. We shortly recall the definition of the predictable projection of a measurable process:

**Definition 5.1.1.** *Let $\mathbb{G}$ be a filtration on $(\Omega, \mathcal{F}, P)$, and let $\mathcal{P}^{\mathbb{G}}$ denote the predictable $\sigma$-field with respect to $\mathbb{G}$. Then, for a measurable process $L$ such that $L$ is positive or bounded, there exists a unique process $\widetilde{L}$, measurable with respect to $\mathcal{P}^{\mathbb{G}}$ such that, for every predictable $\mathbb{G}$-stopping time $T$,*

$$E[L_T|\mathcal{G}_{T-}] = \widetilde{L}_T \text{ } P\text{-a.s. on } \{T < \infty\}.$$

*$\widetilde{L}$ is then called the predictable projection of $L$ on $\mathbb{G}$.*

Note that, when such a process $\widetilde{L}$ exists, we have

$$\widetilde{L}_t = E[L_t|\mathcal{G}_{t-}] \quad P\text{-a.s.,} \quad \forall\, t \geq 0.$$

In the following, we will concentrate our attention on the predictable projection on $\mathbb{F}^X$: because of the properties of the Brownian motion $X$, this actually coincides with the optional projection on $\mathbb{F}^X$. To simplify our notation, the predictable projection of $L$ on $\mathbb{F}^X$ will be denoted by $L^{\mathcal{P}^X}$.

We can now prove the first main result of the chapter: the predictable projection on $\mathbb{F}^X$ maps stochastic integrals with respect to $Z$ onto stochastic integrals with respect to $X$. Moreover, we also obtain a relation between the integrand processes.

**Theorem 5.1.2.** *Let $H$ be a process such that $H \in \mathscr{H}^2(Z)$. Then, the predictable projections $\left(\int HdZ\right)^{\mathcal{P}^X}$ exists, and*

$$\left(\int HdZ\right)_t^{\mathcal{P}^X} = \int_0^t \Pi^X(H)\,dX \ P\text{-a.s. for all } t \geq 0.$$

*Proof.* First of all, we observe that both stochastic integrals can be realized on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$, as $X$ and $Z$ are both continuous martingales on it. Moreover, we notice that the existence of $\left(\int HdZ\right)^{\mathcal{P}^X}$ is a consequence of classical results on filtration shrinkage, which can be found for instance in [63].

We first assume that $H \in b\mathbb{L}$. Fix $t \geq 0$, and consider a sequence $(\pi^n)_{n\in\mathbb{N}}$ of partitions of $[0,t]$ such that $|\pi^n| \to 0$. Because of classical convergence results in stochastic analysis (see [63]), we have that

$$\int_0^t H_s dZ_s = \lim_{n\to\infty} \sum_{t_i\in\pi^n} H_{t_i}(Z_{t_{i+1}} - Z_{t_i}) \quad \text{in } \mathscr{H}^2,$$

and hence there is a subsequence $(\pi^{n_k})_{k\in\mathbb{N}}$ so that $\sum_{t_i\in\pi^{n_k}} H_{t_i}(Z_{t_{i+1}} - Z_{t_i})$ converges to $\int_0^t H_s dZ_s$ $P$-a.s. as $k \to \infty$. Since $H$ is bounded, the bounded convergence theorem gives that

$$\begin{aligned}\left(\int HdZ\right)_t^{\mathcal{P}^X} &= E\left[\int_0^t H_s dZ_s \middle| \mathcal{F}_t^X\right] = \lim_{k\to\infty} \sum_{t_i\in\pi^{n_k}} E[H_{t_i}(Z_{t_{i+1}} - Z_{t_i})|\mathcal{F}_t^X] \\ &= \lim_{k\to\infty} \sum_{t_i\in\pi^{n_k}} \left(E[H_{t_i}(X_{t_{i+1}} - X_{t_i})|\mathcal{F}_t^X] + i\,E[H_{t_i}(Y_{t_{i+1}} - Y_{t_i})|\mathcal{F}_t^X]\right).\end{aligned}$$

We compute the first term. Consider for $t \geq 0$ the class

$$\mathcal{C}_t := \{C \in \mathcal{F} \mid \exists\, A \in \mathcal{F}_t^X, B \in \mathcal{F}_t^Y \text{ such that } C = A \cap B\},$$

which is stable under intersection. Because of the independence of $X$ and $Y$, we can compute that, for all $F \in L^1(\mathcal{F})$ and $C = A \cap B \in \mathcal{C}_{t_i}$,

$$\begin{aligned} E\big[E[F|\mathcal{F}_t^X]\mathbb{1}_C\big] &= E\big[E[F|\mathcal{F}_t^X]\mathbb{1}_A\mathbb{1}_B\big] \\ &= E\big[E[F|\mathcal{F}_t^X]\mathbb{1}_A\big]E[\mathbb{1}_B] \\ &= E\big[E[F|\mathcal{F}_{t_i}^X]\mathbb{1}_A\big]E[\mathbb{1}_B] = E\big[E[F|\mathcal{F}_{t_i}^X]\mathbb{1}_C\big]. \end{aligned}$$

Therefore, by the Dynkin class theorem,

$$E\big[E[F|\mathcal{F}_t^X]\big|\mathcal{F}_{t_i}^Z\big] = E[F|\mathcal{F}_{t_i}^X]$$

since $\mathcal{F}_{t_i}^Z = \sigma(\mathcal{C}_{t_i})$. This implies that $E[H_{t_i}|\mathcal{F}_t^X] = E\big[E[H_{t_i}|\mathcal{F}_{t_i}^Z]\big|\mathcal{F}_t^X\big] = E[H_{t_i}|\mathcal{F}_{t_i}^X] = \Pi^X(H)_{t_i}$, and therefore

$$E[H_{t_i}(X_{t_{i+1}} - X_{t_i})|\mathcal{F}_t^X] = E[H_{t_i}|\mathcal{F}_t^X](X_{t_{i+1}} - X_{t_i}) = \Pi^X(H)_{t_i}(X_{t_{i+1}} - X_{t_i}).$$

On the other hand, we have for the second term that

$$\begin{aligned} E[H_{t_i}(Y_{t_{i+1}} - Y_{t_i})|\mathcal{F}_t^X] &= E[E[H_{t_i}(Y_{t_{i+1}} - Y_{t_i})|\mathcal{F}_{t_i}^Y \vee \mathcal{F}_t^X]|\mathcal{F}_t^X] \\ &= E[H_{t_i}E[(Y_{t_{i+1}} - Y_{t_i})|\mathcal{F}_{t_i}^Y \vee \mathcal{F}_t^X]|\mathcal{F}_t^X] = 0, \end{aligned}$$

and we can hence conclude that

$$\Big(\int H dZ\Big)_t^{\mathcal{P}^X} = \lim_{k\to\infty} \sum_{t_i \in \pi^{n_k}} \Pi^X(H)_{t_i}(X_{t_{i+1}} - X_{t_i}) = \int_0^t \Pi^X(H)\, dX,$$

since $\Pi^X(H)$ remains bounded and left continuous by the general theory of stochastic processes. This proves the claim for $H \in b\mathbb{L}$. The result is extended first to $H \in b\mathcal{P}$ and then to $H \in \mathscr{H}^2(Z)$ by applying respectively the bounded and the monotone convergence theorem. As this procedure is fairly standard, the details are left to the reader. □

In particular, if the predictable projection $H^{\mathcal{P}^X}$ exists for $H \in \mathscr{H}^2(Z)$, then $\left(\int HdZ\right)_t^{\mathcal{P}^X} = \int_0^t H^{\mathcal{P}^X} dX$ $P$-a.s. for all $t \geq 0$. We end this section by observing that Theorem 5.1.2 immediately gives us an explicit expression for the predictable projection on $\mathbb{F}^X$ of two important classes of stochastic processes.

**Corollary 5.1.3.** *For any $t \geq 0$, the following assertions hold:*

*(i)* $\left(e^{a \cdot Z_t}\right)^{\mathcal{P}^X} = \mathcal{E}(a \cdot X)_t$ *$P$-a.s. for all $a \in \mathbb{R}^d$ and $t \geq 0$.*

*(ii) Let $\alpha = (\alpha_1, \cdots, \alpha_d) \in \mathbb{N}^d$ denote a multi-index. Then, for $t \geq 0$,*

$$(Z_t^\alpha)^{\mathcal{P}^X} = H_\alpha(t, X_t) \ P\text{-a.s.},$$

*where $z^\alpha := \prod_{i=1}^d z^{\alpha_i}$ and $H_\alpha$ denotes the $d$-dimensional generalized Hermite polynomial of degree $\alpha$, defined by*

$$H_\alpha(t, X_t) := \prod_{i=1}^d H_{\alpha_i}(t, X_t^i).$$

*In other words, the powers of the conformal Brownian motion project onto the corresponding Hermite polynomials.*

# 5.2 Expansions in Hermite polynomials

The result of Corollary 5.1.3 (*ii*) is particularly interesting because of the importance of Hermite polynomials in stochastic analysis, in particular in regards to their connection to iterated stochastic integrals and to the Wiener chaos expansion. Thus, expansions in Hermite polynomials and some of their properties will be examined more in detail in this section.

In the following, we denote by $K_n$ the homogeneous Wiener chaos of degree $n$ generated by $(X_t)_{t\in[0,1]}$. First of all, we recall that the Ornstein-Uhlenbeck semigroup $(T_t)_{t\geq 0}$ is defined, for $t \geq 0$ and $F \in L^2(\sigma(X_1))$ by

$$T_t F := \sum_{n=0}^{\infty} e^{-nt} \mathcal{J}_n F,$$

where $\mathcal{J}_n$ denotes the orthogonal projection on $K_n$. It is well known that the properties of the Ornstein-Uhlenbeck semigroup lead to useful comparison results about the $L^p$-norms on the Wiener chaos. In particular, $(T_t)_{t\geq 0}$ enjoys the following hypercontractivity property:

**Proposition 5.2.1.** *Assume that we have constants* $1 < p < q < \infty$ *and* $t > 0$ *such that*

$$e^t \geq \left(\frac{q-1}{p-1}\right)^{1/2}.$$

*Then we have that, for all* $F \in L^p(\sigma(X_1))$,

$$\|T_t F\|_q \leq \|F\|_p.$$

The result can be found, for instance, in Nualart [56]. It is then possible to derive the following estimate:

**Lemma 5.2.2.** *Let* $V_n$ *be a random variable in* $K_n$. *Then, for* $1 < p < q < \infty$ *we have that*

$$\|V_n\|_q \leq \left(\frac{q-1}{p-1}\right)^{n/2} \|V_n\|_p.$$

*Proof.* It is well known that, by applying the operator $T_t$ to $V_n$, we get that

$$T_t V_n = e^{-nt} V_n.$$

We now choose $t > 0$ such that $e^t = \left(\frac{q-1}{p-1}\right)^{1/2}$. Then, Proposition 5.2.1 implies that

$$\left(\frac{q-1}{p-1}\right)^{-n/2} \|V_n\|_q = e^{-nt}\|V_n\|_q = \|T_t V_n\|_q \leq \|V_n\|_p. \qquad \square$$

From now on, we will write $L^p$ for the space $L^p(\Omega, \mathcal{F}, P)$, $p \geq 1$. Because of the well known fact that the Hermite polynomials $(H_\alpha(t, X_t))_{\alpha \in \mathbb{N}^d}$ form a complete basis of $L^2(\sigma(X_t))$, we will consider in the sequel series associated to the system $(H_\alpha(t, X_t))_{\alpha \in \mathbb{N}^d}$. We recall that, given a countable index set $\Lambda$, a sequence $(x_\alpha)_{\alpha\in\Lambda}$ in a normed vector space $(V, \|\cdot\|)$ is said to be unconditionally convergent (or summable) to $x \in V$ if, for all bijections $\sigma : \mathbb{N} \to \Lambda$, the series $(x_{\sigma(k)})_{k\in\mathbb{N}}$ converges to $x$.

As a consequence of the hypercontractivity, we can now derive the second main result of this chapter: this extends, via the corresponding Hermite series, the explicit expression of Corollary 5.1.3 to a class of $L^p$-martingales.

**Theorem 5.2.3.** *For $p > 1$ and $T > 0$, let $(M_t)_{t\in[0,T]}$ be an $L^p$-martingale such that $M_t$ is $\sigma(X_t)$-measurable for all $t \in [0,T]$ (i.e. $M_t$ is of the form $g(t, X_t)$ for some function $g$). Moreover, define $b_\alpha$, for $\alpha \in \mathbb{N}^d$, by*

$$b_\alpha := \frac{E[M_t H_\alpha(t, X_t)]}{\|H_\alpha(t, X_t)\|_2^2} = \frac{E[M_t H_\alpha(t, X_t)]}{\alpha! t^{|\alpha|}}.$$

*Then, the function $f : \mathbb{C}^d \to \mathbb{C}$, $f(z) := \sum_{\alpha\in\mathbb{N}^d} b_\alpha z^\alpha$, is well defined, analytic of order 2, and can be represented, for $z \in \mathbb{C}^d$ and $t \in [0,T]$, as*

$$f(z) = E\left[M_t \exp\left(\frac{1}{t}\left(z \cdot X_t - \frac{z^T z}{2}\right)\right)\right]. \tag{5.2.1}$$

*Moreover, let $S < (p \vee p^*)^{-1} T$, where $p^* = \frac{p}{p-1}$ is the conjugate exponent of $p$. Then, $(f(Z_s))_{s\in[0,S]}$ is an $L^p$-martingale such that*

$$\big(f(Z_s)\big)^{\mathcal{P}^X} = M_s, \quad s \in [0,S].$$

*Proof.* First of all, we show that the series $\sum_{\alpha\in\mathbb{N}^d} b_\alpha z^\alpha$ is absolutely convergent for all $z \in \mathbb{C}^d$. Since $\prod_{i=1}^d |z_i|^{\alpha_i} \le |z|^{|\alpha|}$, it is sufficient to prove that

$$\sum_{n=0}^\infty \Big(\sum_{|\alpha|=n} |b_\alpha|\Big)|z|^n < \infty.$$

By Hölder's inequality and Lemma 5.2.2, it is easy to verify that

$$\begin{aligned}
\Big(\sum_{|\alpha|=n} |b_\alpha|\Big)^{1/n} &\le \left(\sum_{|\alpha|=n} \frac{\|M_t\|_p \|H_\alpha(t, X_t)\|_{p^*}}{\alpha! t^{|\alpha|}}\right)^{1/n} \\
&\le \left(\|M_t\|_p \sum_{|\alpha|=n} \big((p^*-1) \vee 1\big)^{|\alpha|/2} \frac{\|H_\alpha(t, X_t)\|_2}{\alpha! t^{|\alpha|}}\right)^{1/n} \\
&= \left(\left(\frac{1/(p-1) \vee 1}{t}\right)^{n/2} \|M_t\|_p \sum_{|\alpha|=n} \frac{1}{\sqrt{\alpha!}}\right)^{1/n}.
\end{aligned}$$

Therefore, by the multinomial theorem,

$$\begin{aligned}\Big(\sum_{|\alpha|=n}|b_\alpha|\Big)^{1/n} &\le \sqrt{\frac{1/(p-1)\vee 1}{t}}\Big(\|M_t\|_p\sqrt{\big|\{\alpha\,|\,|\alpha|=n\}\big|}\sqrt{\sum_{|\alpha|=n}\frac{1}{\alpha!}}\Big)^{1/n}\\ &\le \sqrt{\frac{1/(p-1)\vee 1}{t}}\Big(\|M_t\|_p\frac{d^{n/2}}{\sqrt{n!}}\sqrt{\sum_{|\alpha|=n}\binom{n}{\alpha}}\Big)^{1/n}\\ &= d\sqrt{\frac{1/(p-1)\vee 1}{t}}\Big(\|M_t\|_p\frac{1}{\sqrt{n!}}\Big)^{1/n}.\end{aligned}$$

The last term converges for $n\to\infty$ to 0 because of Stirling's approximation. This shows that $f$ is well defined and analytic. We now prove the representation (5.2.1): by applying Corollary 5.1.3 and the dominated convergence theorem, it is easy to check that, for all $t>0$,

$$\begin{aligned}f(z)=\sum_{\alpha\in\mathbb{N}^d}\frac{1}{\alpha!t^{|\alpha|}}E[M_tH_\alpha(t,X_t)]\cdot z^\alpha &= \sum_{\alpha\in\mathbb{N}^d}\frac{1}{\alpha!t^{|\alpha|}}E[M_tZ_t^\alpha]\cdot z^\alpha\\ &= E\Bigg[M_t\sum_{\alpha\in\mathbb{N}^d}\frac{z^\alpha Z_t^\alpha}{\alpha!t^{|\alpha|}}\Bigg].\end{aligned}$$

Therefore, the multinomial theorem gives the representation:

$$\begin{aligned}f(z)=E\Bigg[M_t\sum_{n\in\mathbb{N}}\sum_{|\alpha|=n}\frac{z^\alpha Z_t^\alpha}{\alpha!t^{|\alpha|}}\Bigg] &= E\Bigg[M_t\sum_{n\in\mathbb{N}}\frac{1}{n!t^n}\sum_{|\alpha|=n}\binom{n}{\alpha}z^\alpha Z_t^\alpha\Bigg]\\ &= E\Bigg[M_t\sum_{n\in\mathbb{N}}\frac{(z\cdot Z_t)^n}{n!t^n}\Bigg]\\ &= E\Bigg[M_t\exp\Big(\frac{z\cdot Z_t}{t}\Big)\Bigg]\\ &= E\Bigg[M_t\exp\Big(\frac{z\cdot X_t}{t}\Big)E\Big[\exp\Big(i\,\frac{z\cdot Y_t}{t}\Big)\Big]\Bigg]\\ &= E\Bigg[M_t\exp\Big(\frac{1}{t}\Big(z\cdot X_t-\frac{z^Tz}{2}\Big)\Big)\Bigg].\end{aligned}$$

On the other hand, by applying Hölder's inequality to this representation we can verify that

$$\begin{aligned}
|f(z)| &\le \|M_t\|_p \left\| \exp\left(\frac{1}{t}\left(z \cdot X_t - \frac{z^T z}{2}\right)\right)\right\|_{p^*} \\
&= \|M_t\|_p E\left[\left|\exp\left(\frac{p^*}{t} z \cdot X_t\right)\right|\right]^{1/p^*} \exp\left(-\frac{\operatorname{Re}(z^T z)}{2t}\right) \\
&= K_t \exp\left(\frac{1}{2t}((p^*-1)\operatorname{Re}(z)^2 + \operatorname{Im}(z)^2)\right) \\
&\le K_t \exp\left(\frac{(p^*-1)\vee 1}{2t}|z|^2\right), \qquad (5.2.2)
\end{aligned}$$

where $K_t := \|M_t\|_p$, and hence $f$ is analytic of order 2. It remains to consider the process $(f(Z_s))_{s\in[0,S]}$. Because of the choice of $S$, we get that $f(Z_s) \in L^p$ for all $s \in [0,S]$; namely, by taking some $t \in [0,T]$ such that $t > (p \vee p^*)s$, the estimate (5.2.2) implies that

$$E\Big[|f(Z_s)|^p\Big] \le K_t E\left[\exp\left(\frac{p \vee p^*}{2t}|Z_s|^2\right)\right] < \infty.$$

Then, the fact that $(f(Z_s))_{s\in[0,S]}$ is a martingale such that $\big(f(Z_s)\big)^{\mathcal{P}^X} = M_s$ for $s \in [0,S]$ is obtained by applying Fubini's theorem and Corollary 5.1.3 to the integral representation (5.2.1) for $f$: the details are left to the reader. □

As a consequence, we obtain the following result for martingales on $[0,\infty)$:

**Corollary 5.2.4.** *Let $p > 1$, and assume that $(M_t)_{t\in[0,\infty)}$ is an $L^p$-martingale such that $M_t$ is $\sigma(X_t)$-measurable for all $t \in [0,T]$, and let $f$ as in Theorem 5.2.3. Then, $(f(Z_t))_{t\in[0,\infty)}$ is an $L^p$-martingale such that*

$$\big(f(Z_t)\big)^{\mathcal{P}^X} = M_t, \quad t \in [0,\infty).$$

By analyzing the proof of Theorem 5.2.3, we immediately notice that the result cannot hold for $p = 1$, due to the unboundedness of the Hermite polynomials. It is however possible to relate the convergence in $L^1$ of an Hermite series to that in $L^p$, $p > 1$. First of all, we derive the following estimate with the help of Lemma 5.2.2 and of the well known interpolation of Hölder's inequality.

**Lemma 5.2.5.** *Let $V_n$ is a random variable in $K_n$, and $p > 1$. Then:*

$$\|V_n\|_p \leq e^{np/2}\|V_n\|_1.$$

*Proof.* Let $q > p$ be arbitrary, and let $\theta = \theta(p,q) \in (0,1)$ be such that $\frac{1}{p} = \frac{1-\theta}{q} + \theta$. Then, the interpolation of Hölder's inequality and Lemma 5.2.2 yield that

$$\|V_n\|_p \leq \|V_n\|_q^{1-\theta}\|V_n\|_1^{\theta} \leq \left(\frac{q-1}{p-1}\right)^{n(1-\theta)/2} \|V_n\|_p^{1-\theta}\|V_n\|_1^{\theta}.$$

By rearranging the terms, this gives us that

$$\|V_n\|_p \leq \left(\frac{q-1}{p-1}\right)^{\frac{n(1-\theta)}{2\theta}} \|V_n\|_1.$$

The claim then follows by observing that

$$\inf_{q\in(p,\infty)} \left(\frac{q-1}{p-1}\right)^{\frac{n(1-\theta(p,q))}{2\theta(p,q)}} = \lim_{q\to p^+} \left(\frac{q-1}{p-1}\right)^{\frac{n(1-\theta(p,q))}{2\theta(p,q)}} = e^{np/2}. \qquad \square$$

Even though the constant $e^{np/2}$ could be further optimized, it is sufficiently small for our purpose. Thanks to the above estimate, we obtain that the convergence in $L^1$ of an Hermite series implies, to a certain extent, that in $L^p$ for $p > 1$:

**Proposition 5.2.6.** *Let $t \geq 0$, and assume that the series $\sum_{\alpha\in\mathbb{N}^d} b_\alpha H_\alpha(t, X_t)$ is unconditionally convergent in $L^1$. Then, for $p > 1$, $\sum_{\alpha\in\mathbb{N}^d} b_\alpha H_\alpha(s, X_s)$ converges unconditionally and absolutely in $L^p$ for $s < \frac{t}{d^2 e^p}$.*

*Proof.* We denote by $\mathcal{X}$ the unconditional $L^1$-limit of $\sum_{\alpha\in\mathbb{N}^d} b_\alpha H_\alpha(t, X_t)$, and fix a bijection $\sigma : \mathbb{N} \to \mathbb{N}^d$. Then, for any $\varepsilon > 0$, there is an $N \in \mathbb{N}$ such that

$$\Big\| \sum_{k=0}^{n} b_{\sigma(k)} H_{\sigma(k)}(t, X_t) - \mathcal{X} \Big\|_1 \leq \varepsilon \quad \text{for all } n \geq N.$$

This implies that, for all $k > N$,

$$|b_{\sigma(k)}|\|H_{\sigma(k)}(t,X_t)\|_1 = \Big\| \sum_{j=0}^{k} b_{\sigma(j)} H_{\sigma(j)}(t,X_t) - \sum_{j=0}^{k-1} b_{\sigma(j)} H_{\sigma(j)}(t,X_t) \Big\|_1$$
$$\leq \Big\| \sum_{j=0}^{k} b_{\sigma(j)} H_{\sigma(j)}(t,X_t) - \mathcal{X} \Big\|_1 + \Big\| \sum_{j=0}^{k-1} b_{\sigma(j)} H_{\sigma(j)}(t,X_t) - \mathcal{X} \Big\|_1 \leq 2\varepsilon.$$

On the other hand, since the Hermite polynomial $H_\alpha(1, X_1)$ lies for $\alpha \in \mathbb{N}^d$ in the Wiener chaos $K_{|\alpha|}$, Lemma 5.2.5 implies that, for $s \geq 0$ and $k \in \mathbb{N}$,

$$\|H_{\sigma(k)}(s,X_s)\|_p = (\sqrt{s})^{|\sigma(k)|/2} \|H_{\sigma(k)}(1,X_1)\|_p \leq (e^p s)^{|\sigma(k)|/2} \|H_{\sigma(k)}(1,X_1)\|_1.$$

As a consequence of these two estimates and of the scaling property of Brownian motion we obtain that, for all $m, n > N$ and $s \geq 0$,

$$\begin{aligned} \Big\| \sum_{k=n}^{m} b_{\sigma(k)} H_{\sigma(k)}(s,X_s) \Big\|_p &\leq \sum_{k=n}^{m} |b_{\sigma(k)}| \|H_{\sigma(k)}(s,X_s)\|_p \\ &\leq \sum_{k=n}^{m} 2\varepsilon (e^p s)^{|\sigma(k)|/2} \frac{\|H_\alpha(1,X_1)\|_1}{\|H_\alpha(t,X_t)\|_1} \\ &= 2\varepsilon \sum_{k=n}^{m} \left( e^p \frac{s}{t} \right)^{|\sigma(k)|/2}. \end{aligned} \tag{5.2.3}$$

Consider now the series $\sum_{k=0}^{\infty} \left(e^p \frac{s}{t}\right)^{|\sigma(k)|/2}$, and assume that $s < \frac{t}{d^2 e^p}$. Since all terms are positive, it is easy to verify that

$$\begin{aligned} \sum_{k=0}^{\infty} \left( e^p \frac{s}{t} \right)^{|\sigma(k)|/2} &= \sum_{j=0}^{\infty} \sum_{|\alpha|=j} \left( e^p \frac{s}{t} \right)^{|\alpha|/2} = \sum_{j=0}^{\infty} \left( e^p \frac{s}{t} \right)^{j/2} \big| \{\alpha \in \mathbb{N}^d | \ |\alpha| = j\} \big| \\ &\leq \sum_{j=0}^{\infty} \left( e^{p/2} d \sqrt{\frac{s}{t}} \right)^j < \infty, \end{aligned}$$

and we can therefore find an $N' \geq N$ such that

$$\sum_{k=n}^{m} \left( e^p \frac{s}{t} \right)^{|\sigma(k)|/2} \leq \frac{1}{2} \quad \text{for all } m, n \geq N'.$$

As a consequence of (5.2.3), we finally obtain that, for $s < \frac{t}{d^2 e^p}$ and $m, n \geq N'$,

$$\Big\| \sum_{k=n}^{m} b_{\sigma(k)} H_{\sigma(k)}(s, X_s) \Big\|_p \leq \varepsilon.$$

Thus, the series $\sum_{k=0}^{n} b_{\sigma(k)} H_{\sigma(k)}(s, X_s)$ converges in $L^p$, and the $L^p$-limit has to be equal to the $L^1$-limit $\mathcal{X}$ because of Jensen's inequality. As $\sigma$ was arbitrary, this proves the unconditional convergence. The absolute convergence follows similarly from (5.2.3) by applying the same arguments. □

On the other hand, the assumption that the series converges unconditionally can be further weakened, since we can prove that the conditional convergence of an Hermite series in $L^1$ with respect to a graded order on $\mathbb{N}^d$ implies its unconditional convergence in the following way:

**Lemma 5.2.7.** *Let $(b_\alpha)_{\alpha \in \mathbb{N}^d}$ be a family of coefficients in $\mathbb{R}$. Moreover, set $\gamma_n(t, X_t) := \sum_{|\alpha|=n} b_\alpha H_\alpha(t, X_t)$ for $t \geq 0$, and assume that $\sum_{n \in \mathbb{N}} \gamma_n(t, X_t)$ converges in $L^1$ for some $t$. Then, $\sum_{\alpha \in \mathbb{N}^d} b_\alpha H_\alpha(s, X_s)$ converges unconditionally in $L^1$ for $s < \frac{t}{d^2 e^2}$.*

*Proof.* Let $\varepsilon > 0$ be arbitrary. By computations similar to those of Proposition 5.2.6, we can easily find an $N \in \mathbb{N}$ such that, for all $n \geq N$ and all $\alpha$ with $|\alpha| = n$, $|b_\alpha| \|H_\alpha(t, X_t)\|_1 \leq \varepsilon$. By the orthogonality of Hermite polynomials and Lemma 5.2.5, it is then easy to check that, for all $n \geq N$ and $s \geq 0$,

$$\begin{aligned}
\big\|\gamma_n(s, X_s)\big\|_2^{1/n} &= \Big( \sum_{|\alpha|=n} |b_\alpha| \|H_\alpha(s, X_s)\|_2 \Big)^{1/n} \\
&= \Big( \sum_{|\alpha|=n} \varepsilon e^n s^{n/2} \frac{\|H_\alpha(1, X_1)\|_2}{\|H_\alpha(t, X_t)\|_2} \Big)^{1/n} \\
&\leq \Big( \varepsilon d^n e^n \Big(\frac{s}{t}\Big)^{n/2} \Big)^{1/n} = \varepsilon^{1/n} de \sqrt{\frac{s}{t}}\, .
\end{aligned}$$

By choosing $s < \frac{t}{d^2 e^2}$, we thus get that

$$\lim_{n \to \infty} \big\|\gamma_n(s, X_s)\big\|_2^{1/n} \leq de \sqrt{\frac{s}{t}} \lim_{n \to \infty} \varepsilon^{1/n} = de \sqrt{\frac{s}{t}} < 1.$$

Hence, the series $\sum_{n\in\mathbb{N}} \gamma_n(s, X_s)$ converges absolutely in $L^2$. By the orthogonality of the Hermite polynomials in $L^2$, we can break up the homogeneous terms in $\gamma_n$, obtaining that $\sum_{\alpha\in\mathbb{N}^d} b_\alpha H_\alpha(s, X_s)$ converges unconditionally in $L^2$. This in turn implies the unconditional convergence in $L^1$. □

We conclude this section by observing that the combination of Proposition 5.2.6 and Lemma 5.2.7 leads to the following result:

**Corollary 5.2.8.** *Assume that, for all $t \geq 0$, the series $\sum_{\alpha\in\mathbb{N}^d} b_\alpha H_\alpha(t, X_t)$ converges conditionally in $L^1$ with respect to a graded order on $\mathbb{N}^d$. Then, $\sum_{\alpha\in\mathbb{N}^d} b_\alpha H_\alpha(t, X_t)$ converges unconditionally and absolutely in $L^p$ for all $t \geq 0$ and $p \geq 1$.*

## 5.3 Widder's representation for Brownian martingales

As an application of the above properties, we derive a characterization of Widder's theorem about the representation of positive martingales. First of all, we recall how the classical formulation of Widder translates in a probabilistic setting. A purely probabilistic proof of this theorem can be found in [53], but we prefer to include a different presentation for the convenience of the reader. Moreover, we observe that our result hold for any dimension $d \in \mathbb{N}$.

**Theorem 5.3.1.** *Let $X$ be a $d$-dimensional Brownian motion, and suppose that $M_t = g(t, X_t)$ is a continuous martingale such that $g(0,0) = 1$ and $g(t, X_t) \geq 0$. Then, there exists a probability measure $\mu$ on $\mathbb{R}^d$ such that, for all $t \geq 0$,*

$$M_t = \int_{\mathbb{R}^d} \mathcal{E}(v \cdot X)_t \ \mu(dv) \quad P\text{-a.s.}.$$

*Proof.* Let $f : \mathbb{R}^d \to \mathbb{C}$ be a bounded, continuous map. Then, for $s \geq t$ one has that

$$E\left[M_t f\left(\frac{X_t}{t}\right)\right] = E\left[M_s f\left(\frac{X_t}{t}\right)\right] = E\left[M_s f\left(\frac{X_s}{s} + \left(\frac{X_t}{t} - \frac{X_s}{s}\right)\right)\right].$$

Note that the random variables $\left(\frac{X_t}{t} - \frac{X_s}{s}\right)$ and $\frac{X_s}{s}$ are independent as they are uncorrelated in a Gaussian space. Therefore, by conditioning on $\frac{X_s}{s}$ one gets that

$$E\left[M_t f\left(\frac{X_t}{t}\right)\right] = E\left[M_s \int_{\mathbb{R}^d} f\left(\frac{X_s}{s} + \sqrt{\frac{1}{t} - \frac{1}{s}}x\right) \frac{\exp(-|x|^2/2)}{(2\pi)^{d/2}} dx\right].$$

We can now proceed with the construction of the measure $\mu$. By taking $f(x) := \exp(iu \cdot x)$, the previous equality yields that

$$E\left[M_t \exp\left(iu \cdot \frac{X_t}{t}\right)\right] = E\left[M_s \exp\left(iu \cdot \frac{X_s}{s}\right)\right] \exp\left(-\frac{1}{2}\left(\frac{1}{t} - \frac{1}{s}\right)|u|^2\right).$$

On the other hand, the process $(M_t)_{t\geq 0}$ is by assumption a density process, and is therefore associated to a measure $Q$ on $C(\mathbb{R}^+, \mathbb{R}^d)$ with $Q\big|_{\mathcal{F}_t} << P\big|_{\mathcal{F}_t}$ via the Radon-Nikodym theorem. This gives us that

$$E_Q\left[\exp\left(iu \cdot \frac{X_t}{t}\right)\right] = E_Q\left[\exp\left(iu \cdot \frac{X_s}{s}\right)\right] \exp\left(-\frac{1}{2}\left(\frac{1}{t} - \frac{1}{s}\right)|u|^2\right).$$

By taking $t = 1$, this implies in particular that

$$\varphi^Q_{X_1}(u) = \varphi^Q_{\frac{X_s}{s}}(u) \exp\left(-\frac{1}{2}\left(1 - \frac{1}{s}\right)|u|^2\right),$$

where $\varphi^Q_Y$ denotes the characteristic function of the random variable $Y$ under the measure $Q$.

As a consequence, we get that $\varphi^Q_{\frac{X_s}{s}}(u)$ converges pointwise for $s \to \infty$ to a continuous function $\varphi^Q_{X_1}(u) \exp\left(\frac{1}{2}|u|^2\right)$. Therefore, Lévy's continuity theorem yields the existence of a measure $\mu$ such that $\frac{X_s}{s}$ converges weakly to $\mu$ under $Q$ as $s \to \infty$.

We now have to check that $\mu$ satisfies the desired property. By the above convergence in distribution, we have that, for any $f$ bounded and continuous and for any fixed $t > 0$,

$$E_Q\left[\int_{\mathbb{R}^d} f\left(\frac{X_s}{s} + \sqrt{\frac{1}{t} - \frac{1}{s}}x\right) \frac{e^{-|x|^2/2}}{(2\pi)^{d/2}} dx\right] \to \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f\left(v + \sqrt{\frac{1}{t}}x\right) \frac{e^{-|x|^2/2}}{(2\pi)^{d/2}} dx\ \mu(dv)$$

when $s \to \infty$. Since the left hand side equals $E\left[M_t f\left(\frac{X_t}{t}\right)\right]$ for any $s \geq t$, we get that

$$E\left[M_t f\left(\frac{X_t}{t}\right)\right] = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f\left(v + \sqrt{\frac{1}{t}}x\right)\frac{e^{-|x|^2/2}}{(2\pi)^{d/2}}dx\ \mu(dv).$$

On the other hand, by using Girsanov's theorem one can easily check that, for any $f$ bounded and continuous,

$$\begin{aligned} E\left[\int_{\mathbb{R}^d}\mathcal{E}(vX)_t\ \mu(dv) f\left(\frac{X_t}{t}\right)\right] &= \int_{\mathbb{R}^d} E\left[\mathcal{E}(vX)_t\ f\left(\frac{X_t}{t}\right)\right]\mu(dv) \\ &= \int_{\mathbb{R}^d} E\left[f\left(\frac{X_t}{t} + v\right)\right]\mu(dv) \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f\left(v + \sqrt{\frac{1}{t}}x\right)\frac{e^{-|x|^2/2}}{(2\pi)^{d/2}}dx\ \mu(dv). \end{aligned}$$

By approximation arguments, this equality can then be extended to any $f$ bounded and measurable, obtaining the desired representation for $M_t$. □

The condition that $g(t, X_t)$ forms a martingale on the whole interval $[0,\infty)$ is essential. This can easily be shown by separation arguments: let $d = 1$, $T > 0$, and denote by $\mathcal{K}$ the set of all the martingales having the desired representation, i.e.

$$\mathcal{K} = \Big\{N = (N_t)_{0\leq t\leq T}\ \Big|\ \text{There is a probability measure } \mu \text{ such that } N_t = \int_{\mathbb{R}}\mathcal{E}(vX)_t\ \mu(dv), 0 \leq t \leq T\Big\}.$$

Clearly, $\mathcal{K}$ is a convex set. Now, for an arbitrary $f \in L^\infty(\mathbb{R}^d)$ and any $N \in \mathcal{K}$ we have that

$$\begin{aligned} E\left[f(X_t)N_t\right] &= \int_\Omega\int_{\mathbb{R}} f(X_t)\mathcal{E}(vX)_t\ \mu(dv)dP \\ &= \int_{\mathbb{R}} E\left[f(X_t)\mathcal{E}(vX)_t\right]\mu(dv) \\ &= \int_{\mathbb{R}} E\left[f(X_t + vt)\right]\mu(dv). \end{aligned}$$

The choice $f_k(x) := \sin(kx)$ gives

$$\begin{aligned} E\left[f_k(X_t)N_t\right] &= \int_{\mathbb{R}} E\left[\sin(kX_t + kvt)\right]\mu(dv) \\ &= \int_{\mathbb{R}} E\left[\cos(kX_t)\right]\sin(kvt)\mu(dv) \\ &\geq E\left[\cos(kX_t)\right]\inf_{v\in\mathbb{R}}\sin(kvt) = -e^{-k^2t/2}. \end{aligned}$$

However, by defining

$$M_T = \frac{\mathbb{1}_{\{\sin(kX_T) < -e^{-k^2T/2}\}}}{P(\sin(kX_T) < -e^{-k^2T/2})}$$

and by setting $M_t := E[M_T|\mathcal{F}_t] = g(t, X_t)$, we get a martingale on $[0, T]$ of the desired form and such that $E\left[f_k(X_t)M_t\right] < -e^{-k^2t/2}$. This proves the claim.

On the other hand, Theorem 5.3.1 can easily be extended to any continuous $L^1$-bounded Brownian martingale on $[0, \infty)$. First of all, we recall the well known Krickeberg decomposition for $L^1$-bounded martingales:

**Theorem 5.3.2.** *Let $M$ denote a martingale on $[0, \infty)$. Then $M$ is $L^1$-bounded if and only if it can be written ($P$-a.s.) as the difference of two positive martingales $M^1$, $M^2$. Moreover, one can choose $M^1$, $M^2$ so that*

$$\sup_{t\geq 0}\|M_t\|_1 = E[M_0^1] + E[M_0^2],$$

*and the decomposition is then given by*

$$M_t^1 = \sup_{s\geq t} E[M_s^+|\mathcal{F}_t], \quad M_t^2 = \sup_{s\geq t} E[M_s^-|\mathcal{F}_t] \quad P\text{-a.s.}, \quad t \geq 0.$$

The proof of this well known result can be found for instance in [24]. Krickeberg's decomposition allows us to extend Widder's representation theorem, obtaining the following result:

**Proposition 5.3.3.** *Let $(M_t)_{t\geq 0}$ be a continuous, $L^1$-bounded martingale of the form $M_t = g(t, X_t)$. Then there is a signed measure $\mu$ on $\mathbb{R}^d$ such that*

$$M_t = \int_{\mathbb{R}^d} \mathcal{E}(v\cdot X)_t \; \mu(dv) \quad P\text{-a.s. for all } t \geq 0.$$

*Moreover, we have that*

$$\sup_{t\geq 0} \|M_t\|_1 = \|\mu\|.$$

*Proof.* We apply the Krickeberg decomposition to $M$, obtaining two positive martingales $M^1$, $M^2$ such that $M = M^1 - M^2$ and $\sup_{t\geq 0} \|M_t\|_1 = M_0^1 + M_0^2$.

By construction, $M^i$ is of the form $M^i = f^i(t, X_t)$. We can assume, without loss of generality, that $M_0^i \neq 0$. Then, we can apply Widder's representation to the positive martingales $\frac{M^1}{M_0^1}$, $\frac{M^2}{M_0^2}$, obtaining two probability measures $\hat{\mu}^1$, $\hat{\mu}^2$ such that

$$\frac{M_t^i}{M_0^i} = \int_{\mathbb{R}^d} \mathcal{E}(v \cdot X)_t \; \hat{\mu}^i(dv) \quad P\text{-a.s.}.$$

We thus set $\mu^i := M_0^i \hat{\mu}^i$ and $\mu := \mu^1 - \mu^2$. Then, for all $t$,

$$M_t = M_t^1 - M_t^2 = \int_{\mathbb{R}^d} \mathcal{E}(v \cdot X)_t \; \mu(dv) \quad P\text{-a.s.}.$$

On the other hand, we have that

$$\|M_t\|_1 \leq \int_\Omega \int_{\mathbb{R}^d} \mathcal{E}(v \cdot X)_t \; |\mu|(dv) dP = \int_{\mathbb{R}^d} E\left[\mathcal{E}(v \cdot X)_t\right] \; |\mu|(dv) = \|\mu\|,$$

so that we finally get that

$$\|\mu^1\| + \|\mu^2\| = M_0^1 + M_0^2 = \sup_{t\geq 0} \|M_t\|_1 \leq \|\mu\| \leq \|\mu^1\| + \|\mu^2\|. \qquad \square$$

## 5.4 A characterization of Widder's measure

We now present the aforementioned characterization of the measure $\mu$ appearing in Proposition 5.3.3. To the best of our knowledge, this characterization is new and shows interesting analogies with results from Fourier analysis.

**Theorem 5.4.1.** *Let $(g(t, X_t))_{t\geq 0}$ be a continuous $L^1$-bounded martingale on $[0, \infty)$ with Widder's representation $g(t, X_t) = \int_{\mathbb{R}^d} \mathcal{E}(v \cdot X)_t \; \mu(dv)$. Then, the following assertions hold:*

(*i*) *If the measure $|\mu|$ has quadratic exponential moments of all orders, i.e.*

$$\int_{\mathbb{R}^d} e^{\lambda |v|^2} |\mu|(dv) < \infty \textit{ for all } \lambda > 0,$$

*then there is a family $(b_\alpha)_{\alpha\in\mathbb{N}^d}$ of coefficients in $\mathbb{R}$ such that, for all $t \geq 0$, the series $\sum_{\alpha\in\mathbb{N}^d} b_\alpha H_\alpha(t, X_t)$ converges unconditionally to $g(t, X_t)$ in $L^1$, and the coefficients $b_\alpha$ can be represented as*

$$b_\alpha = \frac{1}{\alpha!}\int_{\mathbb{R}^d} v^\alpha \ \mu(dv), \quad \alpha \in \mathbb{N}^d.$$

(*ii*) *Conversely, if $\mu$ is positive and there is a family $(b_\alpha)_{\alpha\in\mathbb{N}^d}$ such that $\sum_{\alpha\in\mathbb{N}^d} b_\alpha H_\alpha(t, X_t)$ converges unconditionally to $g(t, X_t)$ in $L^1$ for all $t \geq 0$, then $\mu$ has quadratic exponential moments of all orders. Moreover, the function $f(z) := \int_{\mathbb{R}^d} e^{v\cdot z}\mu(dv)$, $z \in \mathbb{C}^d$, is well defined, analytic of order 2, and can be represented, for $z \in \mathbb{C}^d$ and $t > 0$, as*

$$f(z) = \sum_{\alpha\in\mathbb{N}^d} b_\alpha z^\alpha = E\Bigg[g(t, X_t)\exp\bigg(\frac{1}{t}\Big(z\cdot X_t - \frac{z^T z}{2}\Big)\bigg)\Bigg].$$

*Proof.* We first show (*i*). By the properties of Hermite polynomials we have that, for all bijections $\sigma : \mathbb{N} \to \mathbb{N}^d$, $t \geq 0$ and $v \in \mathbb{R}^d$,

$$\begin{aligned}\mathcal{E}(v\cdot X)_t = \prod_{i=1}^d \mathcal{E}(v_i\cdot X^i)_t &= \prod_{i=1}^d\Big(\sum_{n\in\mathbb{N}}\frac{v_i^n}{n!}H_n(t, X_t^i)\Big)\\ &= \sum_{k\in\mathbb{N}}\frac{v^{\sigma(k)}}{\sigma(k)!}H_{\sigma(k)}(t, X_t),\end{aligned}$$

where the last equality holds pointwise $P$-a.s.; moreover, by Corollary 5.1.3 we obtain that, for all $n \in \mathbb{N}$,

$$\begin{aligned}\Big|\sum_{k=0}^n \frac{v^{\sigma(k)}}{\sigma(k)!}H_{\sigma(k)}(t, X_t)\Big| &= \Big|E\Big[\sum_{k=0}^n\frac{v^{\sigma(k)}}{\sigma(k)!}Z_t^{\sigma(k)}\Big|\mathcal{F}_t^X\Big]\Big|\\ &\leq E\Big[\sum_{k\in\mathbb{N}}\frac{|v^{\sigma(k)}Z_t^{\sigma(k)}|}{\sigma(k)!}\Big|\mathcal{F}_t^X\Big]\\ &= E\Big[\exp\Big(\sum_{i=1}^d|v_i||Z_t^i|\Big)\Big|\mathcal{F}_t^X\Big] \quad P\text{-a.s.},\end{aligned}$$

where the last equality is valid because of monotone convergence. On the other hand, it is not difficult to verify that $E\Big[\exp\Big(\sum_{i=1}^d |v_i||Z_t^i|\Big)\Big|\mathcal{F}_t^X\Big]$ lies in $L^1(P\otimes\mu)$. Namely, by Cauchy-Schwarz inequality and the assumption on the moments of $|\mu|$, we have that, for some constant $C>0$,

$$\begin{aligned}\int_{\mathbb{R}^d} E\Big[E\Big[\exp\Big(\sum_{i=1}^d |v_i||Z_t^i|\Big)\Big|\mathcal{F}_t^X\Big]\Big]|\mu|(dv) &\le \int_{\mathbb{R}^d} E\Big[\exp\big(|v||Z_t|\big)\Big]|\mu|(dv)\\ &\le C\int_{\mathbb{R}^d} E\Big[\exp\big(|v|Z_t\big)\Big]|\mu|(dv)\\ &= \frac{C}{\sqrt{2\pi t}}\int_{\mathbb{R}^d}\exp\Big(\frac{t|v|^2}{2}\Big)|\mu|(dv)<\infty.\end{aligned}$$

We can therefore apply the bounded convergence theorem, obtaining that $\sum_{k\in\mathbb{N}}\frac{v^{\sigma(k)}}{\sigma(k)!}H_{\sigma(k)}(t,X_t)$ converges absolutely in $L^1(P\otimes\mu)$ to $\mathcal{E}(v\cdot X_t)$. By integrating with respect to $\mu$, we finally get the unconditional convergence in $L^1(P)$ of $\sum_{\alpha\in\mathbb{N}^d}\frac{1}{\alpha!}\big(\int_{\mathbb{R}^d} v^\alpha\mu(dv)\big)H_\alpha(t,X_t)$ to $g(t,X_t)$, as $\sigma$ was arbitrary. This proves the first implication as well as the representation for the coefficients $b_\alpha$.

We now prove the second implication. First of all, we consider the function $\tilde f:\mathbb{C}^d\to\mathbb{C}$, $\tilde f(z):=\sum_{\alpha\in\mathbb{N}^d} b_\alpha z^\alpha$. By Proposition 5.2.6, we have that $g(t,X_t)\in L^2$ for $t\ge 0$: thus, we have that $b_\alpha=\frac{1}{\alpha! t^{|\alpha|}}E[g(t,X_t)H_\alpha(t,X_t)]$, since the Hermite polynomials are orthogonal to each other. We therefore see that the series $\tilde f$ corresponds to the one considered in Corollary 5.2.4, and this implies that $\tilde f$ is well defined, analytic of order 2, and

$$\tilde f(z)=\sum_{\alpha\in\mathbb{N}^d} b_\alpha z^\alpha = E\Big[g(t,X_t)\exp\Big(\frac{1}{t}\Big(z\cdot X_t-\frac{z^Tz}{2}\Big)\Big)\Big].$$

On the other hand, by substituting $g(t,X_t)$ with its Widder's representation in the above expression and by applying Fubini's theorem, it is easy to verify that $\tilde f(z)=f(z)=\int_{\mathbb{R}^d} e^{v\cdot z}\mu(dv)$. It remains to prove that $\mu$ has quadratic exponential moments: however, the estimate (5.2.2) implies similarly to Theorem 5.2.3 that $f(Z_s)\in L^1$ for all $s\ge 0$, since by taking $t>s$ we can verify that

$$E|f(Z_s)|\le K_t E\Big[\exp\Big(\frac{|Z_s|^2}{2t}\Big)\Big]<\infty.$$

On the other hand, since $f(z) = \int_{\mathbb{R}^d} e^{v \cdot z} \mu(dv)$ and $\mu$ is positive, by applying Fubini's theorem we obtain that

$$E[f(Z_t)] = \int_{\mathbb{R}^d} e^{\frac{t}{2}|v|^2} \mu(dv), \quad t > 0.$$

As $t > 0$ is arbitrary, this proves the existence of all quadratic exponential moments. □

We conclude this chapter with an example illustrating the fact that the quadratic exponential moments of $\mu$ are needed in order to have an expansion in Hermite series of the corresponding martingale. Let $\mu$ denote the standard Gaussian measure on $\mathbb{R}$: we can then verify that, for $t \geq 0$,

$$g(t, X_t) = \int_{\mathbb{R}} \mathcal{E}(v \cdot X)_t \, \mu(dv) = \frac{1}{\sqrt{t+1}} \exp\left(\frac{X_t^2}{2(t+1)}\right).$$

Since not all the quadratic exponential moments of $\mu$ exist, Theorem 5.4.1 implies that $g(t, X_t)$ cannot be represented as an Hermite series in $L^1$. Moreover, it is easy to check that $g(t, X_t)$ corresponds to the counterexample introduced in a deterministic setting by Pollard [62] in order to prove that the Hermite polynomials do not form a basis of $L^p$ for $p \neq 2$, and his conclusions can then be recovered from the results on $L^p$-convergence proved in Section 5.2.

We can therefore observe that Theorem 5.4.1 gives a full explanation of Pollard's counterexample: namely, the non-convergence of the corresponding Hermite series is simply due to the non-existence of quadratic exponential moments of any order for the corresponding Widder's measure. This observation allows us to construct several other counterexamples whose Hermite series do not converge in $L^1$: after choosing a measure on $\mathbb{R}$ which does not have quadratic exponential moments of all orders, it suffices to consider the martingale given by the corresponding Widder representation.

# Bibliography


[1] F. Antonelli: Backward-forward stochastic differential equations. *Ann. Appl. Probab.* **3**(3) (1993), 777–793.

[2] V. Bally: Approximation scheme for solutions of BSDE. In *Backward stochastic differential equations*, volume 364 of *Pitman Res. Notes Math. Ser.*, Longman, Harlow (1997), 177–191.

[3] P. Barrieu, N. Cazanave, and N. El Karoui: Closedness results for BMO semi-martingales and application to quadratic BSDEs. *C. R. Math. Acad. Sci. Paris* **346**(15-16) (2008), 881–886.

[4] C. Bender and R. Denk: A forward scheme for backward SDEs. *Stochastic Processes and their Applications* **117**(12) (2007), 1793–1812.

[5] C. Bender and J. Zhang: Time discretization and Markovian iteration for coupled FBSDEs. *Ann. Appl. Probab.* **18**(1) (2008), 143–177.

[6] J.-M. Bismut: Conjugate convex functions in optimal stochastic control. *J. Math. Anal. Appl.* **44** (1973), 384–404.

[7] B. Bouchard and N. Touzi: Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stochastic Processes and their Applications* **111**(2) (2004), 175–206.

[8] P. Briand, B. Delyon, and J. Mémin: Donsker-type theorem for BSDEs. *Electron. Comm. Probab.* **6** (2001), 1–14.

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

[10] P. Briand and Y. Hu: Quadratic BSDEs with convex generators and unbounded terminal conditions. *Probab. Theory Related Fields* **141**(3-4) (2008), 543–567.

[11] M. Casserini and F. Delbaen: A note on quadratic BSDEs and associated functional differential equations. *Forthcoming.* ETH Zurich (2011).

[12] M. Casserini and F. Delbaen: Predictable projections of conformal stochastic integrals: an application to Hermite series and to Widder's representation. *Forthcoming.* ETH Zurich (2011).

[13] M. Casserini and G. Liang: Functional differential equations for coupled forward-backward stochastic dynamics and related numerical approximations. *Forthcoming.* ETH Zurich and Oxford-Man Institute (2011).

[14] P. Cheridito and M. Stadje: BSΔEs and BSDEs with non-Lipschitz drivers: comparison, convergence and robustness. *arXiv:1002.0175v1* (2010).

[15] D. Chevance: Numerical methods for backward stochastic differential equations. In *Numerical methods in finance*, Publ. Newton Inst., Cambridge Univ. Press, Cambridge (1997), 232–244.

[16] D. Crisan and K. Manolarakis: Solving backward stochastic differential equations using the cubature method. *Preprint.* Available at http://www2.imperial.ac.uk/~dcrisan/Cub_on_BSDE.pdf (2007).

[17] B. Davis: Brownian motion and analytic functions. *Ann. Probab.* **7**(6) (1979), 913–932.

[18] F. Delarue: On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case. *Stochastic Process. Appl.* **99**(2) (2002), 209–286.

[19] F. Delarue: Estimates of the solutions of a system of quasi-linear PDEs. A probabilistic scheme. In *Séminaire de Probabilités XXXVII*, volume 1832 of *Lecture Notes in Math.*, Springer, Berlin (2003), 290–332.

[20] F. Delarue and S. Menozzi: A forward-backward stochastic algorithm for quasi-linear PDEs. *Ann. Appl. Probab.* **16**(1) (2006), 140–184.

[21] F. Delbaen, Y. Hu, and X. Bao: Backward SDEs with superquadratic growth. *Probability Theory and Related Fields* (2009), 1–48.

[22] F. Delbaen, Y. Hu, and A. Richou: On the uniqueness of solutions to quadratic BSDEs with convex generators and unbounded terminal conditions. *arXiv:0906.0752v1* (2009), to appear in *Annales de l'Institut Henri Poincaré Probabilités et Statistiques* (2011).

[23] F. Delbaen and S. Tang: Harmonic analysis of stochastic equations and backward stochastic differential equations. *Probab. Theory Related Fields* **146**(1-2) (2010), 291–336.

[24] C. Dellacherie and P.-A. Meyer: *Probabilities and potential B. Theory of Martingales*, volume 72 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam (1982).

[25] L. Delong: Applications of time-delayed backward stochastic differential equations to pricing, hedging and portfolio management. *arXiv: 1005.4417v3* (2011).

[26] Ł. Delong and P. Imkeller: Backward stochastic differential equations with time delayed generators—results and counterexamples. *Ann. Appl. Probab.* **20**(4) (2010), 1512–1536.

[27] G. Dos Reis, A. Réveillac, and J. Zhang: FBSDE with time delayed generators: $L^p$-solutions, differentiability, representation formulas and path regularity. *arXiv:1008.1149v1* (2010).

[28] J. J. Douglas, J. Ma, and P. Protter: Numerical methods for forward-backward stochastic differential equations. *The Annals of Applied Probability* **6**(3) (1996), 940–968.

[29] D. Duffie and L. Epstein: Stochastic differential utility. *Econometrica* **60**(2) (1992), 353–394.

[30] N. El Karoui, S. Hamadène, and A. Matoussi: *Indifference pricing: theory and applications*, chapter BSDEs and applications. Princeton University Press (2009), 267–320.

[31] N. El Karoui and L. Mazliak (editors): *Backward stochastic differential equations*, volume 364 of *Pitman Research Notes in Mathematics Series*. Longman, Harlow (1997).

[32] N. El Karoui, S. Peng, and M. C. Quenez: Backward stochastic differential equations in finance. *Mathematical Finance* **7**(1) (1997), 1–71.

[33] L. C. Evans: *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. Second edition, American Mathematical Society, Providence, R.I. (2010).

[34] R. K. Getoor and M. J. Sharpe: Conformal martingales. *Inventiones Mathematicae* **16** (1972), 271–308.

[35] E. Gobet, J. Lemor, and X. Warin: A regression-based Monte Carlo method to solve backward stochastic differential equations. *The Annals of Applied Probability* **15**(3) (2005), 2172–2202.

[36] E. Hille: A class of reciprocal functions. *Ann. of Math. (2)* **27**(4) (1926), 427–464.

[37] E. Hille: Contributions to the theory of Hermitian series. *Duke Math. J.* **5** (1939), 875–936.

[38] E. Hille: Contributions to the theory of Hermitian series II. The representation problem. *Trans. Amer. Math. Soc.* **47** (1940), 80–94.

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

[40] Y. Hu and S. Peng: Solution of forward-backward stochastic differential equations. *Probab. Theory Related Fields* **103**(2) (1995), 273–283.

[41] P. Imkeller and G. Dos Reis: Path regularity and explicit convergence rate for BSDE with truncated quadratic growth. *arXiv:0905.0788v3* (2009).

[42] P. Imkeller, G. Dos Reis, and J. Zhang: Results on numerics for FBSDE with drivers of quadratic growth. *Contemporary Quantitative Finance: Essays in Honour of Eckhard Platen* (2010).

[43] I. Karatzas and S. E. Shreve: *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Second edition, Springer-Verlag, New York (1991).

[44] N. Kazamaki: *Continuous exponential martingales and BMO*, volume 1579 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin (1994).

[45] P. E. Kloeden and E. Platen: *Numerical solution of stochastic differential equations*, volume 23 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin (1992).

[46] M. Kobylanski: Backward stochastic differential equations and partial differential equations with quadratic growth. *Ann. Probab.* **28**(2) (2000), 558–602.

[47] M. Ledoux: Isoperimetry and Gaussian analysis. In *Lectures on probability theory and statistics (Saint-Flour, 1994)*, volume 1648 of *Lecture Notes in Math.*, Springer, Berlin (1996), 165–294.

[48] J.-P. Lepeltier and J. San Martín: Existence for BSDE with superlinear-quadratic coefficient. *Stochastics Stochastics Rep.* **63**(3-4) (1998), 227–240.

[49] G. Liang, T. Lyons, and Z. Qian: Backward stochastic dynamics on a filtered probability space. *arXiv:0904.0377v3* (2010), to appear in *The Annals of Probability* (2011).

[50] J. Ma, P. Protter, and J. M. Yong: Solving forward-backward stochastic differential equations explicitly—a four step scheme. *Probab. Theory Related Fields* **98**(3) (1994), 339–359.

[51] J. Ma and J. Yong: *Forward-backward stochastic differential equations and their applications*, volume 1702 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin (1999).

[52] J. Ma and J. Zhang: Path regularity for solutions of backward stochastic differential equations. *Probability Theory and Related Fields* **122** (2002), 163–190.

[53] R. Mansuy and M. Yor: *Aspects of Brownian motion*. Universitext, Springer-Verlag, Berlin (2008).

[54] G. N. Milstein and M. V. Tretyakov: Numerical algorithms for forward-backward stochastic differential equations. *SIAM J. Sci. Comput.* **28**(2) (2006), 561–582.

[55] M.-A. Morlais: Quadratic BSDEs driven by a continuous martingale and applications to the utility maximization problem. *Finance Stoch.* **13**(1) (2009), 121–150.

[56] D. Nualart: *The Malliavin calculus and related topics*. Second edition, Probability and its Applications, Springer-Verlag, Berlin (2006).

[57] É. Pardoux: Backward stochastic differential equations and viscosity solutions of systems of semilinear parabolic and elliptic PDEs of second order. In *Stochastic analysis and related topics, VI (Geilo, 1996)*, volume 42 of *Progr. Probab.*, Birkhäuser Boston, Boston, MA (1998), 79–127.

[58] É. Pardoux and S. G. Peng: Adapted solution of a backward stochastic differential equation. *Systems Control Lett.* **14**(1) (1990), 55–61.

[59] É. Pardoux and S. G. Peng: Backward stochastic differential equations and quasilinear parabolic partial differential equations. In *Stochastic partial differential equations and their applications (Charlotte, NC, 1991)*, volume 176 of *Lecture Notes in Control and Inform. Sci.*, Springer, Berlin (1992), 200–217.

[60] É. Pardoux and S. Tang: Forward-backward stochastic differential equations and quasilinear parabolic PDEs. *Probab. Theory Related Fields* **114**(2) (1999), 123–150.

[61] S. Peng and Z. Wu: Fully coupled forward-backward stochastic differential equations and applications to optimal control. *SIAM J. Control Optim.* **37**(3) (1999), 825–843.

[62] H. Pollard: The mean convergence of orthogonal series II. *Trans. Amer. Math. Soc.* **63** (1948), 355–367.

[63] P. E. Protter: *Stochastic integration and differential equations*, volume 21 of *Stochastic Modelling and Applied Probability*. Second edition (version 2.1), Springer-Verlag, Berlin (2005).

[64] D. Revuz and M. Yor: *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften*. Third edition, Springer-Verlag, Berlin (1999).

[65] A. Richou: Numerical simulation of BSDEs with drivers of quadratic growth. *arXiv:1001.0401v3* (2010).

[66] B. L. Rozovsky: A note on strong solutions of stochastic differential equations with random coefficients. In *Stochastic differential systems (Proc. IFIP-WG 7/1 Working Conf., Vilnius, 1978)*, volume 25 of *Lecture Notes in Control and Information Sci.*, Springer, Berlin (1980), 287–296.

[67] M. Stadje: *Backward stochastic differential equations and dynamic risk measures: from discrete to continuous time*. Ph.D. thesis, Princeton University (2009).

[68] G. Szegő: *Orthogonal polynomials*, volume 23 of *American Mathematical Society Colloquium Publications*. Fourth edition, American Mathematical Society, Providence, R.I. (1975).

[69] R. Tevzadze: Solvability of backward stochastic differential equations with quadratic growth. *Stochastic Process. Appl.* **118**(3) (2008), 503–515.

[70] D. V. Widder: Positive temperatures on an infinite rod. *Trans. Amer. Math. Soc.* **55** (1944), 85–95.

[71] J. Yong: Finding adapted solutions of forward-backward stochastic differential equations: method of continuation. *Probab. Theory Related Fields* **107**(4) (1997), 537–572.

[72] J. Zhang: *Some fine properties of backward stochastic differential equations*. Ph.D. thesis, Purdue University (2001).

[73] J. Zhang: A numerical scheme for BSDEs. *The Annals of Applied Probability* **14**(1) (2004), 459–488.

[74] J. Zhang: The wellposedness of FBSDEs. *Discrete Contin. Dyn. Syst. Ser. B* **6**(4) (2006), 927–940.

[75] J. Zhang: The wellposedness of FBSDEs (II). *Preprint.* Available at http://www-bcf.usc.edu/~jianfenz/Papers/fbsde2.pdf (2006).