

# Backward Stochastic Differential Equations with Jumps

by

Jan Olav Halle

**THESIS**  
*for the degree of*  
**MASTER OF SCIENCE**

*(Master i Modelling og dataanalyse)*



*Faculty of Mathematics and Natural Sciences*  
*University of Oslo*

*June 2010*

*Det matematisk- naturvitenskapelige fakultet*  
*Universitetet i Oslo*



*One's work may be finished some day,  
but one's education never.*

Alexandre Dumas

## Acknowledgments

I could not have done this thesis without the help and support of the good people around me, the great environment of the 6th floor, my sturdy lap top, my Spotify-playlist, coffee and Coca-Cola. I would like to specifically thank the following people (with undefined order of priority):

- Bernt Øksendal: My supervisor and excellent advisor. Feeling a little low, a meeting with him always increased the motivation.
- Madeleine G. Innholt: My girlfriend. Her patience and support has been incalculable for me.
- My family: Maybe especially my mom which has been very patient not being able to see her son as much as she would like.
- Giulia Di Nunno: My teacher in most of my stochastic analysis courses. Much of my knowledge and intuition comes from her.



# Contents

<b>1</b>	<b>Introduction</b>	<b>6</b>
<b>2</b>	<b>Preliminaries</b>	<b>7</b>
2.1	Basics on Lévy processes . . . . .	7
2.2	The problem . . . . .	8
2.3	Why study BSDEs . . . . .	11
<b>3</b>	<b>Existence and uniqueness</b>	<b>13</b>
3.1	The a priori estimates . . . . .	14
3.2	The existence and uniqueness theorem . . . . .	18
<b>4</b>	<b>Linear BSDEs with jumps</b>	<b>21</b>
4.1	Explicit formula of linear BSDEs . . . . .	21
4.2	The importance of linear BSDEs . . . . .	27
<b>5</b>	<b>Comparison of solutions</b>	<b>27</b>
5.1	The comparison theorem . . . . .	27
5.2	Why we need a strong generator . . . . .	29
<b>6</b>	<b>Dependence on a parameter</b>	<b>30</b>
6.1	The continuous case . . . . .	30
6.2	The differentiable case . . . . .	31
6.3	The differentiable case for quasi-strong generators . . . . .	37
<b>7</b>	<b>Malliavin derivative</b>	<b>40</b>
7.1	Definition of the Malliavin derivative and preliminary results	42
7.2	Differentiability in $\tilde{N}$ . . . . .	49
<b>8</b>	<b>Combining the derivatives</b>	<b>57</b>
8.1	The idea . . . . .	58
8.2	The results on the combined derivative . . . . .	59

# 1 Introduction

This Master thesis is about backward stochastic differential equations (BSDEs) with jumps. The motivation for this topic was a suggestion from my supervisor Bernt Øksendal after we both attended classes in the course MAT4760 about BSDEs with Brownian noise. I wanted to write mainly a theoretical thesis and this topic seemed like a good way to do this in addition to learn some of the theory and calculus of Lévy processes.

The purpose of my thesis will be to look at theory about BSDEs from the article [11] we used in MAT4760 and see if it is possible to generalize it to include jumps. The formulation we gave on the application for approval of study plan was the following:

The purpose of this Master project is to study properties of backward stochastic differential equations (BSDEs) with jumps. Topics that will be investigated include existence and uniqueness theorems, comparison theorems, dependence on parameters, concave generators, BSDEs with larger filtrations, Malliavin calculus and applications to finance.

Writing this introduction just before deadline, I realize that I have achieved many, but not all of these tasks. I have focused on the theory, and not on applications. I have also not looked at concave generators. The rest of the topics, I have managed to dive into. All but *BSDEs with larger filtrations* I have managed to generalize. Also, because both *dependence of parameters* and *Malliavin calculus* include some sort of derivative, I have studied the combination of these two derivatives.

The thesis be organized as follows. In section 2 I give a brief introduction to Lévy processes and some basic definitions before we state the general problem. I will also discuss why studying BSDEs is interesting. In section 3 I prove existence and uniqueness which shows that there is something to study. Then I go on studying a special case known as linear BSDEs in section 4. Here I am able to give an explicit solution. This is further used to prove a comparison theorem in section 5. After this I study dependence on a parameter, both continuity and differentiability, in section 6. In section 7, I first use some of space both defining the Malliavin derivative for the combined Wiener and jump case and stating some preliminary results. Then I prove a result regarding differentiability of the solution. In section 8 I combine section 6 and 7 looking at the derivative in both the parameter and the Malliavin sense combined.

## 2 Preliminaries

### 2.1 Basics on Lévy processes

Let us briefly recall some basic properties of Lévy processes. The notation and topics are mainly from [10], chapter 9. This is briefly what we need this thesis about Lévy processes. Additional theory about the Fréchet derivative and Malliavin calculus will be given in chapter 6 and 7 respectively.

Let  $\eta_t$  be a one-dimensional Lévy process on a complete probability space  $(\Omega, \mathcal{F}, P)$ . From theorem 30 in [24] we know that  $\eta$  has a càdlàg modification, and we will always use this modification. We define the jumps of  $\eta$  by  $\Delta\eta_t = \eta_t - \eta_{t-}$ , where  $\eta_{0-} = \eta_0 = 0$ .

Let  $\mathbb{R}_0 = \mathbb{R} \setminus 0$  and define  $\mathcal{B}(\mathbb{R}_0)$  as the  $\sigma$ -algebra generated by the Borel sets  $U \subset \mathbb{R}$  such that  $\bar{U} \subset \mathbb{R}_0$ . We may now define a family of set functions on  $\mathcal{B}(\mathbb{R}_0)$  by

$$N(t, U) = \sum_{0 \leq s \leq t} I(\Delta\eta_s \in U), \quad U \in \mathcal{B}(\mathbb{R}_0), \quad t \geq 0. \quad (1)$$

Because  $\eta$  is càdlàg, we know that  $N(t, U) < \infty$  for all  $U \in \mathcal{B}(\mathbb{R}_0)$  with closure not containing zero. Further, (1) defines a Poisson random measure  $N$  on  $\mathcal{B}(0, \infty) \times \mathcal{B}(\mathbb{R}_0)$  given by

$$(a, b] \times U \rightarrow N(b, U) - N(a, U), \quad 0 < a \leq b, \quad U \in \mathcal{B}(\mathbb{R}_0),$$

with its standard extension. This random measure is called the jump measure of  $\eta$ , and we use the notation  $N(dt, dz), t > 0, z \in \mathbb{R}_0$ , for the differential form of the jump measure.

The Lévy measure  $\nu$  of  $\eta$  is defined by

$$\nu(U) = E[N(1, U)], \quad U \in \mathcal{B}(\mathbb{R}_0),$$

which is known to satisfy the condition

$$\int_{\mathbb{R}_0} \min(1, z^2) \nu(dz) < \infty.$$

Further, we define the compensated jump measure  $\tilde{N}$  on  $\mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R}_0)$  by

$$\tilde{N}(dt, dz) = N(dt, dz) - \nu(dz)dt.$$

Let  $W$  be a brownian motion on  $(\Omega, \mathcal{F}, P)$ , and let  $\{\mathcal{F}_t\}_{t \geq 0}$  be the filtration generated by  $W$  and  $\tilde{N}$  where  $\mathcal{F}_0$  contains all the  $P$ -null sets of  $\mathcal{F}$ .

Now, from [10], theorem 9.3, we know that  $\eta$  has the following representation

$$\eta_t = at + \sigma W_t + \int_0^t \int_{|z| < 1} z \tilde{N}(ds, dz) + \int_0^t \int_{|z| \geq 1} z N(ds, dz)$$

for some real numbers  $a_1, \sigma$ . We will, however, assume that  $\eta$  satisfies  $E[\eta_t^2] < \infty$  for all  $t \geq 0$ . Then  $\int_{|z| \geq 1} |z|^2 \nu(dz) < \infty$ , and  $\eta$  has the representation

$$\eta_t = a_1 t + \sigma W_t + \int_0^t \int_{\mathbb{R}_0} z \tilde{N}(ds, dz),$$

where  $a = a_1 + \int_{|z| \geq 1} |z| \nu(dz)$ .

Motivated by this representation of  $\eta$ , we define a class of processes  $V = V_t, t \geq 0$ , admitting the stochastic integral representation in the form

$$V_t = x + \int_0^t \alpha_s ds + \int_0^t \beta_s dW_s + \int_0^t \int_{\mathbb{R}_0} \gamma_s(z) \tilde{N}(ds, dz) \quad (2)$$

with  $\alpha, \beta$  and  $\gamma$  predictable. The term predictable will be defined in section 2.2. For the integrals to be defined, we need the following assumption:

$$\int_0^T [|\alpha_s| + \beta_s^2 + \int_{\mathbb{R}_0} \gamma_s^2(z) \nu(dz)] ds < \infty \quad a.s.$$

Now all the integrals are well defined, and it is well known that the stochastic integrals are local martingales. If we strengthen this condition to

$$E \left[ \int_0^T [|\alpha_s| + \beta_s^2 + \int_{\mathbb{R}_0} \gamma_s^2(z) \nu(dz)] ds \right] < \infty, \quad (3)$$

then the corresponding stochastic integrals are martingales. This will mainly be the case in this thesis.

Finally, we will use the following short-hand differential form of (2) to some degree:

$$dV_t = \alpha_t dt + \beta_t dW_t + \int_{\mathbb{R}_0} \gamma_t(z) \tilde{N}(dt, dz)$$

$$V_0 = x.$$

We call these processes Itô-Lévy processes.

## 2.2 The problem

Let  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$ ,  $\eta$ ,  $\tilde{N}$  and  $W$  be as in the previous section. The underlying problem for this thesis, will be to find solutions  $(X_t, Z_t, K_t(z))$  in an appropriate space satisfying

$$dX_t = -f(t, X_t, Z_t, K_t(\cdot)) dt + Z_t dW_t + \int_{\mathbb{R}} K_t(z) \tilde{N}(dt, dz)$$

$$X_T = \xi \quad (4)$$



for some appropriate  $(f, \xi)$ , a fixed  $T > 0$  and where  $f$  depends on the operator  $K_t(\cdot)$  which from now on, for notational simplicity, will be written as  $K_t$ . In all the following, we will need definitions that follows.

Let  $\mathcal{P}$  be the smallest  $\sigma$ -algebra on  $\Omega \times [0, T]$  such that all left-continuous, adapted processes are measurable. A  $\mathcal{P}$ -measurable process  $\phi : \Omega \times [0, T] \rightarrow \mathbb{R}$  will be called predictable. Further, let  $\mathcal{P}'$  be the smallest  $\sigma$ -algebra on  $\Omega \times [0, T] \times \mathbb{R}_0$  such that all mappings  $\theta : \Omega \times [0, T] \times \mathbb{R}_0 \rightarrow \mathbb{R}$  with the following properties are measurable:

1.  $\forall t > 0, (\omega, x) \rightarrow \theta(\omega, t, x)$  is  $\mathcal{F}_t \times \mathcal{B}(\mathbb{R}_0)$ -measurable
2.  $\forall (\omega, x), t \rightarrow \theta(\omega, t, x)$  is left continuous.

A mapping  $\theta' : \Omega \times [0, T] \times \mathbb{R}_0 \rightarrow \mathbb{R}$  which is  $\mathcal{P}'$ -measurable, will also be called predictable.

Now, define the following spaces for all  $\beta \geq 0$ :

- $L_{T,\beta}^2(\mathbb{R})$ : The space of all  $\mathcal{F}_T$ -measurable random variables  $X : \Omega \rightarrow \mathbb{R}$  such that  $\|X\|_\beta^2 = E[e^{\beta T} |X|^2] < \infty$ .
- $S_{T,\beta}^2$ : The space of all adapted, càdlàg processes  $\gamma : \Omega \times [0, T] \rightarrow \mathbb{R}$  such that  $\|\gamma\|_{S_{T,\beta}^2}^2 = E[e^{\beta T} \sup_{0 \leq t \leq T} |\gamma_t|^2] < \infty$ .
- $H_{T,\beta}^2(\mathbb{R})$ : The space of all predictable processes  $\phi : \Omega \times [0, T] \rightarrow \mathbb{R}$  such that  $\|\phi\|_{H_{T,\beta}^2(\mathbb{R})}^2 = E[\int_0^T e^{\beta t} |\phi_t|^2 dt] < \infty$ .
- $\hat{H}_{T,\beta}^2(\mathbb{R})$ : The space of all predictable mappings  $\theta : \Omega \times [0, T] \times \mathbb{R}_0 \rightarrow \mathbb{R}$  such that  $\|\theta\|_{\hat{H}_{T,\beta}^2(\mathbb{R})}^2 = E[\int_0^T \int_{\mathbb{R}} e^{\beta t} |\theta_t(z)|^2 \nu(dz) dt] < \infty$ .
- $\mathcal{V}_\beta = S_{T,\beta}^2 \times H_{T,\beta}^2(\mathbb{R}) \times \hat{H}_{T,\beta}^2(\mathbb{R})$ , which will be our solution space.
- $\hat{L}_T^2(\mathbb{R}_0, \mathcal{B}(\mathbb{R}_0), \nu)$ : The space of all  $\mathcal{B}(\mathbb{R}_0)$ -measurable mappings  $\psi : \mathbb{R}_0 \rightarrow \mathbb{R}$  such that  $\|\psi\|_{\hat{L}_T^2(\mathbb{R}_0, \mathcal{B}(\mathbb{R}_0), \nu)}^2 = \int_{\mathbb{R}_0} |\psi(z)|^2 \nu(dz) < \infty$ .

For notational simplicity, we skip the subscripts on the norms when the norm being used is clear, but specify  $\beta$ , e.g.  $\|\phi\|_{H_{T,\beta}^2(\mathbb{R})} = \|\phi\|_\beta$ . When  $\beta = 0$  we also skip the  $\beta$ , so that  $\|\phi\|_{H_{T,0}^2(\mathbb{R})} = \|\phi\|$ . Notice that all the norms defined will be equivalent for different choices  $\beta$ . The motivation for introducing the  $\beta$  is from [11] and makes the answers slightly nicer.

**Definition 2.1.** Suppose  $\xi \in L_{T,\beta}^2(\mathbb{R})$  and that  $f : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R} \times \hat{L}_T^2(\mathbb{R}_0, \mathcal{B}(\mathbb{R}_0), \nu) \rightarrow \mathbb{R}$  satisfies

- $f$  is  $\mathcal{P} \times \mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R}) \times \mathcal{B}(\hat{L}_T^2(\mathbb{R}_0, \mathcal{B}(\mathbb{R}_0), \nu))$ -measurable

- there exists a  $C_1$  such that for all  $(x^i, z^i, k^i) \in \mathbb{R} \times \mathbb{R} \times \hat{L}_T^2(\mathbb{R}_0, \mathcal{B}(\mathbb{R}_0), \nu)$ ,  $i = 1, 2$  and for  $\mathcal{F} \times \mathcal{B}([0, T])$ -a.a.  $(\omega, t)$ ,

$$f(t, 0, 0, 0) \in H_T^2$$

$$|f(t, x^1, z^1, k^1) - f(t, x^2, z^2, k^2)| \leq C(|x^1 - x^2| + |z^1 - z^2| + \|k^1 - k^2\|).$$

Then  $(\xi, f)$  is called a standard parameter. We will sometimes call  $\xi$  the end point and  $f$  the generator.

**Definition 2.2.** Given a standard parameter  $(f, \xi)$ , a solution to the BSDE associated to  $(f, \xi)$  is a triple  $(X, Z, K) \in \mathcal{V}$  that satisfies (4).

Note that  $X$  is a Itô-Lévy process satisfying (3). Now, let us define two other classes of parameters:

**Definiton 2.3.** Suppose the pair  $(f, \xi)$  is such that  $\xi \in L_T^2$  and  $f$  satisfies

$$f(t, x, z, k) = g(t, x, z, \int_{\mathbb{R}_0} \Psi_t(y) k(y) \nu(dy)), \quad (5)$$

where  $\Psi$  is predictable and satisfies

$$c_{\Psi}^2(1 \wedge |y|) \leq \Psi_t(y) \leq c_{\Psi}^1(1 \wedge |y|), \quad \mathcal{F} \times \mathcal{B}([0, T]) \times \mathcal{B}(\mathbb{R}_0) - a.e.,$$

for a  $c_{\Psi}^2 \in (-1, 0]$  and a  $c_{\Psi}^1 \geq 0$ . Further, suppose  $g$  is a  $\mathcal{P} \times \mathcal{B} \times \mathcal{B} \times \mathcal{B}$ -measurable function and there exists a constant  $C_g \in \mathbb{R}$  such that for all  $(x^i, z^i, r^i) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ ,  $i = 1, 2$  and for  $\mathcal{F} \times \mathcal{B}([0, T])$ -a.a.  $(\omega, t)$ , we have

$$g(t, 0, 0, 0) \in H_T^2$$

$$|g(t, x^1, z^1, r^1) - g(t, x^2, z^2, r^2)| \leq C_g(|x^1 - x^2| + |z^1 - z^2| + |r^1 - r^2|).$$

Then we call  $(f, \xi)$  a quasi-strong standard parameter. If in addition we have that for all  $(x, z) \in \mathbb{R} \times \mathbb{R}$  and for all  $r^i \in \mathbb{R}$ ,  $i = 1, 2$

$$|g(t, x, z, r^1) - g(t, x, z, r^2)| \leq |r^1 - r^2|, \quad \mathcal{F} \times \mathcal{B}([0, T]) - a.e.$$

we call  $(f, \xi)$  a strong standard parameter.

Note that a quasi-strong standard parameter is also a standard parameter, because from Hölder's inequality we have

$$|g(t, x, z, \int_{\mathbb{R}_0} \Psi_t(y) k_1(y) \nu(dy)) - g(t, x, z, \int_{\mathbb{R}_0} \Psi_t(y) k_2(y) \nu(dy))| \leq$$

$$C_g \left| \int_{\mathbb{R}_0} \Psi_t(y) (k_1(y) - k_2(y)) \nu(dy) \right|$$

$$C_g \left| \int_{\mathbb{R}_0} \Psi_t(y)^2 \nu(dy) \right|^{\frac{1}{2}} \left| \int_{\mathbb{R}_0} (k_1(y) - k_2(y))^2 \nu(dy) \right|^{\frac{1}{2}}$$

$$\leq D \|k_1 - k_2\|,$$

for a  $D > 0$ , because  $\int_{\mathbb{R}_0} \Psi_t(y)^2 \nu(dy)$  is bounded by assumption.

The reason why we define this subset of the standard parameters is because some of the results, e.g. the comparison theorem, will not be true for the general standard parameters, but only for strong or quasi-strong standard parameters. The idea of this subset is from [25], where just this comparison theorem is proved. Later we will see that the important class of linear generators is a subset of the quasi-strong generators.

### 2.3 Why study BSDEs

In addition to be interesting theoretically, the problem of solving a BSDE often occurs in different circumstances. Maybe the most well-known is that of finding the replicating portfolio or the price of a contingent claim. This problem must necessarily be of a backward kind, because we know where we are supposed to end, the contingent claim, and we want to find where to start and how to get there. For instance, suppose the market is of the following type (of course with sufficient conditions for all the integrals to be well defined):

1. A risk free asset, where the price at time  $t$ ,  $S_t^0$  solves

$$\begin{aligned} dS_t^0 &= r_t S_t^0 dt \\ S_T^0 &= 1. \end{aligned}$$

2. A risky asset, which price at time  $t$ ,  $S_t$ , solves

$$\begin{aligned} dS_t &= S_{t-} [\mu_t dt + \sigma_t dW_t + \int_{\mathbb{R}_0} \gamma_t(z) \tilde{N}(dt, dz)] \\ S_T &> 0. \end{aligned}$$

Suppose further that the process  $(\phi_t^0, \phi_t)$  is the amount of each asset held at each time  $t$ ,  $X$  is the wealth process and  $\xi$  is the claim. Then we can show that  $X$  solves the equation

$$\begin{aligned} dX_t &= [X_t r_t + \phi_t S_t (\beta_t - r_t)] dt + \phi_t S_t \sigma_t dW_t + \int_{\mathbb{R}_0} \phi_t S_{t-} \gamma_t(z) \tilde{N}(dt, dz) \\ X_T &= \xi. \end{aligned}$$

If  $\xi$  is attainable, and we give further conditions, the solution of the BSDE

$$\begin{aligned} dX_t &= [X_t r_t + Y_t \frac{\beta_t - r_t}{\sigma_t}] dt + Y_t dW_t + \int_{\mathbb{R}_0} K_t(z) \tilde{N}(dt, dz) \\ X_T &= \xi \end{aligned}$$

gives us the replicating portfolio and the price.

As is well known, all claims are replicable in the classical Brownian Black and Scholes market. In this market, where jumps are allowed, we see that the solution of the BSDE have to “match” both  $Z_t = \phi_t S_t \sigma_t$  and  $K_t(z) = \phi_t S_{t-\gamma_t}(z)$ . Thus this market is, except for very trivial cases, incomplete. Often there are additional assumptions added, which would make the Black and Scholes market incomplete and hence make the jump market even “more incomplete”.

So to give a price to a non-replicable claim, one has to define what the price should be and find a technique to get the price. Often this means to pick the “best” claim or claims that are replicable and use its or their price to define the price of the claim in question.

An example I have seen, is when this takes the form of a sequence of BSDEs in [7]. Here, pricing of American options with constrained portfolios is the topic. This problem is solved using a penalization method, where a sequence of BSDEs is penalized more and more for going outside the constraint by adding an extra term in the generator. The limit is then what we are interested in. Of course, limit results about BSDEs are needed here.

Another exciting application of BSDEs, is solving a classical differential equation. This is done by studying a coupled forward-backward system. If the solution is  $X_t^u$ , where  $t$  is the time variable and  $u$  is the starting time, one can show that  $X_t^t$  is a deterministic function which also is a viscosity solution of the differential equation. This is done in the Brownian case in theorem 4.2 in [11], the article I have been working with, and was originally a problem I wanted to study. Unfortunately for me, it is generalized to include jumps in [3].

After finishing chapter 7 in the beginning of March, I wanted a last project to work on. As you will see, this project became a theoretical result combining two derivatives. However, my wish was to have an application, both because then I could add “with applications” to the name of the thesis, and, most important, because it would be interesting. In fact, I worked with two problems before I studied the topics in chapter 8. For different reasons, I had to abandon both researches and instead, I will mention them here as more exciting applications of BSDE theory.

The first problem was a suggestion from my supervisor. That is, to try to generalize the results from [13] to include jumps. The problem is to maximize the utility of a small investor with a constraint on his portfolio in a market with noise. This is another example where the penalization method is applied, and limit results from [22] is used for this to work. The results of [22] I found generalized to jumps in [18]. So I started researching this. However, after a week or two I stumbled upon the preprints [21] and [20] from a PhD thesis where this is done.[13] is generalized.

The other problem was to generalize [5]. Here, an optimal superhedging portfolio with constraints is found in a Brownian market, also this using a

penalization method. I worked with this for a few weeks, but the complexity of the problem soon became clear. In the Brownian setting, we approximate the problem, the unattainable portfolio, with attainable portfolios. This we cannot do in the jump setting. So this problem I would have to overcome. I have not found any articles looking at this problem for jumps, so as far as I know, this is an open problem.

The conclusion is that there are many good reasons for studying BSDEs. In finance the jump markets are known to give more flexibility to modelling. The book [8], for instance, has much material on this. Even though I do not have any applications, there certainly is potential. After some of the results, I will comment on this.

### 3 Existence and uniqueness

For any analysis about BSDEs on the form (4) to be done, a discussion about existence and uniqueness must be done. The existence and uniqueness result has been done by many authors. Examples I was able to find include [26] and [3].

The classical way to prove existence and uniqueness is by giving some sort of a priori estimates. Using these estimates, a contraction is constructed which fixpoint is the unique solution of the BSDE. This is because the contraction lives on a Banach space.

I will do this with the equivalent  $\beta$ -norms, which is not done in [26] and [3]. This is because this is the way it is done in [11] which is the article I am generalizing. Still, the idea is the same.

The difficulty in my approach is to make some terms vanish in the a priori estimates. But when these estimates are established, the existence and uniqueness are proven exactly the same way as in [11]. This is due to the martingale representation theorem for jumps.

Before the results, we make an observation.

**Proposition 3.1.** *Let  $(f, \xi)$  be a standard parameter and suppose  $(X, Z, K) \in H_T^2 \times H_T^2 \times \hat{H}_T^2$  solves (4). Then  $X \in S_{T,\beta}^2$ .*

*Proof.* We observe that

$$\begin{aligned} \sup_{0 \leq t \leq T} |X_t| \leq & |\xi| + \int_0^T |f(s, X_s, Z_s, K_s)| ds + \\ & \sup_{0 \leq t \leq T} \left| \int_t^T Z_s dW_s \right| + \sup_{0 \leq t \leq T} \left| \int_t^T \int_{\mathbb{R}_0} K_s(z) \tilde{N}(ds, dz) \right|. \end{aligned}$$

Now,  $E[|\xi|^2 + \int_0^T |f(s, X_s, Z_s, K_s)|^2 ds] < \infty$  because  $(f, \xi)$  are standard

parameters. Further we have that

$$\begin{aligned} & \sup_{0 \leq t \leq T} \left| \int_t^T \int_{\mathbb{R}_0} K_s(z) \tilde{N}(ds, dz) \right| \leq \\ & \left| \int_0^T \int_{\mathbb{R}_0} K_s(z) \tilde{N}(ds, dz) \right| + \sup_{0 \leq t \leq T} \left| \int_0^t \int_{\mathbb{R}_0} K_s(z) \tilde{N}(ds, dz) \right|. \end{aligned}$$

Here  $\left| \int_0^T \int_{\mathbb{R}_0} K_s(z) \tilde{N}(ds, dz) \right|$  is square integrable by assumption. Because  $M_t := \int_0^t \int_{\mathbb{R}_0} K_s(z) \tilde{N}(ds, dz)$  is a local martingale, Burkholder's inequality, see e.g. the discussion after theorem 4.74 in [24], with  $M_u^* := \sup_{0 \leq t \leq u} \left| \int_0^t \int_{\mathbb{R}_0} K_s(z) \tilde{N}(ds, dz) \right|^2$  gives for a constant  $c > 0$  that we have

$$E[|M_T^*|^2] \leq cE[[M]_T].$$

Here  $[\cdot]$  is the quadratic variation process. From example 1.29 in [16] we have

$$[M]_t = \int_0^t \int_{\mathbb{R}_0} |K_s(z)|^2 \nu(dz) ds + \int_0^t \int_{\mathbb{R}_0} |K_s(z)|^2 \tilde{N}(ds, dz).$$

Further, from page 62 in [14] or the discussion in section 4.3.2 in [2], we have that  $\int_0^t \int_{\mathbb{R}_0} |K_s(z)|^2 \tilde{N}(ds, dz)$  is a martingale. So

$$E[|M_T^*|^2] \leq cE\left[\int_0^T \int_{\mathbb{R}_0} |K_s(z)|^2 \nu(dz) ds\right] < \infty$$

and

$$\sup_{0 \leq t \leq T} \left| \int_t^T \int_{\mathbb{R}_0} K_s(z) \tilde{N}(ds, dz) \right|^2$$

is integrable. The same argument holds for  $\sup_{0 \leq t \leq T} \left| \int_t^T Z_s dW_s \right|$ . Thus,  $X \in S_{T, \beta}^2$ .  $\square$

### 3.1 The a priori estimates

Let  $\beta \geq 0$  and  $(f^i, \xi^i), i = 1, 2$  be two standard parameters, and suppose their respective BSDEs admit solutions  $(X^i, Z^i, K^i), i = 1, 2$ . We define

$$\begin{aligned} \delta X_t &= X_t^1 - X_t^2 \\ \delta Z_t &= Z_t^1 - Z_t^2 \\ \delta K_t(z) &= K_t^1(z) - K_t^2(z) \\ \delta_2 f_t &= f^1(t, X_t^2, Z_t^2, K_t^2) - f^2(t, X_t^2, Z_t^2, K_t^2) \\ \delta \xi &= \xi^1 - \xi^2, \end{aligned} \tag{6}$$

and for notational simplicity, we let

$$f_t^{i_1 i_2 i_3 i_4} = f^{i_1}(t, X_t^{i_2}, Z_t^{i_3}, K_t^{i_4}), \quad i_1, i_2, i_3, i_4 \in \{1, 2\}. \tag{7}$$

**Theorem 3.2.** *Under the assumptions over, for any  $\kappa > 0$ ,  $\lambda^2, \mu^2 > C$ , where  $C$  is the Lipschitz constant of  $f^1$ , and  $\beta \geq \kappa^2 + C(\lambda^2 + \mu^2 + 2)$ , then  $(X^i, Z^i, K^i)$  satisfies*

$$\begin{aligned}\|\delta X\|_{H_{T,\beta}^2}^2 &\leq T[e^{\beta T} E[|\delta\xi|^2] + \frac{1}{\kappa^2} \|\delta_2 f\|_{\beta}^2] \\ \|\delta Z\|_{H_{T,\beta}^2}^2 &\leq \frac{\lambda^2}{\lambda^2 - C} [e^{\beta T} E[|\delta\xi|^2] + \frac{1}{\kappa^2} \|\delta_2 f\|_{\beta}^2] \\ \|\delta K\|_{\hat{H}_{T,\beta}^2}^2 &\leq \frac{\mu^2}{\mu^2 - C} [e^{\beta T} E[|\delta\xi|^2] + \frac{1}{\kappa^2} \|\delta_2 f\|_{\beta}^2].\end{aligned}$$

*Proof.*  $\delta X_t$  is an Itô-Lévy process, so we can use Itô's formula, in the form of [16], on  $Y_t = e^{\beta t} \delta X_t^2$ . Doing this, we get

$$\begin{aligned}d(e^{\beta t} \delta X_t^2) &= \beta e^{\beta t} \delta X_t^2 dt + 2e^{\beta t} \delta X_t [\delta Z_t dW_t - (f_t^{1111} - f_t^{2222}) dt] + \\ &\quad e^{\beta t} \delta Z_t^2 dt + \int_{\mathbb{R}_0} e^{\beta t} \delta K_t(z)^2 \nu(dz) dt + \\ &\quad \int_{\mathbb{R}_0} e^{\beta t} [\delta K_t(z)^2 - 2\delta X_{t-} \delta K_t(z)] \tilde{N}(dt, dz),\end{aligned}$$

which implies

$$\begin{aligned}E[e^{\beta t} \delta X_t^2] &= E[e^{\beta T} \delta\xi^2] - \beta E\left[\int_t^T e^{\beta s} \delta X_s^2 ds\right] - \\ &\quad E\left[2 \int_t^T e^{\beta s} \delta X_s \delta Z_s dW_s\right] + \\ &\quad E\left[\int_t^T 2e^{\beta s} \delta X_s (f_s^{1111} - f_s^{2222}) ds\right] - \\ &\quad E\left[\int_t^T e^{\beta s} \delta Z_s^2 ds\right] - E\left[\int_t^T \int_{\mathbb{R}_0} e^{\beta s} \delta K_s(z)^2 \nu(dz) ds\right] - \\ &\quad E\left[\int_t^T \int_{\mathbb{R}_0} e^{\beta s} [\delta K_s(z)^2 - 2\delta X_{s-} \delta K_s(z)] \tilde{N}(ds, dz)\right].\end{aligned}\quad (8)$$

To get the result, we have to show that

$$\int_t^T e^{\beta s} \delta X_s \delta Z_s dW_s \quad (9)$$

$$\int_t^T \int_{\mathbb{R}_0} e^{\beta s} [\delta K_s(z)^2 - 2\delta X_{s-} \delta K_s(z)] \tilde{N}(ds, dz). \quad (10)$$

have mean equal to zero. To do this we need a lemma.

**Lemma 3.3.** *Let  $M$  be a local martingale satisfying  $E[[M]_t^{\frac{1}{2}}] < \infty$  for all  $0 \leq t \leq T$ . Then  $M$  is a martingale.*

*Proof.* Let  $M_s^* = \sup_{0 \leq t \leq s} |M_t|$ . Then Burkholder's inequality gives us

$$E[M_T^*] \leq KE[[M]_T^{\frac{1}{2}}] < \infty.$$

Now, let  $\sigma$  be a stopping time with  $\sigma \in [0, T]$ , and  $k \geq 0$ . Then

$$\int_{|M_{t \wedge \sigma}| > k} |M_{t \wedge \sigma}| dP \leq \int_{M_T^* > k} M_T^* dP.$$

Since the expression on the right is independent of  $\sigma$ , and goes to zero as  $k \rightarrow \infty$ , we conclude that  $\{M_{t \wedge \sigma}\}_{\{\sigma \in [0, T] \text{ stopping time}\}}$  is uniformly integrable, and we conclude from [15] exercise 1.5.19 (i) that  $M$  is a martingale.  $\square$

If we let  $\pi_t = e^{\beta t} \delta X_t \delta Z_t$ , we have from [15], proposition 3.2.24, that  $\Pi_t = \int_0^t \pi_s dW_s$  is a continuous local martingale and because  $2ab \leq a^2 + b^2$  for all  $a, b \in \mathbb{R}$ , we have

$$\begin{aligned} E \left[ \sqrt{\int_0^T |e^{\beta t} \delta X_t \delta Z_t|^2 dt} \right] &\leq e^{\beta T} E \left[ \sup_{0 \leq s \leq T} |\delta X_s| \sqrt{\int_0^T |\delta Z_t|^2 dt} \right] \leq \\ &2e^{\beta T} \left( E \left[ \sup_{0 \leq s \leq T} |\delta X_s|^2 \right] + E \left[ \int_0^T |\delta Z_t|^2 dt \right] \right) < \infty. \end{aligned}$$

So we conclude that  $E[[\Pi]_T^{\frac{1}{2}}] < \infty$ . Hence, lemma 3.3 gives that  $\Pi$  is a martingale, and  $E[\Pi_t] = E[\Pi_0] = 0$ . This proves that (9) has zero mean.

To show that (10) has zero mean, we get from [14] p.62, that if  $\theta : \Omega \times [0, T] \times \mathbb{R}_0 \rightarrow \mathbb{R}$  is predictable and satisfies  $E[\int_0^T \int_{\mathbb{R}_0} |\theta_t(z)| \nu(dz) dt] < \infty$ , then  $E[\int_0^T \int_{\mathbb{R}_0} |\theta_t(z)| \tilde{N}(dz, dt)] = 0$ . This is satisfied by  $e^{\beta t} \delta K_t(z)^2$  by assumption, and hence  $E[\int_t^T \int_{\mathbb{R}_0} e^{\beta s} \delta K_s(z)^2 \tilde{N}(ds, dz)] = 0$ . We notice further that if we define

$$Y_t = \int_0^t \int_{\mathbb{R}_0} \delta X_s - \delta K_s(z) \tilde{N}(ds, dz),$$

which is a local martingale, and has quadratic variation  $[Y]_t$  given by

$$[Y]_t = \int_0^t \int_{\mathbb{R}_0} \delta X_s^2 - \delta K_s(z)^2 N(ds, dz),$$

we have, for a constant  $c$ , by Hölder's inequality

$$\begin{aligned} E[[Y]_T^{\frac{1}{2}}] &= E \left[ \left( \int_0^T \int_{\mathbb{R}_0} \delta X_s^2 - \delta K_s(z)^2 N(ds, dz) \right)^{\frac{1}{2}} \right] \leq \\ &cE \left[ \sup_{0 \leq u \leq T} |X_u| \left( \int_0^t \int_{\mathbb{R}_0} \delta K_s(z)^2 N(ds, dz) \right)^{\frac{1}{2}} \right] \leq \\ &cE \left[ \sup_{0 \leq u \leq T} |X_u|^2 \right]^{\frac{1}{2}} E \left[ \int_0^t \int_{\mathbb{R}_0} \delta K_s(z)^2 N(ds, dz) \right]^{\frac{1}{2}} < \infty. \end{aligned}$$



Here, the finiteness follows from the same argumentation as in the proof of proposition 3.1. Therefore we have from lemma 3.3 that (10) has zero mean.

Now (8) is reduced to

$$\begin{aligned} & E[e^{\beta t} \delta X_t^2] + E \left[ \int_t^T e^{\beta s} \delta Z_s^2 ds \right] + \\ & \beta E \left[ \int_t^T e^{\beta s} \delta X_s^2 ds \right] + E \left[ \int_t^T \int_{\mathbb{R}} e^{\beta s} \delta K_s(z)^2 \nu(dz) ds \right] \leq \\ & E[e^{\beta T} \delta \xi^2] + E \left[ \int_t^T 2e^{\beta s} |\delta X_s| |(f_s^{1111} - f_s^{2222})| ds \right]. \end{aligned}$$

Now, let us split  $f$  in the following way:

$$\begin{aligned} |f_s^{1111} - f_s^{2222}| & \leq |f_s^{1111} - f_s^{1112}| + |f_s^{1112} - f_s^{1122}| \\ & \quad |f_s^{1122} - f_s^{1222}| + |\delta_2 f_s| \\ & \leq C(|\delta X_s| + |\delta Z_s| + \|\delta K_s\|) + |\delta_2 f_s|. \end{aligned}$$

Further, because  $\forall a, b \in \mathbb{R}$  and  $\lambda \neq 0$ ,  $2ab \leq \lambda^2 a^2 + \frac{1}{\lambda^2} b^2$ , we have the inequality

$$\begin{aligned} 2X[C(X + Z + K) + D] & \leq [\kappa^2 + C(\lambda^2 + \mu^2 + 2)]X^2 + \\ & \quad \frac{C}{\lambda^2}Z^2 + \frac{C}{\mu^2}K^2 + \frac{1}{\kappa^2}D^2 \end{aligned}$$

for  $X, Z, K, D \in \mathbb{R}$  and  $\lambda, \mu, \kappa > 0$ . Using this, we obtain

$$\begin{aligned} & E[e^{\beta t} \delta X_t^2] + E \left[ \int_t^T e^{\beta s} \delta Z_s^2 ds \right] + \\ & \beta E \left[ \int_t^T e^{\beta s} \delta X_s^2 ds \right] + E \left[ \int_t^T \int_{\mathbb{R}} e^{\beta s} \delta K_s(z)^2 \nu(dz) ds \right] \leq \\ & E[e^{\beta T} \delta \xi^2] + [\kappa^2 + C(\lambda^2 + \mu^2 + 2)]E \left[ \int_t^T e^{\beta s} \delta X_s^2 ds \right] + \\ & \frac{C}{\lambda^2}E \left[ \int_t^T e^{\beta s} \delta Z_s^2 ds \right] + \frac{C}{\mu^2}E \left[ \int_t^T \int_{\mathbb{R}} e^{\beta s} \delta K_s(z)^2 \nu(dz) ds \right] + \\ & \frac{1}{\kappa^2}E \left[ \int_t^T e^{\beta s} \delta_2 f_s^2 ds \right]. \end{aligned}$$

If we now choose  $\kappa > 0$ ,  $\lambda^2, \mu^2 > C$  and  $\beta \geq \alpha^2 + C(\lambda^2 + \mu^2 + 2)$  we get

$$\begin{aligned} & E[e^{\beta t} \delta X_t^2] + \frac{\lambda^2 - C}{\lambda^2}E \left[ \int_t^T e^{\beta s} \delta Z_s^2 ds \right] + \\ & \epsilon E \left[ \int_t^T e^{\beta s} \delta X_s^2 ds \right] + \frac{\mu^2 - C}{\mu^2}E \left[ \int_t^T \int_{\mathbb{R}} e^{\beta s} \delta K_s(z)^2 \nu(dz) ds \right] \leq \\ & E[e^{\beta T} \delta \xi^2] + \frac{1}{\kappa^2}E \left[ \int_t^T e^{\beta s} \delta_2 f_s^2 ds \right], \end{aligned}$$

where  $\epsilon = \beta - \kappa^2 + C(\lambda^2 + \mu^2 + 2) > 0$ . To finish the argument, we see that the a priori estimates for  $Z$  and  $K$  follows immediately. The estimate for  $X$  follows by integrating from 0 to  $T$  on both sides of the inequality.  $\square$

These a priori estimates will be used in almost all arguments in the thesis, either directly or indirectly. This is because we got the estimates from Itô's formula, the fundamental theorem of stochastic calculus. So when we apply the a priori estimates, we have actually done a part of the argument where we would have used Itô's formula.

Note that with the same argumentation as in proposition 3.1, using Burkholder's inequality and the a priori estimates, it is straight forward to show that there exists a  $c > 0$  such that

$$\|X\|_{S_{T,\beta}^2} \leq c \|X\|_{H_{T,\beta}^2}. \quad (11)$$

As mentioned, most arguments will depend on the a priori estimate in some form. So because of (11) we can prove convergence and similar arguments under the norm of  $H_T^2 \times H_T^2 \times \hat{H}_T^2$  instead of under the norm of  $S_T^2 \times H_T^2 \times \hat{H}_T^2$ .

### 3.2 The existence and uniqueness theorem

In the following comes the fundamental result existence result. Because the existence is proved by the contraction result, uniqueness will automatically follow. We will prove the result in  $H_{T,\beta}^2 \times H_{T,\beta}^2 \times \hat{H}_{T,\beta}^2$  with the norm

$$\|(X', Z', K')\|_{H_{T,\beta}^2 \times H_{T,\beta}^2 \times \hat{H}_{T,\beta}^2}^2 = \|X'\|_{H_{T,\beta}^2}^2 + \|Z'\|_{H_{T,\beta}^2}^2 + \|K'\|_{\hat{H}_{T,\beta}^2}^2$$

which is known from [4], theorem 6.2, to be a Banach space. But when  $(X', Z', K')$  is the solution of a BSDE, it is also contained in  $\mathcal{V}$ , as we have seen in proposition 3.1.

**Theorem 3.4.** *Let  $(f, \xi)$  be a standard parameter. Then the BSDE (4) has a unique solution  $(X, Z, K) \in \mathcal{V}$ .*

*Proof.* Let  $(X', Z', K') \in H_{T,\beta}^2 \times H_{T,\beta}^2 \times \hat{H}_{T,\beta}^2$ , and define

$$M_t = E \left[ \xi + \int_0^T f(s, X'_s, Z'_s, K'_s) ds \middle| \mathcal{F}_t \right].$$

Because  $(f, \xi)$  is a standard parameter,  $\xi + \int_0^T f(s, X'_s, Z'_s, K'_s) ds$  is square integrable and hence  $M$  is a square integrable martingale. Then we know from [26], lemma 2.3, that there exists  $Z \in H_\beta$ ,  $K \in \hat{H}_\beta$  such that

$$M_t = M(0) + \int_0^t Z_s dW_s + \int_0^t \int_{\mathbb{R}_0} K_s(z) \tilde{N}(ds, dz).$$

If we now define  $X_t = E[\xi + \int_t^T f(s, X'_s, Z'_s, K'_s) ds | \mathcal{F}_t]$ , we have

$$M_T - M_t = \xi + \int_t^T f(s, X'_s, Z'_s, K'_s) ds - X_t = \int_t^T Z_s dW_s + \int_t^T \int_{\mathbb{R}_0} K_s(z) \tilde{N}(ds, dz),$$

and thus

$$X_t = \xi + \int_t^T f(s, X'_s, Z'_s, K'_s) ds - \int_t^T Z_s dW_s + \int_t^T \int_{\mathbb{R}_0} K_s(z) \tilde{N}(ds, dz). \quad (12)$$

We have now defined a mapping

$$\Xi : H_{T,\beta}^2 \times H_{T,\beta}^2 \times \hat{H}_{T,\beta}^2 \rightarrow H_{T,\beta}^2 \times H_{T,\beta}^2 \times \hat{H}_{T,\beta}^2$$

by  $(X', Z', K') \rightarrow (X, Z, K)$  from the discussion above. Note that because of proposition 11, this is also a mapping

$$\Xi : H_{T,\beta}^2 \times H_{T,\beta}^2 \times \hat{H}_{T,\beta}^2 \rightarrow \mathcal{V}.$$

Now, because  $(H_{T,\beta}^2 \times H_{T,\beta}^2 \times \hat{H}_{T,\beta}^2, \|\cdot\|_{H_{T,\beta}^2 \times H_{T,\beta}^2 \times \hat{H}_{T,\beta}^2})$  is a Banach space we only need to show that  $\Xi$  is a contraction. Therefore, let  $\xi \in L_{T,\beta}^2$  and  $(X^i, Z^i, K^i) \in H_{T,\beta}^2 \times H_{T,\beta}^2 \times \hat{H}_{T,\beta}^2$  for  $i = 1, 2$ , and let  $(X^i, Z^i, K^i) = \Xi(X^i, Z^i, K^i)$  be defined as above for  $i = 1, 2$ . Then, from (12) we see that  $(X^i, Z^i, K^i)$  is a solution of the BDSE (4) with end point  $\xi$  and generator  $\tilde{f}^i(t, x, z, k) = f(t, X_t^i, Z_t^i, K_t^i)$  (and thus  $\tilde{f}^i$  is constant in the  $(x, z, k)$  parameter). Then  $(\tilde{f}^i, \xi)$  trivially satisfies the assumptions in definition 2.1 with the Lipschitz constant  $C = 0$  for  $i = 1, 2$ .

If we use the a priori estimates from theorem 3.2 with  $C = 0$ ,  $\beta = \kappa^2$ , then

$$\begin{aligned} \|X^1 - X^2\|_\beta^2 &\leq \frac{T}{\beta} E \left[ \int_0^T e^{\beta s} |f(s, X_s^1, Z_s^1, K_s^1) - f(s, X_s^2, Z_s^2, K_s^2)|^2 ds \right] \\ \|Z^1 - Z^2\|_\beta^2 &\leq \frac{1}{\beta} E \left[ \int_0^T e^{\beta s} |f(s, X_s^1, Z_s^1, K_s^1) - f(s, X_s^2, Z_s^2, K_s^2)|^2 ds \right] \\ \|K^1 - K^2\|_\beta^2 &\leq \frac{1}{\beta} E \left[ \int_0^T e^{\beta s} |f(s, X_s^1, Z_s^1, K_s^1) - f(s, X_s^2, Z_s^2, K_s^2)|^2 ds \right], \end{aligned}$$

and because  $f$  is Lipschitz with constant  $C_f$ , this means that

$$\begin{aligned} &\| (X^1 - X^2, Z^1 - Z^2, K^1 - K^2) \|_{H_{T,\beta}^2 \times H_{T,\beta}^2 \times \hat{H}_{T,\beta}^2}^2 = \\ &\|X^1 - X^2\|_\beta^2 + \|Z^1 - Z^2\|_\beta^2 + \|K^1 - K^2\|_\beta^2 \leq \\ &\frac{3C_f^2(2+T)}{\beta} \| (X^1 - X^2, Z^1 - Z^2, K^1 - K^2) \|_{\mathcal{V}_\beta}^2. \end{aligned}$$

By choosing  $\beta > 3C_f^2(T+2)$  we get the desired property of  $\Xi$ , and the BSDE (4) must have a solution which is unique under the norm  $\|\cdot\|_{H_{T,\beta}^2 \times H_{T,\beta}^2 \times \hat{H}_{T,\beta}^2}$  for all  $\beta > 3C_f^2(T+2)$ . Since the  $\beta$ -norms are equivalent, this must hold for all  $\beta \geq 0$ . So because the range of  $\Xi$  is  $\mathcal{V}$ , the theorem is proved.  $\square$

As in [11], we can make a Picard iterative scheme that converges to the solution of the BSDE.

**Proposition 3.5.** *Suppose a sequence in  $\mathcal{V}$  is defined by  $(X^0, Z^0, K^0)$  and  $(X^{k+1}, Z^{k+1}, K^{k+1})$  as the solution of*

$$\begin{aligned} dX_t^{k+1} &= -f(t, X_t^k, Z_t^k, K_t^k)dt + Z_t^{k+1}dW_t + \int_{\mathbb{R}_0} K_t^{k+1}(z)\tilde{N}(dt, dz) \\ X_T^{k+1} &= \xi. \end{aligned}$$

Then  $(X^k, Z^k, K^k)$  converges in  $\mathcal{V}$  to  $(X, Z, K)$ , where  $(X, Z, K)$  is the solution to the BSDE corresponding to  $(f, \xi)$ . Further, we have the following inequalities for all  $\beta > 6C_f^2(T+2)$ , where  $C$  is the Lipschitz constant of  $f$  and the norm used for  $(X, Z, K)$  is  $\|\cdot\|_{H_{T,\beta}^2 \times H_{T,\beta}^2 \times \hat{H}_{T,\beta}^2}$ :

$$\|(X^k - X, Z^k - Z, K^k - K)\|_{\beta}^2 \leq \left(\frac{3C_f^2(T+2)}{\beta}\right)^k \|(X, Z, K)\|_{\beta}^2$$

$$\begin{aligned} \|(X^k, Z^k, K^k)\|_{\beta}^2 &\leq \frac{\beta(T+2)}{\beta - 6C_f^2(T+2)} \|\xi\|_{\beta}^2 + \\ &\quad \frac{2(T+2)}{\beta - 6C_f^2(T+2)} \|f(\cdot, 0, 0, 0)\|_{\beta}^2. \end{aligned}$$

*Proof.* Note that the first inequality verifies the convergence. Now, both inequalities are easily verified by using the a priori estimates recursively. In the first inequality, we use  $f^1 = f(\cdot, X^{k-1}, Z^{k-1}, K^{k-1})$ ,  $f^2 = f$ . Because  $f^1$  has Lipschitz constant 0, we choose  $\lambda = \mu = 0$  and  $\kappa^2 = \beta$ . Then we get

$$\begin{aligned} \|(X^k - X, Z^k - Z, K^k - K)\|_{\beta}^2 &\leq \\ &\quad \frac{T+2}{\beta} \|f(\cdot, X^{k-1}, Z^{k-1}, K^{k-1}) - f(\cdot, X, Z, K)\|_{\beta}^2 \leq \\ &\quad \frac{3C_f^2(T+2)}{\beta} \|(X^{k-1} - X, Z^{k-1} - Z, K^{k-1} - K)\|_{\beta}^2, \end{aligned}$$

and the first inequality follows by recursion.

To prove the second inequality, we use  $f^1 = 0$  and  $f^2 = f(\cdot, X^{k-1}, Z^{k-1}, K^{k-1})$  with  $\lambda = \mu = 0$  and  $\kappa^2 = \beta$ . Then

$$\begin{aligned} \|(X^k, Z^k, K^k)\|_\beta^2 &\leq (T+2) \|\xi\|_\beta^2 + \frac{T+2}{\beta} \|f(\cdot, X^{k-1}, Z^{k-1}, K^{k-1})\|_\beta^2 \leq \\ &(T+2) \|\xi\|_\beta^2 + \frac{2(T+2)}{\beta} \|f(\cdot, 0, 0, 0)\|_\beta^2 + \\ &\frac{2(T+2)}{\beta} \|f(\cdot, X^{k-1}, Z^{k-1}, K^{k-1}) - f(\cdot, 0, 0, 0)\|_\beta^2 \leq \\ &(T+2) \|\xi\|_\beta^2 + \frac{2(T+2)}{\beta} \|f(\cdot, 0, 0, 0)\|_\beta^2 + \\ &\frac{6(T+2)C_f^2}{\beta} \|(X^{k-1}, Z^{k-1}, K^{k-1})\|_\beta^2. \end{aligned}$$

Again, the result follows by a simple recursion argument.  $\square$

## 4 Linear BSDEs with jumps

In this section I will discuss a special class of standard parameters, where the generator has a linear shape, and we will find an explicit solution to  $X$ . This was maybe one of the most fascinating result from the BSDE course MAT4760 because of its simplicity, and my supervisor told me this worked in the jump case.

However, when I worked with this result I had not found any articles providing a proof. So this proof is my own in the sense that I have investigated the generalizability of the argumentation in [11]. This is possible, but in [11] the theorem is easily proven because a term trivially is a martingale. Here, much more argumentation is needed to conclude that this term is a martingale. Later in the process I have found this proven for jumps. One example is [12]. It probably is a quite obvious result, but I did this early in the process, and at this stage quite a bit of research had to be done.

### 4.1 Explicit formula of linear BSDEs

**Theorem 4.1.** *Suppose  $(X, Z, K) \in \mathcal{V}$  satisfies*

$$\begin{aligned} dX_t &= - \left[ \phi_t + \beta_t X_t + \gamma_t Z_t + \int_{\mathbb{R}_0} \Psi_t(z) K_t(z) \nu(dz) \right] dt + \\ &Z_t dW_t + \int_{\mathbb{R}_0} K_t(z) \tilde{N}(dt, dz), \end{aligned}$$

where  $\phi \in H_T^2$ ,  $\beta$  and  $\gamma$  are predictable and bounded, and  $\Psi$  is predictable and satisfies

$$D_2(1 \wedge |z|) \leq \Psi_s(z) \leq D_1(1 \wedge |z|), \mathcal{P}' - a.e. \quad (13)$$

for a  $D_1 \geq 0$  and a  $D_2 \in (-1, 0]$ . Then  $X$  attains the following representation

$$X_t = E \left[ \xi \Gamma_T^t + \int_t^T \Gamma_s^t \phi_s ds \middle| \mathcal{F}_t \right], \quad (14)$$

where  $\Gamma$  is a process which satisfies

$$\begin{aligned} d\Gamma_s^u &= \Gamma_s^u [\beta_s ds + \gamma_s dW_s + \int_{\mathbb{R}_0} \Psi_s(z) \tilde{N}(ds, dz)] & u < s \leq T \\ \Gamma_s^u &= 1 & 0 \leq s \leq u \end{aligned} \quad (15)$$

for  $u \in [0, T]$ .

Some observations are needed before we prove this theorem. First, from calculus we have that the condition (13) on  $\Psi$  gives us the following inequalities

$$c_2(1 \wedge |z|) \leq \ln(1 + \Psi_t(z)) \leq c_1(1 \wedge |z|) \quad (16)$$

$$c_3(1 \wedge |z|^2) \leq \ln(1 + \Psi_t(z)) - \Psi_t(z) \leq 0 \quad (17)$$

for some constants  $c_i \in \mathbb{R}$ ,  $i = 1, 2, 3$ . In particular, this holds because we have chosen conditions on  $\Psi$  not to be arbitrary close to  $-1$ .

Second, we have from example 9.6 in [10] that  $\Gamma$  has an explicit solution given by

$$\begin{aligned} \Gamma_t^u &= \exp \left\{ \int_u^t [\beta_s - \frac{\gamma_s^2}{2}] ds + \int_u^t \gamma_s dW_s + \right. \\ &\quad \left. \int_u^t \int_{\mathbb{R}_0} [\ln(1 + \Psi_s(z)) - \Psi_s(z)] \nu(dz) ds + \right. \\ &\quad \left. \int_u^t \int_{\mathbb{R}_0} \ln(1 + \Psi_s(z)) \tilde{N}(ds, dz) \right\} \end{aligned}$$

where (16) and (17) assure us that the integrals in the exponent are well defined. For simplicity, we write  $\Gamma_t = \Gamma_t^0$ . Notice that we trivially have  $\Gamma_t^u = \frac{\Gamma_t}{\Gamma_u}$ .

Finally, we see that the linear BSDE here defined, trivially satisfies the assumptions in definition 2.1. In fact we have a quasi-strong standard parameter, so we have a unique solution  $(X, Z, K) \in \mathcal{V}$ .

*Proof of theorem 4.1.* By applying Itô's formula on  $X_t\Gamma_t$ , we get

$$\begin{aligned}
d(X_t\Gamma_t) &= X_t\Gamma_t[\beta_t dt + \gamma_t dW_t] - \\
&\Gamma_t[\phi_t + \beta_t X_t + \gamma_t Z_t + \int_{\mathbb{R}_0} \Psi_t(z)K_t(z)\nu(dz)]dt + \\
&\Gamma_t Z_t dW_t + \Gamma_t Z_t \gamma_t dt + \Gamma_t \int_{\mathbb{R}_0} \Psi_t(z)K_t(z)\nu(dz)dt + \\
&\int_{\mathbb{R}_0} [K_t(z)\Gamma_{t-} + X_{t-}\Gamma_{t-}\Psi_t(z) + K_t(z)\Gamma_{t-}\Psi_t(z)]\tilde{N}(dt, dz) \\
&= -\Gamma_t\phi_t dt + X_t\Gamma_t\gamma_t dW_t + \Gamma_t Z_t dW_t + \\
&\int_{\mathbb{R}_0} [K_t(z)\Gamma_{t-} + X_{t-}\Gamma_{t-}\Psi_t(z) + \Gamma_{t-}K_t(z)\Psi_t(z)]\tilde{N}(dt, dz).
\end{aligned}$$

We see that  $X_t\Gamma_t + \int_0^t \Gamma_s\phi_s ds$  is a local martingale. We will show that it actually is a martingale, and hence is the conditional expectation of its end point. This will imply that

$$\begin{aligned}
X_t\Gamma_t + \int_0^t \Gamma_s\phi_s ds &= E[X_T\Gamma_T + \int_0^T \Gamma_s\phi_s ds | \mathcal{F}_t] && \Rightarrow \\
X_t\Gamma_t &= E[X_T\Gamma_T + \int_t^T \Gamma_s\phi_s ds | \mathcal{F}_t] && \Rightarrow \\
X_t &= E[\xi\Gamma_T^t + \int_t^T \Gamma_s^t\phi_s ds | \mathcal{F}_t],
\end{aligned}$$

and the theorem will be proved.

From [24], theorem 1.51, we have that a sufficient condition for  $X_t\Gamma_t + \int_0^t \Gamma_s\phi_s ds$  to be a proper martingale is that  $E[\sup_{0 \leq t \leq T} |X_t\Gamma_t + \int_0^t \Gamma_s\phi_s ds|] < \infty$ . If  $E[\sup_{0 \leq t \leq T} |\Gamma_t|^2] < \infty$ , this condition is satisfied. To see this, remember that  $\bar{X} \in S_T^2$  by assumption and therefore

$$\begin{aligned}
E[\sup_{0 \leq t \leq T} |X_t\Gamma_t|] &\leq E[\sup_{0 \leq t \leq T} |X_t| \sup_{0 \leq u \leq T} |\Gamma(u)|] \\
&\leq E[\sup_{0 \leq t \leq T} |X_t|^2]^{\frac{1}{2}} E[\sup_{0 \leq t \leq T} |\Gamma_t|^2]^{\frac{1}{2}} < \infty
\end{aligned}$$

by Hölder's inequality. Further, we have

$$E\left[\sup_{0 \leq t \leq T} \left|\int_0^t \Gamma_s\phi_s ds\right|\right] \leq T^{\frac{1}{2}} E\left[\int_0^T \phi_s^2 ds\right]^{\frac{1}{2}} E\left[\sup_{0 \leq t \leq T} |\Gamma_t|^2\right]^{\frac{1}{2}} < \infty$$

by Hölder's inequality and Jensen's inequality. So to complete the proof, we will show  $E[\sup_{0 \leq t \leq T} |\Gamma_t|^2] < \infty$ .

With  $u = 0$  the equation (15) attains the following explicit solution

$$\begin{aligned} \Gamma_t = \exp \left\{ \int_0^t [\beta_s - \frac{\gamma_s^2}{2}] ds + \int_0^t \gamma_s dW_s + \right. \\ \left. \int_0^t \int_{\mathbb{R}_0} [\ln(1 + \Psi_s(z)) - \Psi_s(z)] \nu(dz) ds + \right. \\ \left. \int_0^t \int_{\mathbb{R}_0} \ln(1 + \Psi_s(z)) \tilde{N}(ds, dz) \right\}. \end{aligned} \quad (18)$$

Notice that example 9.6 in [10] also gives that if  $\beta \equiv 0$  then the expression on the right in (18) is a local martingale. Therefore,  $\Gamma_t$  is equal to the bounded process  $\exp\{\int_0^t \beta_s ds\}$  multiplied by the local martingale  $V$  given by

$$\begin{aligned} V_t = \exp \left\{ - \int_0^t \frac{\gamma_s^2}{2} ds + \int_0^t \gamma_s dW_s + \right. \\ \left. \int_0^t \int_{\mathbb{R}_0} [\ln(1 + \Psi_s(z)) - \Psi_s(z)] \nu(dz) ds + \right. \\ \left. \int_0^t \int_{\mathbb{R}_0} \ln(1 + \Psi_s(z)) \tilde{N}(ds, dz) \right\}, \end{aligned}$$

or equivalently

$$\begin{aligned} dV_t &= V_{t-} [\gamma_t dW_t + \int_{\mathbb{R}_0} \Psi_t(z) \tilde{N}(dt, dz)] \\ V_t &= 1. \end{aligned}$$

Therefore, from [16], example 1.29, the quadratic variation of  $V$  is given by

$$\begin{aligned} [V]_t &= \int_0^t V_s^2 \gamma_s^2 ds + \int_0^t \int_{\mathbb{R}_0} V_s^2 \Psi_s(z)^2 \nu(dz) ds + \\ &\int_0^t \int_{\mathbb{R}_0} V_{s-}^2 \Psi_s(z)^2 \tilde{N}(ds, dz). \end{aligned}$$



So because  $\exp\{\int_0^t \beta_s ds\}$  is bounded, Burkholder's inequality gives

$$\begin{aligned}
E[\sup_{0 \leq t \leq T} \Gamma_t^2] &\leq C_0 E[\sup_{0 \leq t \leq T} V_t^2] \leq C_1 E[[V]_T] = \\
&C_1 E\left[\int_0^T V_{t^-}^2 [\gamma_t^2 + \int_{\mathbb{R}_0} \Psi_t(z)^2 \nu(dz)] dt\right] + \\
&C_1 E\left[\int_0^T \int_{\mathbb{R}_0} V_{t^-}^2 \Psi_t(z)^2 \tilde{N}(dt, dz)\right] \\
&\leq C_1 E\left[\int_0^T V_{t^-}^2 [C_2 + \int_{\mathbb{R}_0} C_3 |z|^2 \nu(dz)] dt\right] + \\
&C_1 E\left[\int_0^T \int_{\mathbb{R}_0} V_{t^-}^2 \Psi_t(z)^2 \tilde{N}(dt, dz)\right] \\
&\leq C_4 E\left[\int_0^T V_t^2 dt\right] + C_1 E\left[\int_0^T \int_{\mathbb{R}_0} V_{t^-}^2 \Psi_t(z)^2 \tilde{N}(dt, dz)\right]
\end{aligned}$$

for some constants  $C_i > 0$ ,  $i = 0, 1, 2, 3, 4$ . To control this expression, we rewrite  $V^2$  in the following way

$$\begin{aligned}
V_t^2 &= \exp\left\{-\int_0^t \gamma_s^2 ds + \int_0^t 2\gamma_s dW_s + \right. \\
&\quad \left. \int_0^t \int_{\mathbb{R}_0} [\ln((1 + \Psi_s(z))^2) - 2\Psi_s(z)] \nu(dz) ds + \right. \\
&\quad \left. \int_0^t \int_{\mathbb{R}_0} \ln((1 + \Psi_s(z))^2) \tilde{N}(ds, dz)\right\} \\
&= \exp\left\{\int_0^t \gamma_s^2 ds + \int_0^t \int_{\mathbb{R}_0} \Psi_s(z)^2 \nu(dz) ds + \right. \\
&\quad \left. \int_0^t 2\gamma_s dW_s - \int_0^t 2\gamma_s^2 ds + \right. \\
&\quad \left. \int_0^t \int_{\mathbb{R}_0} (\ln(1 + [(1 + \Psi_s(z))^2 - 1]) - \right. \\
&\quad \left. [(1 + \Psi_s(z))^2 - 1]) \nu(dz) ds + \right. \\
&\quad \left. \int_0^t \int_{\mathbb{R}_0} \ln(1 + [(1 + \Psi_s(z))^2 - 1]) \tilde{N}(ds, dz)\right\} \\
&= \exp\left\{\int_0^t \gamma_s^2 du + \int_0^t \int_{\mathbb{R}_0} \Psi_s(z)^2 \nu(dz) ds\right\} U_t
\end{aligned}$$

where

$$\exp\left\{\int_0^t \gamma_s^2 du + \int_0^t \int_{\mathbb{R}_0} \Psi_s(z)^2 \nu(dz) ds\right\}$$

is by (13) a bounded process and  $U$  satisfies

$$dU_t = U_{t-} \left[ 2\gamma_t dW_t + \int_{\mathbb{R}_0} [(1 + \Psi_t(z))^2 - 1] \tilde{N}(dt, dz) \right]$$

$$U(0) = 1,$$

which, from [16], theorem 1.36, is a martingale if the following Novikov condition for Lévy processes is satisfied

$$E \left[ \exp \left\{ \frac{1}{2} \int_0^T (2\gamma_t)^2 dt + \int_0^T \int_{\mathbb{R}_0} 2(1 + \Psi_t(z))^2 \ln(1 + \Psi_t(z)) + 1 - (1 + \Psi_t(z)) \nu(dz) dt \right\} \right] < \infty.$$

If we study the function

$$h(x) = 2(1+x)^2 \ln(1+x) + 1 - (1+x)^2$$

using standard calculus techniques, we obtain that there exists a  $1 > \epsilon > 0$  and a constant  $c_\epsilon$  such that for all  $x \in (-\epsilon, \epsilon)$  we have

$$0 \leq h(x) \leq c_\epsilon |x|^2.$$

This means that there exists a  $c_\epsilon^2$  such that for all  $z \in (-\epsilon, \epsilon) \setminus \{0\}$  we have that for a.a.  $t \in [0, T]$  that  $0 \leq h(\Psi_t(z)) \leq c_\epsilon^2 |z|^2$ . Then we also have that the exponent in the Novikov condition is bounded, and hence the Novikov condition is satisfied and  $U$  is a martingale.

We are now ready to complete our proof. We had

$$E \left[ \sup_{0 \leq t \leq T} \Gamma_t^2 \right] \leq C_4 E \left[ \int_0^T V_t^2 dt \right] + C_1 E \left[ \int_0^T \int_{\mathbb{R}_0} V_{t-}^2 \Psi_t(z)^2 \tilde{N}(dt, dz) \right].$$

Now we also know for a constant  $c_V$

$$E \left[ \int_0^T V_t^2 dt \right] \leq c_V E \left[ \int_0^T U_t dt \right] =$$

$$c_V \int_0^T E[U_t] dt = c_V T E[U_0] < \infty.$$

We also know

$$E \left[ \int_0^T \int_{\mathbb{R}_0} V_{t-}^2 \Psi_t(z)^2 \tilde{N}(dt, dz) \right] = 0$$

because for constants  $c_V^2, c_V^3$  we have

$$E \left[ \int_0^T \int_{\mathbb{R}_0} V_{t-}^2 \Psi_t(z)^2 \nu(dz) dt \right] \leq c_V^2 E \left[ \int_0^T \int_{\mathbb{R}_0} U_{t-} |z|^2 \nu(dz) dt \right] \leq$$

$$c_V^3 E \left[ \int_0^T U_{t-} dt \right] = c_V^3 \int_0^T E[U_t] dt < \infty.$$

The last inequality is valid because  $U$  only have a countable number of jumps. So  $E[\sup_{0 \leq t \leq T} \Gamma_t^2] < \infty$  and the theorem is proved.  $\square$

## 4.2 The importance of linear BSDEs

In the discussion of the importance of studying BSDEs in section 2.3 we saw that the BSDE equivalent to a replicable claim is a linear one. Because the last “constant” element  $\phi$  only needs the  $L^2$  requirement, the linear case can be very flexible. One example, which is also mentioned in section 2.3, is this the method of penalization. Here,  $\phi$  is the penalization term. Even though the article spoken of, that is, [7], does not use the explicit representation, it is imaginable that the explicit representation may give rise to useful inequalities in different settings. This is because  $\phi$  can be dependent on processes.

One application we will see, and which is an example of how a BSDE may be written as a linear BSDE, is the comparison theorem. We will investigate this in the next chapter.

## 5 Comparison of solutions

In this section we will apply the closed formula for a linear BSDE to see that if the strong standard parameter  $(f^1, \xi^1)$  and the standard parameter (not necessarily strong)  $(f^2, \xi^2)$  satisfies  $\xi^1 \geq \xi^2$  and  $f^1 \geq f^2$  in some sense, then we must have  $X_t^1 \geq X_t^2$ ,  $\Omega \times \mathcal{B}([0, T])$ -a.e.

This was proven in [25], which is one of the first articles on BSDEs with jumps I found. Here it is proven by a change of measure. But having done the effort to prove the linear case theorem, a much easier proof is done here. That is, the idea in theorem 2.2. in [11] works well with jumps. However, as in [25] we can not prove it for the general case. We need one of the standard parameters to be strong.

### 5.1 The comparison theorem

**Theorem 5.1.** *Let  $(f^1, \xi^1)$  be a strong standard parameter and  $(f^2, \xi^2)$  be a standard parameter, and let their solutions be denoted by  $(X^1, Z^1, K^1)$  and  $(X^2, Z^2, K^2)$ . Using the same notation as in (6) and (7), suppose that*

$$\begin{aligned} \xi^1 &\geq \xi^2 && a.s. \\ \delta_2 f_t &\geq 0 && \Omega \times \mathcal{B}([0, T]) - a.e. \end{aligned}$$

*Then we have  $X_t^1 \geq X_t^2$ ,  $\Omega \times \mathcal{B}([0, T])$  - a.e.*

*Proof.* We split  $f$  in a slightly more complex manner as in the proof of

theorem 3.2 by defining the processes

$$\begin{aligned} A_t &= \frac{f_t^{1,1,1,1} - f_t^{1,2,1,1}}{\delta X_t} \mathbf{1}\{\delta X_t \neq 0\} \\ B_t &= \frac{f_t^{1,2,1,1} - f_t^{1,2,2,1}}{\delta Z_t} \mathbf{1}\{\delta Z_t \neq 0\} \\ C_t &= \frac{f_t^{1,2,2,1} - f_t^{1,2,2,2}}{\int_{\mathbb{R}_0} \Psi_t(z) \delta K_t(z) \nu(dz)} \mathbf{1}\left\{ \int_{\mathbb{R}_0} \Psi_t(z) \delta K_t(z) \nu(dz) \neq 0 \right\}, \end{aligned}$$

where  $\Psi$  is the function from definition 5 for  $f^1$ . It follows easily from the Lipschitz condition on  $f^1$  that these are bounded functions, where  $C$  is bounded from below by  $-1$ . We easily see that  $\delta X_t$  has the following differential form

$$\begin{aligned} d(\delta X_t) &= - \left[ A_t \delta X_t + B_t \delta Z_t + C_t \int_{\mathbb{R}_0} \Psi_t(z) \delta K_t(z) \nu(dz) + \delta_2 f_t \right] dt + \\ &\quad \delta Z_t dW_t + \int_{\mathbb{R}_0} \delta K_t(z) \tilde{N}(dt, dz) \\ \delta X_t &= \xi. \end{aligned}$$

Because  $C\Psi \geq D_2 > -1$  uniformly, this BSDE satisfies the conditions of theorem 4.1 and  $\delta X$  obtains the representation

$$\delta X_t = E \left[ \Gamma_T^t \delta \xi + \int_t^T \Gamma_s^t \delta_2 f_s ds \middle| \mathcal{F}_t \right],$$

where

$$\begin{aligned} d\Gamma_t^u &= \Gamma_{t-} [A_t dt + B_t dW_t + \int_{\mathbb{R}_0} C_t \Psi_t(z) \nu(dz) dt] \\ \Gamma_u^u &= 1. \end{aligned}$$

All the terms in the representation (14) are positive, and hence we have  $\delta X_t \geq 0$ ,  $dP \times dt$ -a.e., and the theorem is proved.  $\square$

In [11], it is also proved that the comparison is strict. This is also the case for jumps. With this we mean:

**Proposition 5.2.** *Suppose  $(f^1, \xi^1)$  and  $(f^2, \xi^2)$  satisfies the conditions of theorem 5.1. In addition, suppose for a  $t \in [0, T]$  that we have  $X_t^1 = X_t^2$  a.s. on a set  $F \in \mathcal{F}_t$ . Then we have  $X_s^1 = X_s^2$  a.e. on  $[t, T] \times F$ ,  $\xi_1 = \xi_2$  a.s. on  $F$  and  $f^1(s, X_s^2, Z_s^2, K_s^2) = f^2(s, X_s^2, Z_s^2, K_s^2)$  a.e. on  $[t, T] \times F$ .*

*Proof.* This follows directly from the form of  $\delta X_t$ , namely

$$\delta X_t = E \left[ \Gamma_T^t \delta \xi + \int_t^T \Gamma_s^t \delta_2 f_s ds \middle| \mathcal{F}_t \right].$$

Again, all terms in the conditional expectation are positive, so the proposition follows from the classical measure theory. See e.g. corollary 4.10 in [4].  $\square$

In particular we have that a BSDE has solution with nonnegative  $X$  if  $\xi \geq 0$  a.s.,  $f(t, 0, 0, 0) \geq 0$  a.s. and  $X_0 \geq 0$  a.s. If the inequalities are equalities, then  $X = 0$  a.s.

## 5.2 Why we need a strong generator

To complete the topic of comparison, let us investigate a little deeper why we need one generator to be strong and not only quasi-strong. The answer is found in [24], theorem 2.37, which states that if  $V$  is a general semimartingale with  $V_0 = 0$ , then there exists a semimartingale  $W$  such that

$$W_t = 1 + \int_0^t W_{s-} dV_s, \quad (19)$$

where  $W$  obtains the representation

$$W_t = \exp \left\{ V_t - \frac{1}{2} [V]_t \right\} \prod_{0 < s \leq t} (1 + \Delta V_s) \exp \left\{ -\Delta X_s + \frac{1}{2} (\Delta V_s)^2 \right\}.$$

It is well known that an Itô-Lévy process is a semimartingale, so if we let

$$dV_t = \beta_t dt + \gamma_t dW_t + \int_{\mathbb{R}_0} \Psi_t(z) \tilde{N}(dt, dz)$$

$$V_T = 0;$$

with  $\beta$ ,  $\gamma$  and  $\Psi$  as defined in theorem 4.1, we see that  $\Gamma = \Gamma^0$  in (15) satisfies

$$\Gamma_t = 1 + \int_0^t \Gamma_{s-} dV_s.$$

That is, it is on the form of (19). Now, because the  $dt$  and  $dW_t$  terms vanish in  $\Delta V_s$  we have

$$\Delta V_s = V_s - V_{s-} = \Psi(s, \Delta \eta_s).$$

See e.g. [2], corollary 4.4.9. for verification. So if  $\Psi$  is not bounded below by  $-1$ ,  $(1 + \Delta V_s)$  can be negative and thus  $\Gamma$  is not in general positive. The extra condition for a quasi-strong generator to be strong is precisely what is needed for  $C_t$  in the proof of theorem 5.1 to be bounded below by  $-1$ .

## 6 Dependence on a parameter

In this section I will consider a family of standard parameters  $\{f^\alpha(\cdot), \xi^\alpha, \alpha \in \mathbb{R}\}$ . Let  $(X^\alpha, Z^\alpha, K^\alpha)$  be their respective solutions. I will show that if the standard parameters depends “nicely” on the parameter  $\alpha$ , then the solutions also will depend “nicely” on  $\alpha$ .

As in [11], “nicely” will mean continuity and differentiability. Continuity is proven more or less the same way as in [11]. I state it here, because we need it in the section about differentiability. The differentiability is proved also using the same basic idea as in [11], but the notion of Fréchet differentiability must be introduced. I will also look at differentiation when  $f^\alpha$  is quasi-strong.

One remark to this section is that many assumptions in  $\alpha$  will be given globally. Many of these could easily be generalized to only apply locally. I chose my formulation to be concise with [11] and because I think local conditions are not essentially stronger. The proofs will be the same.

In this section, all convergence results will be for  $\mathcal{V}$  but will be proved in  $H_T^2 \times H_T^2 \times \hat{H}_T^2$ . This is no problem because the size of the constant  $c$  in (11) depends on  $(f, \xi)$ , and when the bounds on  $(f^\alpha, \xi^\alpha)$  are uniform, we can find a  $c$  that is valid for all  $\alpha$ , or as we discussed, valid locally in  $\alpha$ .

Potential applications will be given in section 8 after we have discussed the Malliavin derivative and the combination of differentiation in the parameter and in the Malliavin sense.

### 6.1 The continuous case

Make the following assumptions about a family  $\{f^\alpha(\cdot), \xi^\alpha, \alpha \in \mathbb{R}\}$  of standard parameters:

**Assumption 6.1.** *The family  $\{f^\alpha(\cdot), \alpha \in \mathbb{R}\}$  is equi-Lipschitz, in the sense that there exist a  $C > 0$  such that if  $C_\alpha$  is the Lipschitz constant of  $f^\alpha$ , then  $C_\alpha \leq C \forall \alpha \in \mathbb{R}$ .*

**Assumption 6.2.**  *$\alpha \rightarrow (f^\alpha(\cdot), \xi^\alpha)$  is continuous in the sense that if  $\alpha_0 \in \mathbb{R}$ , then we have that as  $\alpha \rightarrow \alpha_0$*

$$\|f^\alpha(\cdot, X^{\alpha_0}, Z^{\alpha_0}, K^{\alpha_0}) - f^{\alpha_0}(\cdot, X^{\alpha_0}, Z^{\alpha_0}, K^{\alpha_0})\|_{H_T^2} \rightarrow 0$$

$$\|\xi^\alpha - \xi^{\alpha_0}\|_{L_T^2} \rightarrow 0.$$

**Theorem 6.3.** *Let the assumptions 6.1 and 6.2 hold. Then we have*

$$\|(X^\alpha, Z^\alpha, K^\alpha) - (X^{\alpha_0}, Z^{\alpha_0}, K^{\alpha_0})\|_{\mathcal{V}} \rightarrow 0,$$

as  $\alpha \rightarrow \alpha_0$ .

*Proof.* This is a direct consequence of the a priori estimates from theorem 3.2. We choose  $\kappa > 0$ ,  $\mu^2, \lambda^2 > C$  and  $\beta \geq \kappa + C(\lambda^2 + \mu^2 + 2)$  and obtain

$$\begin{aligned} & \| (X^\alpha, Z^\alpha, K^\alpha) - (X^{\alpha_0}, Z^{\alpha_0}, K^{\alpha_0}) \|_\beta^2 \leq \\ & [T + \frac{\lambda^2}{\lambda^2 - C} + \frac{\mu^2}{\mu^2 - C}] \left( \| \xi^\alpha - \xi^{\alpha_0} \|_\beta^2 + \right. \\ & \left. \frac{1}{\kappa^2} \| f^\alpha(\cdot, X^{\alpha_0}, Z^{\alpha_0}, K^{\alpha_0}) - f^{\alpha_0}(\cdot, X^{\alpha_0}, Z^{\alpha_0}, K^{\alpha_0}) \|_{H_{T,\beta}^2}^2 \right) \end{aligned}$$

which by assumption has limit 0 as  $\alpha \rightarrow \alpha_0$ . The theorem then follows from (11).  $\square$

## 6.2 The differentiable case

The goal of this section is to show a differentiability result for a standard parameter which satisfies some differentiability conditions such that the solution of a limit argument coincides with using the classical chain rule on the BSDE. That is, to formally differentiate in  $\alpha$ .

We will show this for a general standard parameter under some conditions, and for this reason we need a way to differentiate the generator  $f^\alpha(t, x, z, k)$  in the parameter  $k$ . To do this we notice that  $k$  is an element in the Banach space  $\hat{L}_T^2(\mathbb{R}_0, \mathcal{B}(\mathbb{R}_0), \nu)$ . So a natural derivative for our purpose is the so-called Fréchet derivative.

We define the Fréchet derivative in the following way. Let  $A$  and  $B$  be two Banach spaces with norms  $\| \cdot \|_A$  and  $\| \cdot \|_B$  respectively. We say that a function  $h : A \rightarrow B$  is Fréchet differentiable in a point  $a_0 \in A$  if there exists a bounded linear operator  $l : A \rightarrow B$  such that

$$\frac{\| h(a) - h(a_0) - l(a - a_0) \|_B}{\| a - a_0 \|_A} \rightarrow 0$$

as  $a \rightarrow a_0$  in the topology induced by  $\| \cdot \|_A$ . In this case we denote the derivative by  $l = \nabla h(a_0)$ . Of course, if  $h$  is Fréchet differentiable in each point  $a_0 \in A$ , we say that  $h$  is Fréchet differentiable. In this case, it is worth noticing that for each  $a_0 \in A$ ,  $\nabla h(a_0)$  defines a linear operator. So if  $v \in A$ , then  $l(v)$  in the definition is given by  $\nabla h(a_0)(v)$ . Further, a linear operator  $l : A \rightarrow B$  is bounded if there exists a  $C > 0$  such that for all  $v \in A$ ,  $\| l(v) \|_B \leq C \| v \|_A$ .

If  $h$  is Fréchet differentiable, we have defined a mapping  $A \rightarrow L_0(A, B)$  by  $a \rightarrow \nabla h(a)(\cdot)$ , where  $L_0(A, B)$  is the set of all bounded linear operators from  $A$  to  $B$ . For  $l \in L_0(A, B)$  we define the norm

$$\| l \|_{L_0(A,B)} = \sup\{ \| l(a) \|_B : a \in A, \| a \|_A \leq 1 \},$$

and we say that  $h$  is  $C_1$ , i.e. differentiable with continuous derivatives, in  $a_0$  if  $a \rightarrow \nabla h(a)(\cdot)$  is continuous in  $a_0$ . Note that  $L_0(A, B)$  is a Banach space under this norm. The definitions and results on the Fréchet derivative can e.g. be found in [6], chapter 2.1.C.

Now, consider the following assumptions for a family of standard parameters  $\{f^\alpha(\cdot), \xi^\alpha, \alpha \in \mathbb{R}\}$ :

**Assumption 6.4.** *For all fixed  $\alpha \in \mathbb{R}$ ,  $f^\alpha(\cdot)$  is differentiable in  $(x, z, k)$  with derivatives denoted by  $\partial_x f^\alpha(t, x, z, k)$ ,  $\partial_z f^\alpha(t, x, z, k)$  and  $\nabla_k f^\alpha(t, x, z, k)$  satisfying the following uniformly boundedness and uniform continuity conditions*

1. *There exists a  $C > 0$  such that for all  $(\alpha, x, z, k)$  and  $h \in \hat{L}_T^2(\mathbb{R}_0, \mathcal{B}(\mathbb{R}_0), \nu)$  we have  $|\partial_x f^\alpha(t, x, z, k)|$ ,  $|\partial_z f^\alpha(t, x, z, k)| \leq C$  and  $|\nabla_k f^\alpha(t, x, z, k)(h)| \leq C \|h\|_{\hat{L}_T^2}$ ,  $dP \times dt$ -a.e.*
2. *For all  $\epsilon > 0$  there exists  $\delta > 0$  such that for all  $(\alpha, x, z, k)$  and  $h \in \mathbb{R}$  and  $h_k \in \hat{L}_T^2(\mathbb{R}_0, \mathcal{B}(\mathbb{R}_0), \nu)$  with  $|h|, \|h_k\| < \delta$  we have*

$$\begin{aligned} |\partial_x f^\alpha(t, x+h, z, k) - \partial_x f^\alpha(t, x, z, k)| &< \epsilon \\ |\partial_x f^\alpha(t, x, z+h, k) - \partial_x f^\alpha(t, x, z, k)| &< \epsilon \\ |\partial_x f^\alpha(t, x, z, k+h_k) - \partial_x f^\alpha(t, x, z, k)| &< \epsilon, \end{aligned}$$

*$dP \times dt$ -a.e., where same inequalities apply for  $\partial_z f$  and for  $\nabla_k f^\alpha$  under the norm  $\|\cdot\|_{L_0(\hat{L}_T^2(\mathbb{R}_0, \mathcal{B}(\mathbb{R}_0), \nu), \mathbb{R})}$ .*

3. *For all  $\epsilon > 0$  there exists  $\delta > 0$  such that for all  $(\alpha, x, z, k)$  and  $h \in \mathbb{R}$  with  $|h| < \delta$  we have for  $dP \times dt$ -a.a.  $(\omega, t)$  that*

$$\begin{aligned} |\partial_x f^{\alpha+h}(t, x, z, k) - \partial_x f^\alpha(t, x, z, k)| &< \epsilon \\ |\partial_z f^{\alpha+h}(t, x, z, k) - \partial_z f^\alpha(t, x, z, k)| &< \epsilon \\ \|\nabla_k f^{\alpha+h}(t, x, z, k) - \nabla_k f^\alpha(t, x, z, k)\| &< \epsilon. \end{aligned}$$

**Assumption 6.5.** *The function  $\alpha \rightarrow (f(\alpha, \cdot), \xi(\alpha))$  is differentiable in the sense that for each  $\alpha_0 \in \mathbb{R}$  the expressions*

$$\frac{1}{\alpha - \alpha_0} [\xi^\alpha - \xi^{\alpha_0}] \tag{20}$$

$$\frac{1}{\alpha - \alpha_0} [f^\alpha(\cdot, X^{\alpha_0}, Z^{\alpha_0}, K^{\alpha_0}) - f^{\alpha_0}(\cdot, X^{\alpha_0}, Z^{\alpha_0}, K^{\alpha_0})] \tag{21}$$

*have limits in  $L_T^2$  and  $H_t^2$  respectively, which we will denote by  $\partial_\alpha \xi^{\alpha_0}$  and  $\partial_\alpha f^{\alpha_0}(t, X_t^{\alpha_0}, Z_t^{\alpha_0}, K_t^{\alpha_0})$  respectively.*



When the assumptions 6.4 and 6.5 applies, we will simplify the notation in the following way. We fix an  $\alpha_0 \in \mathbb{R}$ . Then for all  $\alpha_i \in \mathbb{R}$ ,  $i = 1, 2, 3, 4$  and  $\alpha \in \mathbb{R} \setminus \{\alpha_0\}$  we write

$$\begin{aligned}
\Delta_\alpha X_t &= \frac{X_t^\alpha - X_t^{\alpha_0}}{\alpha - \alpha_0} \\
\Delta_\alpha Z_t &= \frac{Z_t^\alpha - Z_t^{\alpha_0}}{\alpha - \alpha_0} \\
\Delta_\alpha K_t(z) &= \frac{K_t^\alpha(z) - K_t^{\alpha_0}(z)}{\alpha - \alpha_0} \\
f^{\alpha_1}(\alpha_2, \alpha_3, \alpha_4, t) &= f^{\alpha_1}(t, X_t^{\alpha_2}, Z_t^{\alpha_3}, K_t^{\alpha_4}) \\
\Delta_\alpha \xi &= \begin{cases} \frac{\xi^\alpha - \xi^{\alpha_0}}{\alpha - \alpha_0} & \text{if } \alpha \neq \alpha_0 \\ \partial_\alpha \xi^{\alpha_0} & \text{otherwise.} \end{cases} \tag{22}
\end{aligned}$$

**Theorem 6.6.** *Suppose that the assumptions 6.4 and 6.5 hold. Then  $(X^\alpha, Z^\alpha, K^\alpha)$  is differentiable in the sense that for a given  $\alpha_0 \in \mathbb{R}$ ,  $(\Delta_\alpha X, \Delta_\alpha Z, \Delta_\alpha K)$  has a limit in  $\mathcal{V}$  as  $\alpha \rightarrow \alpha_0$ . This limit is denoted by  $(\partial_\alpha X^{\alpha_0}, \partial_\alpha Z^{\alpha_0}, \partial_\alpha K^{\alpha_0})$ , and solves the following BSDE for  $\alpha = \alpha_0$ :*

$$\begin{aligned}
d(\partial_\alpha X_t^\alpha) &= - \left[ \partial_x f^\alpha(t, X_t^\alpha, Z_t^\alpha, K_t^\alpha) \partial_\alpha X_t^\alpha + \partial_y f^\alpha(t, X_t^\alpha, Z_t^\alpha, K_t^\alpha) \partial_\alpha Z_t^\alpha + \right. \\
&\quad \left. \nabla_k f^\alpha(t, X_t^\alpha, Z_t^\alpha, K_t^\alpha) (\partial_\alpha K^\alpha) + \partial_\alpha f^\alpha(t, X_t^\alpha, Z_t^\alpha, K_t^\alpha) \right] dt + \\
&\quad \partial_\alpha Z_t^\alpha dW_t + \int_{\mathbb{R}_0} \partial_\alpha K_t^\alpha(z) \tilde{N}(dt, dz) \\
\partial_\alpha X_T^\alpha &= \partial_\alpha \xi^\alpha. \tag{23}
\end{aligned}$$

*Proof.* We start with the obvious observation that the assumptions 6.4 and 6.5 implies the assumptions 6.1 and 6.2. Also, (23) is well defined because of part 1 and 2 of assumption 6.4 and because of assumption 6.5.

To show the differentiability, we prove this in the point  $\alpha_0 \in \mathbb{R}$ , which will be fixed during the proof.

The idea is to split  $f$  in the same manner as in the proof of theorem 5.1 by defining the processes

$$\begin{aligned}
A_t^\alpha &= \begin{cases} \frac{f^\alpha(\alpha, \alpha, \alpha, t) - f^\alpha(\alpha_0, \alpha, \alpha, t)}{X_t^\alpha - X_t^{\alpha_0}} & \text{if } X_t^\alpha \neq X_t^{\alpha_0} \\ \partial_x f^\alpha(t, X_t^{\alpha_0}, Z_t^\alpha, K_t^\alpha) & \text{otherwise} \end{cases} \\
&= \int_0^1 \partial_x f^\alpha(t, X_t^{\alpha_0} + \lambda(X_t^\alpha - X_t^{\alpha_0}), Z_t^\alpha, K_t^\alpha) d\lambda
\end{aligned}$$

$$\begin{aligned}
B_t^\alpha &= \begin{cases} \frac{f^\alpha(\alpha_0, \alpha, \alpha, t) - f^\alpha(\alpha_0, \alpha_0, \alpha, t)}{Z_t^\alpha - Z_t^{\alpha_0}} & \text{if } Z_t^\alpha \neq Z_t^{\alpha_0} \\ \partial_z f^\alpha(t, X_t^{\alpha_0}, Z_t^{\alpha_0}, K_t^\alpha) & \text{otherwise} \end{cases} \\
&= \int_0^1 \partial_z f^\alpha(t, X_t^{\alpha_0}, Z_t^{\alpha_0} + \lambda(Z_t^\alpha - Z_t^{\alpha_0}), K_t^\alpha) d\lambda \\
D_t^\alpha &= \begin{cases} \frac{f^\alpha(\alpha_0, \alpha_0, \alpha_0, t) - f^{\alpha_0}(\alpha_0, \alpha_0, \alpha_0, t)}{\alpha - \alpha_0} & \text{if } \alpha \neq \alpha_0 \\ \partial_\alpha f^{\alpha_0}(t, X_t^{\alpha_0}, Z_t^{\alpha_0}, K_t^{\alpha_0}) & \text{otherwise.} \end{cases}
\end{aligned}$$

We also define the operators  $C_t^\alpha : \hat{L}_T^2(\mathbb{R}_0, \mathcal{B}(\mathbb{R}_0), \nu) \rightarrow \mathbb{R}$  by

$$C_t^\alpha(h) = \int_0^1 \nabla_k f^\alpha(t, X_t^{\alpha_0}, Z_t^{\alpha_0}, K_t^{\alpha_0} + \lambda(K_t^\alpha - K_t^{\alpha_0}))(h) d\lambda.$$

By assumption 6.4,  $C_t^\alpha$  is a bounded linear operator, and by (2.1.11) in [6] we have

$$\begin{aligned}
\int_0^1 \nabla_k f^\alpha(t, X_t^{\alpha_0}, Z_t^{\alpha_0}, K_t^{\alpha_0} + \lambda(K_t^\alpha - K_t^{\alpha_0}))(K_t^\alpha - K_t^{\alpha_0}) d\lambda = \\
f^\alpha(t, X_t^{\alpha_0}, Z_t^{\alpha_0}, K_t^\alpha) - f^\alpha(t, X_t^{\alpha_0}, Z_t^{\alpha_0}, K_t^{\alpha_0}).
\end{aligned}$$

Because  $f^\alpha$  from assumption 6.4 is assumed to have uniformly bounded derivatives, we together with the Lipschitz condition of  $f$ , trivially have  $A_t^\alpha$ ,  $B_t^\alpha$  and  $C_t^\alpha$  to be uniformly bounded. Of course, the boundedness of  $C_t^\alpha$  is in the sense of bounded linear operators. Also observe that  $D_t^\alpha \in H_T^2$  for all  $\alpha \in \mathbb{R}$ .

Thus, if we define the function  $\chi : \mathbb{R} \times [0, T] \times \mathbb{R} \times \mathbb{R} \times \hat{L}_T^2(\mathbb{R}_0, \mathcal{B}(\mathbb{R}_0), \nu) \rightarrow \mathbb{R}$  given by

$$\chi^\alpha(t, x, z, k) = A_t^\alpha x + B_t^\alpha z + C_t^\alpha(k) + D_t^\alpha$$

we see that  $\chi$  is a generator for all  $\alpha \in \mathbb{R}$  and the solution of the BSDE

$$\begin{aligned}
d(\tilde{X}_t) &= -\chi^\alpha(t, \tilde{X}_t, \tilde{Z}_t, \tilde{K}_t) dt + \tilde{Z}_t dW_t + \int_{\mathbb{R}_0} \tilde{K}_t(z) \tilde{N}(dt, dz) \\
\tilde{X}_t &= \Delta_\alpha \xi
\end{aligned}$$

is  $(\tilde{X}, \tilde{Z}, \tilde{K}) = (\Delta_\alpha X, \Delta_\alpha Z, \Delta_\alpha K)$  for all  $\alpha \neq \alpha_0$ , and  $(\tilde{X}, \tilde{Z}, \tilde{K}) = (\partial_\alpha X^\alpha, \partial_\alpha Z^\alpha, \partial_\alpha K^\alpha)$  for  $\alpha = \alpha_0$ .

The idea now is to prove that the family of standard parameters  $\{\chi^\alpha(\cdot), \Delta_\alpha \xi, \alpha \in \mathbb{R}\}$  satisfies the assumptions 6.1 and 6.2. Then we have from theorem 6.3 that  $(\Delta_\alpha X, \Delta_\alpha Z, \Delta_\alpha K)$  has limit  $(\partial_\alpha X^{\alpha_0}, \partial_\alpha Z^{\alpha_0}, \partial_\alpha K^{\alpha_0})$  in  $\mathcal{V}$  as  $\alpha \rightarrow \alpha_0$  and the theorem is proved.

Because  $A_t^\alpha$ ,  $B_t^\alpha$  and  $C_t^\alpha$  are uniformly bounded, assumption 6.1 is trivially satisfied. So we have to show

$$\|\chi^\alpha(\cdot, \partial_\alpha X^{\alpha_0}, \partial_\alpha Z^{\alpha_0}, \partial_\alpha K^{\alpha_0}) - \chi^{\alpha_0}(\cdot, \partial_\alpha X^{\alpha_0}, \partial_\alpha Z^{\alpha_0}, \partial_\alpha K^{\alpha_0})\|_{H_T^2} \rightarrow 0$$

as  $\alpha \rightarrow \alpha_0$ . To do this, we show that as  $\alpha \rightarrow \alpha_0$ ,

$$\|(A^\alpha - A^{\alpha_0})\partial_\alpha X^{\alpha_0}\|_{H_T^2} \rightarrow 0 \quad (24)$$

$$\|(B^\alpha - B^{\alpha_0})\partial_\alpha Z^{\alpha_0}\|_{H_T^2} \rightarrow 0 \quad (25)$$

$$\|C^\alpha(\partial_\alpha K^{\alpha_0}) - C^{\alpha_0}(\partial_\alpha K^{\alpha_0})\|_{H_T^2} \rightarrow 0 \quad (26)$$

$$\|D^\alpha - D^{\alpha_0}\|_{H_T^2} \rightarrow 0. \quad (27)$$

Notice that (27) is true by assumption 6.5. To show (24), (25) and (26), we split the expressions to vary each variable in  $\alpha$  separately. For (24) this is obtained by showing that the expressions

$$E \left[ \int_0^T [A_t^\alpha - \partial_x f^\alpha(t, X_t^{\alpha_0}, Z_t^\alpha, K_t^\alpha)]^2 |\partial_\alpha X_t^{\alpha_0}|^2 dt \right] \quad (28)$$

$$E \left[ \int_0^T [\partial_x f^\alpha(t, X_t^{\alpha_0}, Z_t^\alpha, K_t^\alpha) - \partial_x f^\alpha(t, X_t^{\alpha_0}, Z_t^{\alpha_0}, K_t^\alpha)]^2 |\partial_\alpha X_t^{\alpha_0}|^2 dt \right] \quad (29)$$

$$E \left[ \int_0^T [\partial_x f^\alpha(t, X_t^{\alpha_0}, Z_t^{\alpha_0}, K_t^\alpha) - \partial_x f^\alpha(t, X_t^{\alpha_0}, Z_t^{\alpha_0}, K_t^{\alpha_0})]^2 |\partial_\alpha X_t^{\alpha_0}|^2 dt \right] \quad (30)$$

$$E \left[ \int_0^T [\partial_x f^\alpha(t, X_t^{\alpha_0}, Z_t^{\alpha_0}, K_t^{\alpha_0}) - \partial_x f^{\alpha_0}(t, X_t^{\alpha_0}, Z_t^{\alpha_0}, K_t^{\alpha_0})]^2 |\partial_\alpha X_t^{\alpha_0}|^2 dt \right] \quad (31)$$

has limit zero as  $\alpha \rightarrow \alpha_0$ . Note that this is satisfied in (31) because of part 3 of assumption 6.4. One important point here is that part 3 of assumption 6.4 is only used to show that (31), and the equivalent expressions for  $\partial_z f^\alpha$  and  $\nabla_k f^\alpha$ , vanishes. See remark 6.7.

To show (28), we see by Jensen's inequality

$$E \left[ \int_0^T [A_t^\alpha - \partial_x f^\alpha(t, X_t^{\alpha_0}, Z_t^\alpha, K_t^\alpha)]^2 |\partial_\alpha X_t^{\alpha_0}|^2 dt \right] \leq E \left[ \int_0^T \int_0^1 [\partial_x f^\alpha(\alpha, t, X_t^{\alpha_0} + \lambda(X_t^\alpha - X_t^{\alpha_0}), Z_t^\alpha, K_t^\alpha) - \partial_x f^\alpha(t, X_t^{\alpha_0}, Z_t^\alpha, K_t^\alpha)]^2 |\partial_\alpha X_t^{\alpha_0}|^2 d\lambda dt \right]. \quad (32)$$

Let  $\epsilon > 0$ . Because  $\partial_x f^\alpha$  is uniformly continuous, there exists a  $\delta > 0$  such that whenever  $|X_t^\alpha - X_t^{\alpha_0}| < \delta$ , we have

$$|\partial_x f^\alpha(\alpha, t, X_t^{\alpha_0} + \lambda(X_t^\alpha - X_t^{\alpha_0}), Z_t^\alpha, K_t^\alpha) - \partial_x f^\alpha(t, X_t^{\alpha_0}, Z_t^\alpha, K_t^\alpha)| < \epsilon$$

for all  $\lambda \in [0, 1]$ . Further, because  $\partial_x f^\alpha$  is uniformly bounded, we can find a  $J > 0$  such that (32) becomes

$$\begin{aligned} & E \left[ \int_0^T [A_t^\alpha - \partial_x f^\alpha(t, X_t^{\alpha_0}, Z_t^\alpha, K_t^\alpha)]^2 |\partial_\alpha X_t^{\alpha_0}|^2 dt \right] \leq \\ & \epsilon^2 \|\partial_\alpha X^{\alpha_0}\|_{H_T^2}^2 + JE \left[ \int_0^T \mathbf{1}\{|X_t^\alpha - X_t^{\alpha_0}| > \delta\} |\partial_\alpha X_t^{\alpha_0}|^2 dt \right]. \end{aligned}$$

We continue by noting that by splitting up the latter integral into  $\{|\partial_\alpha X_t^{\alpha_0}|^2 > M\}$  and its complement for a  $M > 0$ , we get

$$\begin{aligned} & E \left[ \int_0^T \mathbf{1}\{|X_t^\alpha - X_t^{\alpha_0}| > \delta\} |\partial_\alpha X_t^{\alpha_0}|^2 dt \right] \leq \\ & M \int_0^T P(|X_t^\alpha - X_t^{\alpha_0}| > \delta) dt + \\ & E \left[ \int_0^T \mathbf{1}\{|\partial_\alpha X_t^{\alpha_0}|^2 > M\} |\partial_\alpha X_t^{\alpha_0}|^2 dt \right]. \end{aligned}$$

By Markov's inequality we have

$$\begin{aligned} P(|X_t^\alpha - X_t^{\alpha_0}| > \delta) &= P(|X_t^\alpha - X_t^{\alpha_0}|^2 > \delta^2) \leq \\ & \frac{1}{\delta^2} E[|X_t^\alpha - X_t^{\alpha_0}|^2]. \end{aligned}$$

Thus, we are left with

$$\begin{aligned} & E \left[ \int_0^T [A_t^\alpha - \partial_x f^\alpha(t, X_t^{\alpha_0}, Z_t^\alpha, K_t^\alpha)]^2 |\partial_\alpha X_t^{\alpha_0}|^2 dt \right] \leq \\ & \epsilon^2 \|\partial_\alpha X^{\alpha_0}\|_{H_T^2}^2 + \frac{JM}{\delta^2} \|X^\alpha - X^{\alpha_0}\|_{H_T^2}^2 + \\ & J \|\mathbf{1}\{|\partial_\alpha X^{\alpha_0}|^2 > M\} \partial_\alpha X^{\alpha_0}\|_{H_T^2}^2 \end{aligned}$$

for all  $M > 0$  and  $\epsilon > 0$  with the  $\delta = \delta(\epsilon) > 0$  found as above and  $J$  universal.

Now, let  $\rho > 0$ . From the dominated convergence theorem, we have  $\|\mathbf{1}\{|\partial_\alpha X^{\alpha_0}|^2 > M\} \partial_\alpha X^{\alpha_0}\|_{H_T^2}^2 \rightarrow 0$  as  $M \rightarrow \infty$ . Therefore, choose  $M_0 > 0$  such that  $J \|\mathbf{1}\{|\partial_\alpha X^{\alpha_0}|^2 > M\} \partial_\alpha X^{\alpha_0}\|_{H_T^2}^2 < \frac{\rho}{3}$ . Further, let  $\epsilon_0 > 0$  be such that  $\epsilon_0^2 \|\partial_\alpha X^{\alpha_0}\|_{H_T^2}^2 < \frac{\rho}{3}$ , and let  $\delta_0 = \delta(\epsilon_0)$ . Finally, because  $\alpha \rightarrow X^\alpha$  is continuous in the sense defined in theorem 6.3, choose  $\varphi > 0$  such that whenever  $|\alpha - \alpha_0| < \varphi$  we have that  $\frac{JM_0}{\delta_0^2} \|X^\alpha - X^{\alpha_0}\|_{H_T^2}^2 < \frac{\rho}{3}$ . Now we have

$$E \left[ \int_0^T [A_t^\alpha - \partial_x f^\alpha(t, X_t^{\alpha_0}, Z_t^\alpha, K_t^\alpha)]^2 |\partial_\alpha X_t^{\alpha_0}|^2 dt \right] < \rho$$

whenever  $|\alpha - \alpha_0| < \varphi$ , and thus

$$E \left[ \int_0^T [A_t^\alpha - \partial_x f^\alpha(t, X_t^{\alpha_0}, Z_t^\alpha, K_t^\alpha)]^2 |\partial_\alpha X_t^{\alpha_0}|^2 dt \right] \rightarrow 0$$

as  $\alpha \rightarrow \alpha_0$ .

Except for the argument behind equation (32), the proof of (29) is the same as the proof of (28). The same applies for (30), where we split the expression in the sets  $\{\|K_t^\alpha(\cdot) - K_t^{\alpha_0}(\cdot)\|_{\hat{L}} > \delta\}$ ,  $\{|\partial_\alpha X_t^{\alpha_0}|^2 > M\}$  and their compliments remembering that

$$\|K^\alpha - K^{\alpha_0}\|_{\hat{H}}^2 = E \left[ \int_0^T \|K_t^\alpha - K_t^{\alpha_0}\|_{\hat{L}}^2 dt \right]$$

and thus (24) is proved. The proof of (25) is completely similar. To show (26), we note that

$$\begin{aligned} & [\nabla_k f^\alpha(t, X_t^{\alpha_0}, Z_t^{\alpha_0}, K_t^{\alpha_0} + \lambda(K_t^\alpha - K_t^{\alpha_0}))(\partial_\alpha K_t^{\alpha_0}) \\ & - \nabla_k f^\alpha(t, X_t^{\alpha_0}, Z_t^{\alpha_0}, K_t^{\alpha_0})(\partial_\alpha K_t^{\alpha_0})]^2 \leq \\ & \|\partial_\alpha K_t^{\alpha_0}\|_{\hat{L}}^2 \|\nabla_k f^\alpha(t, X_t^{\alpha_0}, Z_t^{\alpha_0}, K_t^{\alpha_0} + \lambda(K_t^\alpha - K_t^{\alpha_0})) \\ & - \nabla_k f^\alpha(t, X_t^{\alpha_0}, Z_t^{\alpha_0}, K_t^{\alpha_0})\|_{L_0(\hat{L}, \mathbb{R})}, \end{aligned}$$

so because  $\hat{L} \rightarrow L_0(\hat{L}, \mathbb{R})$  given by  $k \rightarrow \nabla_k f^\alpha(t, x, z, k)$  is continuous, the proof of (26) is essentially the same as the proof of (24). But then the theorem is proved.  $\square$

**Remark 6.7.** *As mentioned in the proof, we only used part 3 of assumption 6.4 to show that (31), and the equivalent expressions for  $\partial_z f^\alpha$  and  $\nabla_k f^\alpha$ , vanishes. Thus, if one can show that these expressions have limit zero, we do not need this assumption. This will be the case in section 8.*

### 6.3 The differentiable case for quasi-strong generators

One important case when we would like to have theorem 6.6 valid is when the generator is quasi-strong. Therefore, it is interesting to specify for this case the assumptions 6.4 and 6.5. In addition, it will exemplify the theorem. Specifically, the linear case is contained in the quasi-strong generators. For these reasons, let us use some effort to investigate this.

Now, when  $f^\alpha$  is quasi-strong for each  $\alpha$ , we have

$$f^\alpha(t, x, z, k) = g^\alpha(t, x, z, \int_{\mathbb{R}_0} \Psi_t^\alpha(y) k(y) \nu(dy)),$$

where we will call the last component of  $g^\alpha$ , say,  $r$ , i.e.  $(\alpha, \omega, t, x, z, r) \rightarrow g^\alpha(t, x, z, r)$ .

Let us first find  $\nabla_k f^\alpha(t, x, z, k)$  in terms of  $g^\alpha$ . It is the unique bounded operator  $\nabla_k f^\alpha(t, x, z, k)$  such that

$$\begin{aligned} & |g^\alpha(t, x, z, \int_{\mathbb{R}_0} \Psi_t^\alpha(y)(k(y) + h(y))\nu(dy)) - \\ & g^\alpha(t, x, z, \int_{\mathbb{R}_0} \Psi_t^\alpha(y)k(y)\nu(dy)) - \nabla_k f^\alpha(t, x, z, k)(h)| = o(\|h\|), \end{aligned}$$

as  $\|h\| \rightarrow 0$ . Now, suppose for each  $(\alpha, \omega, t, x, z)$ ,  $g^\alpha(t, x, z, \cdot)$  is differentiable. Then we have, where we for notational simplicity write  $g^\alpha(t, x, z, r) = g(r)$ ,

$$\begin{aligned} & |g(\int_{\mathbb{R}_0} \Psi_t^\alpha(y)(k(y) + h(y))\nu(dy)) - g(\int_{\mathbb{R}_0} \Psi_t^\alpha(y)k(y)\nu(dy)) - \\ & \partial_r g(\int_{\mathbb{R}_0} \Psi_t^\alpha(y)k(y)\nu(dy)) \int_{\mathbb{R}_0} \Psi_t^\alpha(y)h(y)\nu(dy)| = o(|\int_{\mathbb{R}_0} \Psi_t^\alpha(y)h(y)\nu(dy)|), \end{aligned}$$

as  $|\int_{\mathbb{R}_0} \Psi_t^\alpha(y)h(y)\nu(dy)| \rightarrow 0$ . Now, suppose  $\Psi^\alpha$  has the bounds from definition 2.3 uniformly in  $\alpha$ . Then, from Hölder's inequality we have for a constant  $D$ ,

$$\begin{aligned} & \int_{\mathbb{R}_0} |\Psi_t^\alpha(y)h(y)|^2 \nu(dy) \leq \\ & \left( \int_{\mathbb{R}_0} |\Psi_t^\alpha(y)|^2 \nu(dy) \right)^{\frac{1}{2}} \left( \int_{\mathbb{R}_0} |h(y)|^2 \nu(dy) \right)^{\frac{1}{2}} \leq D \|h\| \end{aligned}$$

uniformly in  $\alpha$ . But then

$$\begin{aligned} & |g(\int_{\mathbb{R}_0} \Psi_t^\alpha(y)(k(y) + h(y))\nu(dy)) - g(\int_{\mathbb{R}_0} \Psi_t^\alpha(y)k(y)\nu(dy)) - \\ & \partial_r g(\int_{\mathbb{R}_0} \Psi_t^\alpha(y)k(y)\nu(dy)) \int_{\mathbb{R}_0} \Psi_t^\alpha(y)h(y)\nu(dy)| = o(\|h\|), \end{aligned}$$

as  $\|h\| \rightarrow 0$ . Hence,

$$\begin{aligned} \nabla_k f^\alpha(t, x, z, k)(h) = \\ \partial_r g^\alpha(t, x, z, \int_{\mathbb{R}_0} \Psi_t^\alpha(y)k(y)\nu(dy)) \int_{\mathbb{R}_0} \Psi_t^\alpha(y)h(y)\nu(dy). \end{aligned}$$

Now, which other conditions must we assume on  $g^\alpha$  and  $\Psi^\alpha$  for  $\{f^\alpha, \xi^\alpha, \alpha \in \mathbb{R}\}$  to satisfy the assumptions 6.4 and 6.5? It will be commented in theorem 8.5 that the linear case with deterministic coefficients, which trivially has a quasi-strong generator, satisfies these assumptions. For now, let us try to find more general conditions.

One thing that is obvious is that we need the existence of the derivatives of  $g^\alpha$ , i.e.  $\partial_x g^\alpha$ ,  $\partial_z g^\alpha$  and  $\partial_r g^\alpha$ . In addition, they must be uniformly bounded and uniformly continuous in the  $(x, z, r)$ -parameter, equivalent to part 1 and 2 of assumption 6.4. We will also need the boundedness condition on  $\Psi^\alpha$  given over.

The problem is part 3 of assumption 6.4. That is, the problem is when  $\Psi^\alpha$  varies in  $\alpha$ . For instance, to get  $|\partial_x f^{\alpha+h}(t, x, z, k) - \partial_x f^\alpha(t, x, z, k)|$  small by choosing  $h$  small, we would like to resonate in the following way, skipping the  $(t, x, z)$  parameter for notational simplicity:

$$\begin{aligned} & |\partial_x f^{\alpha+h}(k) - \partial_x f^\alpha(k)| = \\ & |\partial_x g^{\alpha+h}(\int_{\mathbb{R}_0} \Psi_t^{\alpha+h}(y)k(y)\nu(dy)) - \partial_x g^\alpha(\int_{\mathbb{R}_0} \Psi_t^\alpha(y)k(y)\nu(dy))| \leq \\ & |\partial_x g^{\alpha+h}(\int_{\mathbb{R}_0} \Psi_t^{\alpha+h}(y)k(y)\nu(dy)) - \partial_x g^\alpha(\int_{\mathbb{R}_0} \Psi_t^{\alpha+h}(y)k(y)\nu(dy))| + \\ & |\partial_x g^\alpha(\int_{\mathbb{R}_0} \Psi_t^{\alpha+h}(y)k(y)\nu(dy)) - \partial_x g^\alpha(\int_{\mathbb{R}_0} \Psi_t^\alpha(y)k(y)\nu(dy))|. \end{aligned}$$

If  $\partial_x g^\alpha$  is uniformly continuous in  $\alpha$ , the first of these two estimates we may get small by choosing  $h$  small. The second, however, would require that we can get  $\|(\Psi_t^{\alpha+h} - \Psi_t^\alpha)k\|$  small by choosing  $h$  small. This is of course impossible to obtain uniformly in  $k$ , unless  $\Psi^\alpha$  does not vary in  $\alpha$  or that the derivatives of  $g^\alpha$  does not vary in the  $r$  parameter, which is essentially the same as  $g^\alpha$  being linear in  $r$ .

If neither of these suggestions apply, we would need to remove the condition that we need uniform continuity, for instance by studying how often  $K_t^\alpha(z)$  is big. This is interesting, but out of the scope of this thesis. However, the two cases discussed may be part of sufficient conditions for the assumptions 6.4 and 6.5 to hold. This leads us to the following two propositions.

**Proposition 6.8.** *Suppose  $\{f^\alpha, \xi^\alpha, \alpha \in \mathbb{R}\}$  is quasi-strong for each  $\alpha \in \mathbb{R}$ , where  $g^\alpha$  has uniformly bounded, uniformly continuous derivatives in  $(\alpha, x, z, r)$ . Suppose further that  $\Psi^\alpha$  is independent of  $\alpha$  and satisfies*

$$D_2(1 \wedge |z|) \leq \Psi_t(z) \leq D_1(1 \wedge |z|)$$

*uniformly for some  $D_1 \geq 0$  and  $D_2 \in (-1, 0]$ . Finally, suppose  $\{f^\alpha, \xi^\alpha, \alpha \in \mathbb{R}\}$  satisfies assumption 6.5. Then  $\{f^\alpha, \xi^\alpha, \alpha \in \mathbb{R}\}$  also satisfies assumption 6.4.*

*Proof.* This follows from our discussion over. □

**Proposition 6.9.** *Suppose  $g^\alpha(\omega, t, x, z, r) = h^\alpha(\omega, t, x, z) + r$ , where  $h$  satisfies 6.4 and 6.5. Further, suppose  $\xi^\alpha$  satisfies (20). Finally, suppose*

$\Psi^\alpha$  is differentiable in  $\alpha$  with uniformly bounded derivative and that there exists  $D_1 \geq 0$  and  $D_2 \in (-1, 0]$  such that the following holds uniformly:

$$D_2(1 \wedge |y|) \leq \Psi_t^\alpha(y), \quad \partial_\alpha \Psi_t^\alpha(y) \leq D_1(1 \wedge |y|).$$

If  $f^\alpha(t, x, z, k) = g^\alpha(t, x, z, \int_{\mathbb{R}_0} \Psi_t^\alpha(y)k(y)\nu(dy)$ , then  $\{f^\alpha, \xi^\alpha, \alpha \in \mathbb{R}\}$  satisfies 6.4 and 6.5.

*Proof.* Now  $f^\alpha$  is the sum of two generators. One is  $h^\alpha$  which satisfies 6.4 and 6.5 by assumption and does not vary in  $k$ . The other is  $\int_{\mathbb{R}_0} \Psi_t^\alpha(y)k(y)\nu(dy)$  which does not vary in  $(x, z)$ . So all we have to do is to show that the latter also satisfies 6.4 and 6.5. Note that  $\nabla_k \left( \int_{\mathbb{R}_0} \Psi_t^\alpha(y)k(y)\nu(dy) \right) (h) = \int_{\mathbb{R}_0} \Psi_t^\alpha(y)h(y)\nu(dy)$ . Thus, part 1 of assumption 6.4 follows from the uniform bounds on  $\Psi^\alpha$ . Because the derivative is constant, part 2 is trivial. Part 3 follows from the mean value inequality, the uniform bound on  $\partial_\alpha \Psi^\alpha$  and the non-dependence on  $r$  in the derivatives of  $g^\alpha$  as we discussed above. To prove that 6.5 is satisfied, observe that

$$\begin{aligned} & E \left[ \int_0^T \left( \int_{\mathbb{R}_0} \left\{ \frac{\Psi_t^\alpha(y) - \Psi_t^{\alpha_0}(y)}{\alpha - \alpha_0} - \partial_\alpha \Psi_t^{\alpha_0}(y) \right\} K_t^{\alpha_0}(y)\nu(dy) \right)^2 dt \right] \leq \\ & E \left[ \int_0^T \left( \int_{\mathbb{R}_0} \left\{ \frac{\Psi_t^\alpha(y) - \Psi_t^{\alpha_0}(y)}{\alpha - \alpha_0} - \partial_\alpha \Psi_t^{\alpha_0}(y) \right\}^2 \nu(dy) \right) \right. \\ & \quad \left. \left( \int_{\mathbb{R}_0} |K_t^{\alpha_0}(y)|^2 \nu(dy) \right) dt \right] \rightarrow 0 \end{aligned}$$

as  $\alpha \rightarrow \alpha_0$ . This follows from the bounded convergence theorem. Thus, the limit of (21) is  $\partial_\alpha h^{\alpha_0}(t, X_t^{\alpha_0}, Z_t^{\alpha_0}) + \int_{\mathbb{R}_0} \partial_\alpha \Psi_t^{\alpha_0}(y)K_t^{\alpha_0}(y)\nu(dy)$ . This completes the proof.  $\square$

## 7 Malliavin derivative

In this section I will prove a result about the Malliavin derivative of the solution of a BSDE. In [11], where this is done in the pure Wiener setting, the derivatives solve a BSDE obtained by formally differentiating each term in the BSDE. I will show a similar result in the jump direction, but because the chain rule is quite different in the jump component, the result will not have this property.

The work on this section was done in the spring semester while learning about the Malliavin calculus. I followed classes and did research on my own. When I was reading literature, it soon became clear that there are many definitions of the derivative. Many specifies  $\Omega$ , for instance the white noise definition and canonical space-type definitions, and many seemed to be a Master project in their own.



I had to choose one that I could work on within the time frame. This I found in [23] which uses a definition that generalize the one for the pure Wiener and pure jump cases in [10]. That is, the definition uses a chaos expansion based on an iterative approach using a representation theorem, which in my case is the martingale representation theorem for Lévy processes. References for this is e.g. [17] or theorem 5.3.6 in [2].

In the main theorem, many results that are classic for the Malliavin derivative are used. So I could have assumed that I had a definition at hand where these results are valid. This is of course not satisfactory. But because the definition in [23] is a mere generalization of the definitions in [10], I have been able to prove all preliminary results needed in the main theorem, the most important being the chain rule.

The preliminary results, I used a significant amount of time proving. However, my supervisor recommended me not to state the proofs as they are very tedious and similar to the classical proofs. Instead I will comment on the idea behind, leaving the proofs as a good exercise.

Now, in [11], the proof of the Malliavin differentiability in the pure Wiener case is done by first proving in section 5.1 a result about extension of filtration in addition to do it in a  $L^p$  setting with  $p > 1$  instead of  $L^2$ . This gave the fundamental inequality

$$\| X \|_{S^p}^p \leq C_p E \left[ \left( |\xi| + \int_0^T |f(s, X_s, Z_s)| ds \right)^p \right]$$

which is strictly stronger than the a priori estimate. Using Hölder's inequality one obtains

$$\left( \int_0^T |f(s, X_s, Z_s)| ds \right)^p \leq T^{\frac{p}{2}} \int_0^T |f(s, X_s, Z_s)|^2 ds.$$

Thus, one may choose  $T$  sufficiently small as a tool. Also, this theorem uses a so-called *orthogonal decomposition with respect to Brownian motion*. Unfortunately, after much research, I was unable to find any references to this, and much less any references to the jump case. I found a result called the *Kunita-Watanabe decomposition*, but this was not enough. Thus, because of short time, I had to abandon this research.

Because I did not have the *choosing  $T$  sufficiently small* tool, I needed other conditions to get the expressions converge to zero. In an article by Delong and Imkeller, see [9], which shows a similar theorem for time delayed generators, and using another definition of the derivative, the condition that  $D_{\theta, z}^1 \xi$  behave sufficiently nice around  $z = 0$  combined with the ideas of [11], the theorem is proved using appropriate inequalities. Note that I will use slightly different inequalities.

One point to notice is that, as in [11] but for another reason,  $T$  is chosen small in the proof in [9]. This is not required in my argument.

The reason for only proving the main theorem in the jump direction the same as is stated in the introduction in [9]. Here, several articles where the pure Wiener case is studied are mentioned. Also, differentiability in the Wiener direction in models with jumps are studied. The authors of [9], as well as myself, has been unable to find the case I am going to prove. This is enough, I think, for my result to be interesting. I will, in addition, give more reasons in a discussion after the main result.

In this section,  $(f, \xi)$  will be a standard parameter and, to be consistent with [23],  $\nu$  a centered Lévy process.

## 7.1 Definition of the Malliavin derivative and preliminary results

Our definition will consist of mixed iterated Poisson and Wiener integrals. For that purpose we define the following unifying notation, with  $\lambda$  as the Lebesgue measure

$$u_k^l = \begin{cases} t_k & l = 0 \\ (t_k, x_k) & l = 1 \end{cases}$$

$$\begin{aligned} U_0 &= [0, T] \text{ and } U_1 = [0, T] \times \mathbb{R}_0 \\ Q_0(dt^0) &= dW(t^0) \text{ and } Q_1(dt^1, dz^1) = \tilde{N}(dt^1, dz^1) \\ d\langle Q_0 \rangle &= d\lambda \text{ and } d\langle Q_1 \rangle = d\lambda \times d\nu \end{aligned}$$

and define an expanded simplex as follows:

$$G_{j_1, \dots, j_k} = \left\{ (u_1^{j_1}, \dots, u_k^{j_k}) \in \prod_{i=1}^k U_{j_i} : 0 < t_1 < \dots < t_k < T \right\},$$

for  $j_1, \dots, j_k = 0, 1$ .

A chaos expansion is constructed the same way as in [10], section 1.2 and 10.2. using a representation theorem. We use the one in theorem 5.3.5 in [2]. A more general is given in [17]. Because this result decomposes the random variable into both a Wiener and a jump integral, we get in each iterative step twice as many terms in addition to a remainder. So instead of a representation on the form

$$F = \sum_{n=0}^{\infty} I_n(f_n)$$

one must expect that for each  $n$ , one has  $2^n$  integrals representing the  $2^n$  ways to combine  $dW_t$  and  $\tilde{N}(dt, dz)$   $n$  times. We also get  $2^n$  remainders,

and not just the one in [10]. However, we can show that these  $2^n$  remainders are no more problematic than the one in [10]. So if we define the following  $n$ -fold integral  $J_n^{(j_1, \dots, j_n)}(g_{j_1, \dots, j_n})$  as

$$J_n^{(j_1, \dots, j_n)}(g_{j_1, \dots, j_n}) = \int_{G_{j_1, \dots, j_n}} g_{j_1, \dots, j_n}(u_1^{j_1}, \dots, u_n^{j_n}) Q_{j_1}(du_1^{j_1}) \dots Q_{j_n}(du_n^{j_n})$$

for  $g_{j_1, \dots, j_n} \in L^2(G_{j_1, \dots, j_n}) = L^2(G_{j_1, \dots, j_n}, \otimes_{i=1}^n d\langle Q_{j_i} \rangle)$ , we obtain the following theorem from [23]:

**Theorem 7.1.** *For every  $F \in L^2(\mathcal{F}_T, P)$ , there exists a unique sequence  $\{g_{j_1, \dots, j_n}\}$ ,  $j_1, \dots, j_n = 0, 1$ , where  $\{g_{j_1, \dots, j_n}\} \in L^2(G_{j_1, \dots, j_n})$ , such that*

$$F = \sum_{n=0}^{\infty} \sum_{j_1, \dots, j_n=0,1} J_n^{(j_1, \dots, j_n)}(g_{j_1, \dots, j_n})$$

and we have the isometry

$$\|F\|_{L^2(P)}^2 = \sum_{n=0}^{\infty} \sum_{j_1, \dots, j_n=0,1} \|g_{j_1, \dots, j_n}\|_{L^2(G_{j_1, \dots, j_n})}^2.$$

The isometry is true because the  $n$ -folded integrals are orthogonal. This follows from the following proposition.

**Proposition 7.2.** *Suppose  $\phi \in H_T^2(\mathbb{R})$ ,  $\psi \in \hat{H}_T^2(\mathbb{R})$ , and define  $X_t = \int_0^t \phi_s dW_s$  and  $Y_t = \int_0^t \int_{\mathbb{R}_0} \psi_s(z) \tilde{N}(ds, dz)$ . Then  $XY$  is a martingale.*

*Proof.* Using theorem 9.5 in [10] (Itô's formula) on  $f(t, x, y) = xy$ , we obtain

$$X_t Y_t = \int_0^t Y_{s-} \phi_s dW_s + \int_0^t \int_{\mathbb{R}_0} X_{s-} \psi_s(z) \tilde{N}(ds, dz)$$

We always choose Wiener and jump integrals to be càdlàg, and thus by [24], theorem 1.3, the following are well defined stopping times:

$$\begin{aligned} \tau_n^1 &= \inf\{t > 0 : |Y_t| > n\} \\ \tau_n^2 &= \inf\{t > 0 : |X_t| > n\}. \end{aligned}$$

Because  $X$  and  $Y$  are square integrable and càdlàg, they are a.s. finite. So  $\tau_n^i \rightarrow \infty$  a.s. for  $i = 1, 2$ . Therefore, for all  $n \geq 1$ ,

$$\begin{aligned} E \left[ \int_0^T (Y_t \phi_t)^2 \mathbf{1}_{\{\tau_n^1 < t\}} dt \right] &< \infty \\ E \left[ \int_0^T \int_{\mathbb{R}_0} (X_t \psi_t(z))^2 \mathbf{1}_{\{\tau_n^2 < t\}} \nu(dz) dt \right] &< \infty. \end{aligned}$$

So  $XY$  is a local martingale. Further, by Hölder's and Doob's inequalities, we have that for all stopping times  $\tau$ , with  $\tau \in [0, T]$  a.s., there exists a  $C > 0$ , independent of  $\tau$ , such that

$$\begin{aligned} E[X_\tau Y_\tau] &\leq E\left[\sup_{0 \leq t \leq T} X_t^2\right]^{\frac{1}{2}} E\left[\sup_{0 \leq t \leq T} Y_t^2\right]^{\frac{1}{2}} \leq \\ &CE[X_T^2]^{\frac{1}{2}} E[Y_T^2]^{\frac{1}{2}} < \infty. \end{aligned}$$

But then  $XY$  is of class DL and hence a martingale. See [15], definition 1.4.8 and exercise 1.5.19 (i). For Doob's inequality, see [15], theorem 3.8 (iv). This theorem requires a right continuous submartingale. This is satisfied by  $|X_t|$  and  $|Y_t|$  by proposition 3.6, also in [15].  $\square$

As mentioned, this shows that  $J_n^{(j_1, \dots, j_n)^1}(g_{(j_1, \dots, j_n)^1})$  and  $J_n^{(j_1, \dots, j_n)^2}(h_{(j_1, \dots, j_n)^2})$  are orthogonal. Here, orthogonality is in the sense that

$$E\left[J_n^{(j_1, \dots, j_n)^1}(g_{(j_1, \dots, j_n)^1})J_n^{(j_1, \dots, j_n)^2}(h_{(j_1, \dots, j_n)^2})\right] = 0,$$

whenever  $(j_1, \dots, j_n)^1 \neq (j_1, \dots, j_n)^2$ . If  $m > n$ ,  $J_n^{j_1, \dots, j_n}(g_{j_1, \dots, j_n})$  and  $J_m^{k_1, \dots, k_m}(h_{k_1, \dots, k_m})$  also are orthogonal, and

$$E\left[J_n^{j_1, \dots, j_n}(g_{j_1, \dots, j_n})J_m^{k_1, \dots, k_m}(h_{k_1, \dots, k_m})\right] = 0.$$

To show this last point, note that if the  $n$  last components in  $(k_1, \dots, k_m)$  are equal to  $(j_1, \dots, j_n)$ , the reasoning in [10], formula (1.7), can be used. That is, we take the expectation of two Wiener integrals or two jump integrals  $n$  times until we have only the expectation of a single of either integrals. If they differ on one of the components, we can apply proposition 7.2 directly because we get the expectation of a jump integral times a Wiener integral.

We are now ready to define the directional derivatives. As in [23], we define the derivatives in both the Wiener direction and jump direction. To do this, we define an expanded simplex

$$\begin{aligned} G_{j_1, \dots, j_n}^k(t) = &\left\{ (u_1^{j_1}, \dots, \hat{u}_k^{j_k}, \dots, u_n^{j_n}) \in G_{j_1, \dots, j_{k-1}, j_{k+1}, \dots, j_n} : \right. \\ &\left. 0 < t_1 < \dots < t_{k-1} < t < t_{k+1} < \dots < t_n < T \right\} \end{aligned}$$

where the  $\hat{u}_k$ -notation means that we omit the  $k$ -th element. Note that  $G_{j_1, \dots, j_n}^k(t) \cap G_{j_1, \dots, j_n}^l(t) = \emptyset$  whenever  $k \neq l$ . Now we define the directional derivatives as follows:

**Definition 7.3.** 1. Let  $\mathbb{D}_{1,2}^{(l)}$  be defined as

$$\mathbb{D}_{1,2}^{(l)} = \left\{ \xi \in L^2(\Omega), \xi = E[\xi] + \sum_{n=1}^{\infty} \sum_{j_1, \dots, j_n=0,1} J_n^{(j_1, \dots, j_n)}(g_{j_1, \dots, j_n}) : \right. \\ \left. \sum_{n=1}^{\infty} \sum_{j_1, \dots, j_n=0,1} \sum_{i=1}^n \mathbf{1}\{j_i = l\} \int_{U_{j_i}} \|g_{j_1, \dots, j_n}(\cdot, u^l, \cdot)\|_{L^2(G_{j_1, \dots, j_n}^i(t))}^2 d\langle Q_l \rangle(u^l) < \infty \right\}.$$

2. For  $\xi \in \mathbb{D}_{1,2}^{(l)}$  we define the derivative in the  $l$ -th direction as:

$$D_{u^l}^{(l)} \xi = \sum_{n=1}^{\infty} \sum_{j_1, \dots, j_n=0,1} \sum_{i=1}^n \mathbf{1}\{j_i = l\} J_{n-1}^{(j_1, \dots, \hat{j}_i, \dots, j_n)} \left( g_{j_1, \dots, j_n}(\dots, u^l, \dots) \mathbf{1}_{G_{j_1, \dots, j_n}^i(t)} \right).$$

It is worth mentioning that this definition reduces to the classical definition in [10] if only the Wiener or the jump part is considered. Now, because we want to differentiate processes, we will also need to define two classes of processes. These are classical spaces when considering Malliavin differentiability of processes.

**Definition 7.4.** Let  $\mathbb{L}_{1,2}^{(l,1)}$ ,  $l \in \{0, 1\}$ , be two spaces of processes  $\phi : \Omega \times [0, T] \rightarrow \mathbb{R}$  and  $\psi : \Omega \times [0, T] \times \mathbb{R}_0 \rightarrow \mathbb{R}$  for  $l = 0$  and  $l = 1$  respectively, such that

$$\begin{aligned} \phi &\in H_T^2, \psi \in \hat{H}_T^2 \\ \phi_s &\in \mathbb{D}_{1,2}^1, \lambda - a.e. \\ \psi_s(y) &\in \mathbb{D}_{1,2}^1, \lambda \times \nu - a.e. \\ E \left[ \int_0^T \int_0^T \int_{\mathbb{R}_0} D_{t,z}^1 \phi_s^2 \nu(dz) dt ds \right] &< \infty \\ E \left[ \int_0^T \int_{\mathbb{R}_0} \int_0^T \int_{\mathbb{R}_0} D_{t,z}^1 \psi_s(y)^2 \nu(dz) dt \nu(dy) ds \right] &< \infty. \end{aligned}$$

Note that these are a banach spaces under the norms

$$\begin{aligned} \|\phi\|_{\mathbb{L}_{1,2}^{(0,1)}}^2 &= E \left[ \int_0^T \phi_s^2 ds \right] + \\ & E \left[ \int_0^T \int_0^T \int_{\mathbb{R}_0} D_{t,z}^1 \phi_s^2 \nu(dz) dt ds \right] \\ \|\psi\|_{\mathbb{L}_{1,2}^{(1,1)}}^2 &= E \left[ \int_0^T \int_{\mathbb{R}_0} \psi_s(y)^2 \nu(dy) ds \right] + \\ & E \left[ \int_0^T \int_{\mathbb{R}_0} \int_0^T \int_{\mathbb{R}_0} D_{t,z}^1 \psi_s(y)^2 \nu(dz) dt \nu(dy) ds \right]. \end{aligned}$$

We may define equivalent spaces for the derivative in the Wiener direction, but this is not used in this thesis.

As mentioned, we will need a chain rule. It is well known from the pure jump or pure Wiener case, see e.g. [10]. Also, [23] proves one for our setting in the Wiener direction. I will state one for the jump direction. The proof is using the same idea as is used in chapter 12.2 in [10] where we find the derivative of the members of a spanning set.

Let us briefly go through the argument. We start by finding a family of Doléans-Dade exponentials that spans  $L^2(\Omega, \mathcal{F}, P)$ . This is achieved in [23] by defining the continuous function

$$\gamma(z) = \begin{cases} e^z - 1 & z \leq 0 \\ 1 - e^{-z} & z > 0. \end{cases}$$

Note that this is a function bounded by 1, and is known to have the following properties, which is proved in [19] page 872:  $\gamma \in L^2(\nu)$ ,  $e^{\alpha\gamma} - 1 \in L^2(\nu)$   $\forall \alpha \in \mathbb{R}$  and for  $h \in C([0, T])$  we have  $e^{\gamma h} - 1 \in L^2(\lambda \times \nu)$ ,  $h\gamma \in L^2(\lambda \times \nu)$  and  $e^{\lambda h} - 1 - \lambda h \in L^1(\lambda \times \nu)$ . Now we have that [23] gives the following lemma:

**Lemma 7.5.** *The linear span  $\mathcal{S}$  of random variables  $Y_T$ , where the  $Y = \{Y_t : t \in [0, T]\}$  are on the form*

$$\begin{aligned} Y_t &= \exp \left\{ \int_0^t \sigma h_s dW_s + \int_0^t \int_{\mathbb{R}_0} h_s \gamma(z) \tilde{N}(ds, dz) \right. \\ & \quad \left. - \int_0^t \frac{\sigma^2 h_s^2}{2} ds - \int_0^t \int_{\mathbb{R}_0} (e^{h_s \gamma(z)} - 1 - h_s \gamma(z)) \nu(dz) ds \right\} \end{aligned}$$

for  $h \in L^2([0, T])$ , is dense in  $L^2(\Omega, \mathcal{F}, P)$ .

To find the chaos expansion of the elements of  $\mathcal{S}$ , we do a simple exercise with Itô's formula and obtain for a  $Y \in \mathcal{S}$

$$dY_t = Y_{t-} \left[ \sigma h_t dW_t + \int_{\mathbb{R}_0} (e^{h_t \gamma(z)} - 1) \tilde{N}(dt, dz) \right]$$

and thus

$$Y_t = 1 + \int_0^t Y_{s-} \left[ \sigma h_s dW_s + \int_{\mathbb{R}_0} (e^{h_s \gamma(z)} - 1) \tilde{N}(ds, dz) \right].$$

If we do this iteratively on the kernels, observing that the remainders vanish, we obtain

$$Y_T = \sum_{n=0}^{\infty} \sum_{j_1, \dots, j_n=0,1} J_n^{(j_1, \dots, j_n)}(g_{j_1, \dots, j_n}),$$

where

$$g_{j_1, \dots, j_n}(u_1, \dots, u_k) = \left( \prod_{\{i: j_i=0\}} \sigma h(t_i) \right) \left( \prod_{\{i: j_i=1\}} (e^{h(t_i) \gamma(z_i)} - 1) \right).$$

Here the product over an empty index set is 1. Further, we can show that  $Y \in \mathbb{D}_{1,2}^0 \cap \mathbb{D}_{1,2}^1$  with derivatives

$$D_{u^l}^{(l)} Y_T = c^l(u^l) Y_T,$$

where  $c^l(u^l) = \sigma h_t$  for  $l = 0$  and  $c^l(u^l) = e^{h_t \gamma(z)} - 1$  for  $l = 1$ . In this thesis, we are interested in  $l = 1$ , and we see we have the same representation for the spanning set  $\mathcal{S}$  as in [10], section 12.2. So using the same argument as in the proof of theorem 12.8 in [10], we obtain the chain rule. Note that there is a misprint on the condition  $\varphi(F + D_{t,z}^1 F) - \varphi(F) \in L^2(P \times \lambda \times \nu)$ . Also, the proof depends on theorem 7.8, but we state the chain rule first because this is most fundamental for the main result of the section.

**Theorem (chain rule) 7.6.** *Let  $k \in \mathbb{N}$  and suppose  $F_i \in \mathbb{D}_{1,2}^1$  for  $i = 1, \dots, k$  and  $\varphi : \mathbb{R}^k \rightarrow \mathbb{R}$  is continuous. Further, suppose  $\varphi(F) \in L^2(P)$  and  $\varphi(F + D_{t,z}^1 F) - \varphi(F) \in L^2(P \times \lambda \times \nu)$ , where  $F = (F_1, \dots, F_k)$ . Then  $\varphi(F) \in \mathbb{D}_{1,2}^1$  and*

$$D_{t,z}^1 \varphi(F) = \varphi(F + D_{t,z}^1 F) - \varphi(F).$$

Let us give some other results used in the main result which are not stated in [23]. We start by a result regarding the representation of a stochastic variable. The proof is a long exercise with the chaos expansions.

**Lemma 7.7.** *Suppose we have that  $\xi \in \mathbb{D}_{1,2}^l$ ,  $l \in \{1, 2\}$ ,  $\phi \in H_T^2(\mathbb{R})$  and  $\psi \in \hat{H}_T^2(\mathbb{R})$ . Further, suppose that*

$$\xi = \int_0^T \phi_s dW_s + \int_0^T \int_{\mathbb{R}_0} \psi_s(z) \tilde{N}(ds, dz).$$

*Then  $\phi \in \mathbb{L}_{1,2}^{0,1}$  and  $\psi \in \mathbb{L}_{1,2}^{1,1}$ .*

One important tool working with Malliavin calculus is the closeability of the derivative. Often we can show that the terms in an approximating sequence is differentiable. The following theorem then gives conditions for the limit to be differentiable. The proof is essentially the same as the proof of theorem 12.6 in [10].

**Theorem 7.8.** *Suppose  $\xi \in L^2(P)$  and  $\xi_k \in \mathbb{D}_{1,2}^1$  for  $k \in \mathbb{N}$  is such that  $\xi_k \rightarrow \xi$  in  $L^2(P)$  and  $D_{\theta,z}\xi_k$  converges in  $L^2(\lambda \times \nu \times P)$ . Then  $\xi \in \mathbb{D}_{1,2}^1$  and  $D_{\theta,z}\xi_k \rightarrow D_{\theta,z}\xi$  in  $L^2(\lambda \times \nu \times P)$ .*

The next lemma says that the Malliavin derivative of an integral with respect to a finite measure, for instance the Lebesgue measure on a bounded interval, is the integral of the Malliavin derivative of the kernel. The result in a different setting is proved in lemma 3.3 in [9], but the proof may be applied here. The proof depends heavily on a fubini-type result for stochastic and Lebesgue integrals. See e.g. problem 3.6.12 in [15] for the Brownian case. For the jump case, see theorem 5 in [1].

**Lemma 7.9.** *Suppose  $\phi \in \mathbb{L}_{1,2}^{0,1}$  and  $\psi \in \mathbb{L}_{1,2}^{1,1}$ . Further, let  $\mu_1$  and  $\mu_2$  be finite measures on  $([0, T], \mathcal{B}([0, T]))$  and  $([0, T] \times \mathbb{R}_0, \mathcal{B}([0, T]) \times \mathcal{B}(\mathbb{R}_0))$  respectively. Then*

$$\begin{aligned} \int_0^T \phi_t \mu_1(dt) &\in \mathbb{D}_{1,2}^1 \\ \int_0^T \int_{\mathbb{R}_0} \psi_t(y) \mu_2(dt, dy) &\in \mathbb{D}_{1,2}^1, \end{aligned}$$

and the following holds for  $\lambda \times \nu$ -a.a.  $(\theta, z)$ :

$$\begin{aligned} D_{\theta,z}^1 \int_0^T \phi_t \mu_1(dt) &= \int_0^T D_{\theta,z}^1 \phi_t \mu_1(dt) \\ D_{\theta,z}^1 \int_0^T \int_{\mathbb{R}_0} \psi_t(y) \mu_2(dt, dy) &= \int_0^T \int_{\mathbb{R}_0} D_{\theta,z}^1 \psi_t(y) \mu_2(dt, dy). \end{aligned}$$

A similar lemma, which is not a corollary of lemma 7.9, but with a proof that uses the same idea is the following.

**Lemma 7.10.** *Suppose  $K \in \mathbb{L}_{1,2}^{1,1}$  and  $\Psi \in \hat{H}_T^2$ , where  $\Psi$  is deterministic and satisfies uniformly*

$$C_2(1 \wedge |y|^2) \leq \Psi_t(y) \leq C_1(1 \wedge |y|^2)$$

for  $C_2 \in (-1, 0]$  and  $C_1 \geq 0$ . Then, for  $\lambda$ -a.a.  $t$

$$\int_{\mathbb{R}_0} \Psi_t(y) K_t(y) \nu(dy) \in \mathbb{D}_{1,2}^1,$$



and  $\lambda \times \nu$ -a.e. we have

$$D_{\theta,z}^1 \int_{\mathbb{R}_0} \Psi_t(y) K_t(y) \nu(dy) = \int_{\mathbb{R}_0} \Psi_t(y) D_{\theta,z}^1 K_t(y) \nu(dy).$$

The last lemma is the equivalent to proposition 3.12 in [10]. The proof is straight forward though tedious.

**Lemma 7.11.** *Suppose  $\xi \in \mathbb{D}_{1,2}^1$ . Then  $E[\xi|\mathcal{F}_t] \in \mathbb{D}_{1,2}^1$  and, for  $\lambda \times \nu$ -a.a.  $(\theta, z)$ ,*

$$D_{\theta,z}^1 E[\xi|\mathcal{F}_t] = E[D_{\theta,z}^1 \xi|\mathcal{F}_t] \mathbf{1}_{[0,t]}(\theta).$$

## 7.2 Differentiability in $\tilde{N}$

Now that we have established the definition and the preliminary results, let us prove the differentiability for a standard parameter  $(f, \xi)$ . To do this, we need the following assumption:

**Assumption 7.12.**  *$(f, \xi)$  is quasi-strong with  $f$  deterministic and  $\xi \in \mathbb{D}_{1,2}^1$ . Further,  $\xi$  satisfies*

$$\lim_{\epsilon \rightarrow 0^+} E\left[\int_0^T \int_{|z|<\epsilon} |D_{\theta,z}^1 \xi|^2 \nu(dz) d\theta\right] = 0, \quad (33)$$

and  $\Psi$  in definition 2.3 satisfies the stronger condition

$$C_2(1 \wedge |z|^2) \leq \Psi_t(z) \leq C_1(1 \wedge |z|^2), \quad \mathcal{P}' - a.e. \quad (34)$$

for  $C_2 \in (-1, 0]$  and  $C_1 \geq 0$ . Further,  $\Psi$  is deterministic.

Some comments are needed on the motivation of this assumption. We need  $f$  deterministic because we are going to apply the chain rule, which we only have for deterministic functions. Because we need to differentiate the arguments of  $f$ , with the theory available in this thesis we need  $f$  to have only real arguments, not operator arguments. So, with  $\Psi$  deterministic, we have

$$f(t, x, z, k) = g(t, x, z, \int_{\mathbb{R}_0} \Psi_t(y) k(y) \nu(dy)),$$

and because  $\Psi$  satisfies the stronger integrability condition (34), we have from theorem 7.10 that if  $K_t(y) \in \mathbb{D}_{1,2}^1$ , then

$$D_{\theta,z}^1 \int_{\mathbb{R}_0} \Psi_t(y) K_t(y) \nu(dy) = \int_{\mathbb{R}_0} \Psi_t(y) D_{\theta,z}^1 K_t(y) \nu(dy).$$

Now, let  $(X, Z, K)$  be the solution of the corresponding BSDE. We use the  $f$ -representation of the generator for notational simplicity. Define

$$\begin{aligned} \hat{f} &: \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R} \times \hat{L}_T^2(\mathbb{R}_0, \mathcal{B}(\mathbb{R}_0), \nu) \rightarrow \mathbb{R} \\ \hat{f}(t, x, z, k) &= f(t, X_t + x, Z_t + z, K_t + k) - f(t, X_t, Z_t, K_t). \end{aligned} \quad (35)$$

It is easy to see that  $\hat{f}$  is Lipschitz with the same Lipschitz constant as  $f$ , and that it is  $\mathcal{P} \times \mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R}) \times \mathcal{B}(\hat{L}_T^2(\mathbb{R}_0, \mathcal{B}(\mathbb{R}_0), \nu))$ -measurable. Thus, for  $\lambda \times \nu$ -a.a.  $(\theta, z)$ ,  $(\hat{f}, D_{\theta, z}\xi)$  is a standard parameter. Let  $(X^{\theta, z}, Z^{\theta, z}, K^{\theta, z})$  be the solution of the corresponding BSDE, i.e.  $(X^{\theta, z}, Z^{\theta, z}, K^{\theta, z})$  solves

$$\begin{aligned} \tilde{X}_t &= D_{\theta, z}\xi + \int_t^T \hat{f}(s, \tilde{X}_s, \tilde{Z}_s, \tilde{K}_s) ds \\ &\quad - \int_t^T \tilde{Z}_s dW_s - \int_t^T \int_{\mathbb{R}_0} \tilde{K}(s, y) \tilde{N}(ds, dy) \end{aligned} \quad (36)$$

for  $\lambda \times \nu$ -a.a.  $(\theta, z)$ . Note that from picard iteration, i.e. theorem 3.5, we can show that this is the a.e. limit of a  $\lambda \times \nu$  measurable sequence and thus is  $(\theta, z)$ -measurable it self. The argumentation is only slightly different to the argumentation behind (38) in the proof. Further, because  $\hat{f}(\cdot, 0, 0, 0) = 0$ , we see from the inequalities of theorem 3.5 that for some  $C > 0$

$$\begin{aligned} &\int_0^T \int_{\mathbb{R}_0} (\|X^{\theta, z}\|_{\beta}^2 + \|Z^{\theta, z}\|_{\beta}^2 + \|K^{\theta, z}\|_{\beta}^2) \nu(dz) d\theta \\ &\leq C \int_0^T \int_{\mathbb{R}_0} \|D_{\theta, z}^1 \xi\| \nu(dz) d\theta < \infty. \end{aligned}$$

**Theorem 7.13.** *Suppose  $(f, \xi)$  satisfies assumption 7.12. Then  $(X, Z, K) \in \mathbb{L}_{1,2}^{0,1} \times \mathbb{L}_{1,2}^{0,1} \times \mathbb{L}_{1,2}^{1,1}$  and  $(D_{\theta, z}^1 X, D_{\theta, z}^1 Z, D_{\theta, z}^1 K)$  is a version of the solution  $(X^{\theta, z}, Z^{\theta, z}, K^{\theta, z})$  of the BSDE (36) for  $\lambda \times \nu$ -a.a.  $(\theta, z)$  with  $\theta \leq t$ .*

*Proof.* As said, the main idea is from [11] and uses Picard iteration and the fact from theorem 7.8 that the malliavin operators are closed operators. In this proof, we let  $(X^{\theta, z}, Z^{\theta, z}, K^{\theta, z})$  be a version of the solution of BSDE (36) for  $t \geq \theta$  for  $\lambda \times \nu$ -a.a.  $(\theta, z)$ , and  $(X^{\theta, z}, Z^{\theta, z}, K^{\theta, z}) = (0, 0, 0)$  elsewhere.

We define the sequence  $(X^k, Z^k, K^k)$  by  $(X^0, Z^0, K^0) = (0, 0, 0)$  and  $(X^{k+1}, Z^{k+1}, K^{k+1})$  the solution of the BSDE

$$\begin{aligned} X_t^{k+1} &= \xi + \int_t^T f(s, X_s^k, Z_s^k, K_s^k) ds \\ &\quad - \int_t^T Z_s^{k+1} dW_s - \int_t^T \int_{\mathbb{R}_0} K_s^{k+1}(y) \tilde{N}(ds, dy). \end{aligned} \quad (37)$$

We know from proposition 3.5, that  $(X^k, Z^k, K^k) \rightarrow (X, Z, K)$  in  $\mathcal{V}$ . Now, suppose  $(X^k, Z^k, K^k) \in \mathbb{L}_{1,2}^{0,1} \times \mathbb{L}_{1,2}^{0,1} \times \mathbb{L}_{1,2}^{1,1}$ . Because of the Lipschitz condition

on  $f$ , it is continuous and satisfies

$$f(t, X_t^k, Z_t^k, K_t^k) \in L^2(P),$$

and

$$\begin{aligned} f(t, X_t^k + D_{\theta,z}^1 X_t^k, Z_t^k + D_{\theta,z}^1 Z_t^k, K_t^k + D_{\theta,z}^1 K_t^k) - \\ f(t, X_t^k, Z_t^k, K_t^k) \in L^2(P \times \lambda \times \nu). \end{aligned}$$

Thus, by theorem (7.6), we have for  $\lambda$ -a.a.  $t$

$$\xi + f(t, X_t^k, Z_t^k, K_t^k) \in \mathbb{D}_{1,2}^1,$$

with derivative

$$\begin{aligned} D_{\theta,z}^1 \xi + f(t, X_t^k + D_{\theta,z}^1 X_t^k, Z_t^k + D_{\theta,z}^1 Z_t^k, \\ K_t^k + D_{\theta,z}^1 K_t^k) - f(t, X_t^k, Z_t^k, K_t^k). \end{aligned}$$

From the construction of the solution in the proof of theorem 3.4, we have

$$X_t^{k+1} = E \left[ \xi + \int_t^T f(s, X_s^k, Z_s^k, Z_s^k) ds \middle| \mathcal{F}_t \right].$$

Thus  $X_t^{k+1} \in \mathbb{D}_{1,2}^1$ ,  $\lambda$ -a.e., and, for  $\theta \leq t$ , we have from lemma 7.11

$$\begin{aligned} D_{\theta,z}^1 X_t^{k+1} = E \left[ D_{\theta,z}^1 \xi + \int_t^T [f(s, X_s^k + D_{\theta,z}^1 X_s^k, Z_s^k + D_{\theta,z}^1 Z_s^k, \\ K_s^k + D_{\theta,z}^1 K_s^k) - f(s, X_s^k, Z_s^k, K_s^k)] ds \middle| \mathcal{F}_t \right]. \end{aligned}$$

Further, from the Lipschitz condition on  $f$ , we easily see

$$E \left[ \int_0^T \int_0^T \int_{\mathbb{R}_0} D_{\theta,z}^1 X_t^{k+1} \nu(dz) d\theta dt \right] < \infty,$$

and so  $X^{k+1} \in \mathbb{L}_{1,2}^{(0,1)}$ , see definition (7.4). Because  $Z^{k+1}$  and  $K^{k+1}$  is constructed as

$$\begin{aligned} E \left[ \xi + \int_0^T f(s, X_s^k, Z_s^k, Z_s^k) ds \middle| \mathcal{F}_t \right] = E \left[ \xi + \int_0^T f(s, X_s^k, Z_s^k, Z_s^k) ds \right] + \\ \int_0^t Z_s^{k+1} dW_s + \int_0^t \int_{\mathbb{R}_0} K_s^{k+1}(y) \tilde{N}(ds, dy), \end{aligned}$$

we have from lemma 7.7 that  $(X^{k+1}, Z^{k+1}, K^{k+1}) \in \mathbb{L}_{1,2}^{(0,1)} \times \mathbb{L}_{1,2}^{(0,1)} \times \mathbb{L}_{1,2}^{(1,1)}$ . Because we obviously have  $(0, 0, 0) \in \mathbb{L}_{1,2}^{(0,1)} \times \mathbb{L}_{1,2}^{(0,1)} \times \mathbb{L}_{1,2}^{(1,1)}$ , we have proved

that  $(X^k, Z^k, K^k) \in \mathbb{L}_{1,2}^{(0,1)} \times \mathbb{L}_{1,2}^{(0,1)} \times \mathbb{L}_{1,2}^{(1,1)}$  for all  $k \in \mathbb{N}_0$ . When this is established, we use (37) and proposition 6 in [23] and obtain

$$\begin{aligned} D_{\theta,z}^1 X_t^{k+1} &= D_{\theta,z}^1 \xi + \\ &\int_t^T \left( f(s, X_s^k + D_{\theta,z}^1 X_s^k, Z_s^k + D_{\theta,z}^1 Z_s^k, K_s^k + D_{\theta,z}^1 K_s^k) - \right. \\ &\quad \left. f(s, X_s^k, Z_s^k, K_s^k) \right) ds + \int_t^T D_{\theta,z}^1 Z_s^{k+1} dW_s + \\ &\int_t^T \int_{\mathbb{R}_0} D_{\theta,z}^1 K_s^{k+1}(y) \tilde{N}(ds, dy), \end{aligned} \quad (38)$$

for  $\theta \leq t$ , and, of course,  $(D_{\theta,z}^1 X_t^{k+1}, D_{\theta,z}^1 Z_t^{k+1}, D_{\theta,z}^1 K_t^{k+1}) = (0, 0, 0)$ ,  $\lambda \times P$ -a.e. for  $\theta > t$ . If we let

$$\begin{aligned} f^k(s, x, z, k) &:= \\ &f(s, X_s^k + D_{\theta,z}^1 X_s^k, Z_s^k + D_{\theta,z}^1 Z_s^k, K_s^k + D_{\theta,z}^1 K_s^k) - \\ &f(s, X_s^k, Z_s^k, K_s^k), \end{aligned}$$

we obviously have that  $(f^k, D_{\theta,z}^1 \xi)$  is a standard parameter for a.a.  $(\theta, z)$ , and  $(D_{\theta,z}^1 X^{k+1}, D_{\theta,z}^1 Z^{k+1}, D_{\theta,z}^1 K^{k+1})$  is the corresponding solution. Note that  $f^k$  does not vary in  $(x, z, k)$ .

Let us show that  $(X^k, Z^k, K^k)$  converges in  $\mathbb{L}_{1,2}^{(0,1)} \times \mathbb{L}_{1,2}^{(0,1)} \times \mathbb{L}_{1,2}^{(1,1)}$ . Because  $(X^k, Z^k, K^k) \rightarrow (X, Z, K)$  in  $\mathcal{V}$ , we need to verify that as  $k \rightarrow \infty$ ,

$$\begin{aligned} &E \left[ \int_0^T \int_0^T \int_{\mathbb{R}_0} |X_t^{\theta,z} - D_{\theta,z}^1 X_t^{k+1}|^2 \nu(dz) d\theta dt \right] + \\ &E \left[ \int_0^T \int_0^T \int_{\mathbb{R}_0} |Z_t^{\theta,z} - D_{\theta,z}^1 Z_t^{k+1}|^2 \nu(dz) d\theta dt \right] + \\ &E \left[ \int_0^T \int_{\mathbb{R}_0} \int_0^T \int_{\mathbb{R}_0} |K_t^{\theta,z}(y) - D_{\theta,z}^1 K_t^{k+1}(y)|^2 \nu(dz) d\theta \nu(dy) dt \right] \rightarrow 0. \end{aligned} \quad (39)$$

Then we have from theorem 7.8 that  $(X, Z, K)$  is Malliavin differentiable in the jump direction, and the derivatives equals (up to a modification) the solution of the BSDE with standard parameters  $(\hat{f}, D_{\theta,z}^1 \xi)$  for  $\lambda \times \nu - a.a.$   $(\theta, z)$  such that  $\theta \leq t$ .

To show (39), we use (33). Whith  $f^1 = f^k$  and  $f^2 = \hat{f}$  from the a priori estimates we split  $\delta_2 f$  in two different ways. One for very small  $|z|$  and one for the other. To be specific, we choose an  $L > 0$  such that the following two

inequalities are uniformly true:

$$\begin{aligned}
|\delta_2 f_t|^2 &= |\hat{f}(t, X_t^{\theta,z}, Z_t^{\theta,z}, K_t^{\theta,z}) - f^k(t, X_t^{\theta,z}, Z_t^{\theta,z}, K_t^{\theta,z})|^2 \leq \\
&2 \left( |\hat{f}(t, X_t^{\theta,z}, Z_t^{\theta,z}, K_t^{\theta,z})|^2 + |f^k(t, X_t^{\theta,z}, Z_t^{\theta,z}, K_t^{\theta,z})|^2 \right) \leq \\
&L \left( |D_{\theta,z}^1 X_t^k|^2 + |D_{\theta,z}^1 Z_t^k|^2 + \|D_{\theta,z}^1 K_t^k\|^2 + \right. \\
&\quad \left. |X_t^{\theta,z}|^2 + |Z_t^{\theta,z}|^2 + \|K_t^{\theta,z}\|^2 \right), \tag{40}
\end{aligned}$$

and for all  $v > 0$ ,

$$\begin{aligned}
|\delta_2 f_t|^2 &= |\hat{f}(t, X_t^{\theta,z}, Z_t^{\theta,z}, K_t^{\theta,z}) - f^k(t, X_t^{\theta,z}, Z_t^{\theta,z}, K_t^{\theta,z})|^2 \leq \\
(1+v) \left( 2 + \frac{1}{v} \right) L &\left( |X_t^k - X_t|^2 + |Z_t^k - Z_t|^2 + \|K_t^k - K_t\|^2 \right) \\
\left( 1 + \frac{1}{v} \right)^2 L &\left( |X_t^{\theta,z} - D_{\theta,z}^1 X_t^k|^2 + \right. \\
&\quad \left. |Z_t^{\theta,z} - D_{\theta,z}^1 Z_t^k|^2 + \|K_t^{\theta,z} - D_{\theta,z}^1 K_t^k\|^2 \right). \tag{41}
\end{aligned}$$

(41) is true because  $2ab \leq va^2 + \frac{1}{v}b^2$  for all  $a, b, v > 0$ , and thus we have

$$\begin{aligned}
(a + (a+b))^2 &\leq (1+v)a^2 + \left(1 + \frac{1}{v}\right)(a+b)^2 \leq \\
&(1+v)a^2 + \left(1 + \frac{1}{v}\right)(1+v)a^2 + \left(1 + \frac{1}{v}\right)^2 b^2 \leq \\
&(1+v) \left( 2 + \frac{1}{v} \right) a^2 + \left( 1 + \frac{1}{v} \right)^2 b^2.
\end{aligned}$$

From now, we will work with the  $\beta$ -norms, remembering the equivalence of the norms. We need the following assumptions on the size of  $\beta$  where we choose  $\lambda^2, \mu^2 > C_f$ , where  $C_f$  is the Lipschitz constant of  $f$ :

$$\beta > \max\{C_f(2 + \lambda^2 + \mu^2), 3C_f^2(T + 2), L(T + 2)\}. \tag{42}$$

Now, let  $\epsilon > 0$ , and let  $\delta > 0$  be such that

$$E \left[ \int_0^T \int_{0 < |z| < \delta} |D_{\theta,z}^1 \xi|^2 \nu(dz) d\theta \right] < \epsilon.$$

By the a priori estimates and (40) we now have

$$\begin{aligned}
& \int_0^T \int_{0 < |z| < \delta} \left( \| X^{\theta,z} - D_{\theta,z}^1 X^{k+1} \|_{\beta}^2 + \right. \\
& \left. \| Z^{\theta,z} - D_{\theta,z}^1 Z^{k+1} \|_{\beta}^2 + \| K^{\theta,z} - D_{\theta,z}^1 K^{k+1} \|_{\beta}^2 \right) \nu(dz) d\theta \leq \\
& L \left[ \int_0^T \int_{0 < |z| < \delta} \left( \| X^{\theta,z} \|_{\beta}^2 + \| Z^{\theta,z} \|_{\beta}^2 + \| K^{\theta,z} \|_{\beta}^2 \right) \nu(dz) d\theta + \right. \\
& \left. \int_0^T \int_{0 < |z| < \delta} \left( \| D_{\theta,z}^1 X^k \|_{\beta}^2 + \| D_{\theta,z}^1 Z^k \|_{\beta}^2 + \| D_{\theta,z}^1 K^k \|_{\beta}^2 \right) \nu(dz) d\theta \right]. \quad (43)
\end{aligned}$$

Let us control these two integrals. To control the first one we use the a priori estimates with  $f^1 = \hat{f}$ ,  $f^2 = 0$ , choosing  $\lambda^2, \mu^2 > C_f$  according to (42). We now have  $f^1(t, 0, 0, 0) = 0$ , and hence  $\delta_2 f = 0$ . Because  $C_f$  also is the Lipschitz constant of  $\hat{f}$ , we have

$$\begin{aligned}
& \int_0^T \int_{0 < |z| < \delta} \left( \| X^{\theta,z} \|_{\beta}^2 + \| Z^{\theta,z} \|_{\beta}^2 + \| K^{\theta,z} \|_{\beta}^2 \right) \nu(dz) d\theta \leq \\
& e^{\beta T} \left( T + \frac{\lambda^2}{\lambda^2 - C_f} + \frac{\mu^2}{\mu^2 - C} \right) E \left[ \int_0^T \int_{0 < |z| < \delta} |D_{\theta,z}^1 \xi|^2 \nu(dz) d\theta \right] < \\
& e^{\beta T} \left( T + \frac{\lambda^2}{\lambda^2 - C_f} + \frac{\mu^2}{\mu^2 - C} \right) \epsilon.
\end{aligned}$$

Of course, this is valid only if  $\beta > C_f(\lambda^2 + \mu^2 + 2)$ , which is verified by (42).

To control the second integral in (43), we see from the a priori estimates with  $f^1 = 0$ ,  $f^2 = f^k$  that we have

$$\begin{aligned}
|\delta_2 f_t|^2 &= |f(t, X_t^{k-1} + D_{\theta,z}^1 X_t^{k-1}, Z_t^{k-1} + D_{\theta,z}^1 Z_t^{k-1}, K_t^{k-1} + D_{\theta,z}^1 K_t^{k-1}) - \\
& f(t, X_t^{k-1}, Z_t^{k-1}, K_t^{k-1})|^2 \leq \\
& 3C_f^2 (|D_{\theta,z}^1 X_t^{k-1}|^2 + |D_{\theta,z}^1 Z_t^{k-1}|^2 + \| D_{\theta,z}^1 K_t^{k-1} \|^2).
\end{aligned}$$

Note that  $f^2$  has Lipschitz constant 0. We choose  $\kappa$  in the a priori estimates to be  $\kappa^2 = \beta$ , and get

$$\begin{aligned}
& \int_0^T \int_{0 < |z| < \delta} \left( \| D_{\theta,z}^1 X^k \|_{\beta}^2 + \| D_{\theta,z}^1 Z^k \|_{\beta}^2 + \| D_{\theta,z}^1 K^k \|_{\beta}^2 \right) \nu(dz) d\theta \leq \\
& (T + 2) e^{\beta T} E \left[ \int_0^T \int_{0 < |z| < \delta} |D_{\theta,z}^1 \xi|^2 \nu(dz) d\theta \right] + \\
& \frac{3(T + 2)C_f^2}{\beta} \int_0^T \int_{0 < |z| < \delta} \left( \| D_{\theta,z}^1 X^{k-1} \|_{\beta}^2 + \right. \\
& \left. \| D_{\theta,z}^1 Z^{k-1} \|_{\beta}^2 + \| D_{\theta,z}^1 K^{k-1} \|_{\beta}^2 \right) \nu(dz) d\theta.
\end{aligned}$$

If we define  $a = \frac{3(T+2)C_f^2}{\beta}$ , then  $a < 1$  because  $\beta > 3C_f^2(T+2)$  by (42). Thus, we recursively see

$$\begin{aligned} & \int_0^T \int_{0 < |z| < \delta} \left( \| D_{\theta,z}^1 X^k \|_\beta^2 + \right. \\ & \quad \left. \| D_{\theta,z}^1 Z^k \|_\beta^2 + \| D_{\theta,z}^1 K^k \|_\beta^2 \right) \nu(dz) d\theta \leq \\ & (T+2)e^{\beta T} (1+a+\dots+a^{k-1}) E \left[ \int_0^T \int_{0 < |z| < \delta} |D_{\theta,z}^1 \xi|^2 \nu(dz) d\theta \right] + \\ & a^{k-1} \int_0^T \int_{0 < |z| < \delta} \left( \| D_{\theta,z}^1 X^0 \|_\beta^2 + \right. \\ & \quad \left. \| D_{\theta,z}^1 Z^0 \|_\beta^2 + \| D_{\theta,z}^1 K^0 \|_\beta^2 \right) \nu(dz) d\theta \leq \frac{(T+2)e^{\beta T}}{1-a} \epsilon \end{aligned}$$

because  $(X^0, Z^0, K^0) = (0, 0, 0)$ . So both the integrals in (40) are bounded by  $\epsilon$  multiplied by constants independent of  $k$  for a given  $\beta$ . So the integrals can be chosen as small as we want by choosing  $\epsilon$  and  $\delta$  sufficiently small.

Now, by (41) and the reasoning with the a priori estimates, we have, if we choose  $\kappa^2 = \beta$ ,

$$\begin{aligned} & \int_0^T \int_{|z| \geq \delta} \left( \| X^{\theta,z} - D_{\theta,z}^1 X^{k+1} \|_\beta^2 + \right. \\ & \quad \left. \| Z^{\theta,z} - D_{\theta,z}^1 Z^{k+1} \|_\beta^2 + \| K^{\theta,z} - D_{\theta,z}^1 K^{k+1} \|_\beta^2 \right) \nu(dz) d\theta \leq \\ & \frac{(T+2)}{\beta} \int_0^T \int_{|z| \geq \delta} \| \delta_2 f \|_\beta^2 \nu(dz) d\theta \leq \\ & \frac{L(T+2)}{\beta} \left[ \left(1 + \frac{1}{v}\right)^2 \int_0^T \int_{|z| \geq \delta} \left( \| X^{\theta,z} - D_{\theta,z}^1 X^k \|_\beta^2 + \right. \right. \\ & \quad \left. \left. \| Z^{\theta,z} - D_{\theta,z}^1 Z^k \|_\beta^2 + \| K^{\theta,z} - D_{\theta,z}^1 K^k \|_\beta^2 \right) \nu(dz) d\theta + \right. \\ & \quad \left. (1+v)\left(2 + \frac{1}{v}\right) \right. \\ & \quad \left. \left( \| X^k - X \|_\beta^2 + \| Z^k - Z \|_\beta^2 + \| K^k - K \|_\beta^2 \right) \int_0^T \int_{|z| \geq \delta} \nu(dz) d\theta \right]. \quad (44) \end{aligned}$$

Because  $\beta > K(T+2)$  we can choose  $v$  so big that  $\frac{K(T+2)}{\beta} \left(1 + \frac{1}{v}\right)^2 < 1$ . Further, choose  $N \in \mathbb{N}$  such that for all  $k \geq N$ , we have

$$\begin{aligned} & (1+v)\left(2 + \frac{1}{v}\right) \left( \| X^k - X \|_\beta^2 + \right. \\ & \quad \left. \| Z^k - Z \|_\beta^2 + \| K^k - K \|_\beta^2 \right) \int_0^T \int_{|z| \geq \delta} \nu(dz) d\theta < \epsilon. \end{aligned}$$

If we define

$$b_2 = \left(1 + \frac{1}{v}\right)^2 \frac{K(T+2)}{\beta} < 1,$$

using (41) recursively, we have for all  $k \geq N$

$$\begin{aligned} & \int_0^T \int_{|z| \geq \delta} \left( \|X^{\theta,z} - D_{\theta,z}^1 X^k\|_{\beta}^2 + \right. \\ & \left. \|Z^{\theta,z} - D_{\theta,z}^1 Z^k\|_{\beta}^2 + \|K^{\theta,z} - D_{\theta,z}^1 K^k\|_{\beta}^2 \right) \nu(dz) d\theta \leq \\ & b_2^{k-N} \int_0^T \int_{|z| \geq \delta} \left( \|X^{\theta,z} - D_{\theta,z}^1 X^N\|_{\beta}^2 + \right. \\ & \left. \|Z^{\theta,z} - D_{\theta,z}^1 Z^N\|_{\beta}^2 + \|K^{\theta,z} - D_{\theta,z}^1 K^N\|_{\beta}^2 \right) \nu(dz) d\theta + \frac{\epsilon}{1-b_2}. \end{aligned}$$

Here,  $b_2$  does not depend on  $\epsilon$  and we can get this expression as close to  $\frac{\epsilon}{1-b_2}$  as we want by choosing  $k$  big.

Let us complete the proof. We choose  $\beta$  according to (42) so that we can get (41), in the equivalent  $\beta$ -norm, as small as we want by first choosing an  $\epsilon > 0$  and a  $\delta > 0$  to control the small  $|z|$ 's and get (43) to be smaller than a constant independent of  $\epsilon$  times  $\epsilon$  it self. Then we can get (44) small by choosing  $k$  big. But then the convergence is established and hence the theorem is proved.  $\square$

This theorem worked because we had the fundamental chain rule at hand. However, it would be much more interesting if the generator was allowed to be stochastic. The equivalent result in [11], for instance, has this at hand, as well as the result in [9]. Why it is interesting is obvious. For instance, when pricing options, the interest rate and volatility will be coefficients in the generator, and these are rarely deterministic.

To study this, one could first look at the linear case with differentiable coefficients. Then the deterministic chain rule may be applied directly when one uses the Picard method in the proof above assuming that  $(X^k, Z^k, K^k)$  is differentiable to show that  $(X^{k+1}, Z^{k+1}, K^{k+1})$  is differentiable. As have been discussed, the linear case is important, and if time would have allowed this, this investigation would have been part of the thesis.

To find a chain rule in the general case, where the randomness is not only in well-handled coefficients, it seems like one has to specify  $\Omega$  to a certain degree. For instance in equation (3.4) in [9], a shift is done on  $\omega$ , which is allowed because  $\Omega$  is a canonical-type space. In the Wiener direction, one extra term, the derivative of  $f$  it self, will be the result. In the jump direction, however, it seems like we get the same differential operator shape on the chain rule. See equation (4.9) in [9]. So the proof over could very well work out for random generators if we had the stochastic chain rule.



Note that most of the preliminary results in section 7.1 are not affected by a random generator. Also, for other definitions of the derivative, these usually maintain true. The only problem both in the linear case and the general case for quasi-strong generators would be to verify lemma 7.10 with  $\Psi$  random. However, this intuitively seems like an easier task.

If I had time to study the case of random generators, in addition to prove differentiability in the Wiener direction, it would be interesting to look at the Clark-Ocone formula, see e.g. chapter 4 in [10]. We could get a powerful alternative representation of the solution of a BSDE.

Let us end this section by the following corollary. It is a classical one when Malliavin differentiability of BSDEs are discussed. It gives an interpretation of the solution components. Because we have only proved differentiability in the jump direction, we get the corollary only for the jump integral term  $K$ .

**Corollary 7.14.** *Under the assumptions of theorem 7.13 the following holds:*

$$D_{\theta,z}^1 X_\theta = K_\theta(z)$$

for  $\lambda \times \nu$ -a.a.  $(\theta, z)$ , a.s.

For a proof, we may use the same idea as in [9], corollary 4.1., which is a little stronger. The same result in the pure Wiener case is proved in [11].

## 8 Combining the derivatives

In this section I will present a result about the double derivative of a family  $\{f^\alpha, \xi^\alpha, \alpha \in \mathbb{R}\}$ , one time in the parameter and one in the Malliavin sense. After working very much with both derivatives, this seemed like a natural result, even though it is not done in [11]. The two types of derivatives are not dependent on each other in any specific way, and the proof should be a simple verification. At least this was my first thought.

The problem is that to be interesting, the solutions should live in

$$L^2(P \times \lambda \times \nu \times \lambda) \times L^2(P \times \lambda \times \nu \times \lambda) \times L^2(P \times \lambda \times \nu \times \lambda \times \nu).$$

But integrability with respect to  $(\theta, z)$  is clearly not a part of theorem 6.6. So for the general case I have therefore not been able to give what I would call natural conditions for this to be well defined. However, in the linear case, it is possible to make sense of this concept.

This is done in the last months before deadline, so I will sketch the idea I have been working on. Then I will, to make the thesis reflect some of my work, prove part two of the list under. Further, I will explain what I think are to unnatural assumptions for part 3 of the list to work. Finally, I will show that this works in the simple linear case. In this case, the problematic terms vanish.

## 8.1 The idea

**Idea list 8.1.** Consider the following algorithm:

1. Suppose  $\{f^\alpha, \xi^\alpha, \alpha \in \mathbb{R}\}$  satisfies the assumptions 6.4 and 6.5. Suppose further that for each  $\alpha \in \mathbb{R}$ ,  $(f^\alpha, \xi^\alpha)$  satisfies the assumption 7.12. Then both derivatives are well defined.
2. Because the Malliavin derivatives solves for each  $\alpha$  the BSDE corresponding to  $(\hat{f}^\alpha, D_{\theta,z}^1 \xi^\alpha)$ , we give  $(\hat{f}^\alpha, D_{\theta,z}^1 \xi^\alpha)$  sufficient conditions for theorem 6.6 to be true,  $\lambda \times \nu$ -a.e. In particular, the BSDE (23) is then well defined for a.a.  $(\theta, z)$ . Remember the definition of  $\hat{f}^\alpha$  in (35). We call the solution of (23) in this case  $(\partial_\alpha D_{\theta,z}^1 X^\alpha, \partial_\alpha D_{\theta,z}^1 Z^\alpha, \partial_\alpha D_{\theta,z}^1 K^\alpha)$ .
3. We show that

$$\int_0^T \int_{\mathbb{R}_0} \| (\partial_\alpha D_{\theta,z}^1 X^\alpha, \partial_\alpha D_{\theta,z}^1 Z^\alpha, \partial_\alpha D_{\theta,z}^1 K^\alpha) \|_{V_\beta}^2 \nu(dz) d\theta < \infty$$

for all  $\alpha \in \mathbb{R}$ , and hence is well defined with the wanted integrability condition.

4. Because for all  $\alpha_0 \in \mathbb{R}$ ,  $(\Delta_\alpha X, \Delta_\alpha Z, \Delta_\alpha K) \rightarrow (\partial_\alpha X^{\alpha_0}, \partial_\alpha Z^{\alpha_0}, \partial_\alpha K^{\alpha_0})$  from theorem 6.6, and  $(\Delta_\alpha X, \Delta_\alpha Z, \Delta_\alpha K) \in \mathbb{L}_{1,2}^{0,1} \times \mathbb{L}_{1,2}^{0,1} \times \mathbb{L}_{1,2}^{1,1}$ , we show

$$\begin{aligned} \int_0^T \int_{\mathbb{R}_0} \| (D_{\theta,z}^1 \Delta_\alpha X - \partial_\alpha D_{\theta,z}^1 X^{\alpha_0}, D_{\theta,z}^1 \Delta_\alpha Z - \partial_\alpha D_{\theta,z}^1 Z^{\alpha_0}, \\ D_{\theta,z}^1 \Delta_\alpha K - \partial_\alpha D_{\theta,z}^1 K^{\alpha_0}) \|_{V_\beta}^2 \nu(dz) d\theta \rightarrow 0. \end{aligned}$$

Then, by theorem 7.8,  $(D_{\theta,z}^1 \partial_\alpha X^{\alpha_0}, D_{\theta,z}^1 \partial_\alpha Z^{\alpha_0}, D_{\theta,z}^1 \partial_\alpha K^{\alpha_0})$  is well defined with the wanted integrability condition and equal to  $(\partial_\alpha D_{\theta,z}^1 X^{\alpha_0}, \partial_\alpha D_{\theta,z}^1 Z^{\alpha_0}, \partial_\alpha D_{\theta,z}^1 K^{\alpha_0})$ . Because

$$\begin{aligned} (D_{\theta,z}^1 \Delta_\alpha X, D_{\theta,z}^1 \Delta_\alpha Z, D_{\theta,z}^1 \Delta_\alpha K) = \\ (\Delta_\alpha D_{\theta,z}^1 X, \Delta_\alpha D_{\theta,z}^1 Z, \Delta_\alpha D_{\theta,z}^1 K), \end{aligned}$$

we have also proved the convergence

$$\begin{aligned} \int_0^T \int_{\mathbb{R}_0} \| (\Delta_\alpha D_{\theta,z}^1 X - \partial_\alpha D_{\theta,z}^1 X^{\alpha_0}, \Delta_\alpha D_{\theta,z}^1 Z - \partial_\alpha D_{\theta,z}^1 Z^{\alpha_0}, \\ \Delta_\alpha D_{\theta,z}^1 K - \partial_\alpha D_{\theta,z}^1 K^{\alpha_0}) \|_{V_\beta}^2 \nu(dz) d\theta \rightarrow 0. \end{aligned}$$

Note that we must do it in this order, that is, make sense of

$$(\partial_\alpha D_{\theta,z}^1 X^\alpha, \partial_\alpha D_{\theta,z}^1 Z^\alpha, \partial_\alpha D_{\theta,z}^1 K^\alpha)$$

before we make sense of

$$(D_{\theta,z}^1 \partial_\alpha X^\alpha, D_{\theta,z}^1 \partial_\alpha Z^\alpha, D_{\theta,z}^1 \partial_\alpha K^\alpha).$$

This is because we only have the Malliavin chain rule for deterministic generators, and  $(\partial_\alpha X^\alpha, \partial_\alpha Z^\alpha, \partial_\alpha K^\alpha)$  certainly does not have a deterministic generator. Thus, we need to use the closeability of the derivative.

Before we proceed, one remark is needed.

**Remark 8.2.** *When we do a limit argument for  $\lambda \times \nu$ -a.a.  $(\theta, z)$  letting  $\alpha \rightarrow \alpha_0$ , this is strictly speaking only valid for a sequence  $\alpha^k \rightarrow \alpha_0$ ,  $k \rightarrow \infty$ . However, all results used holds in this case, and we use for notational simplicity the limit notation  $\alpha \rightarrow \alpha_0$  we have used everywhere else.*

## 8.2 The results on the combined derivative

Let us give sufficient conditions for part 1 and 2 of the idea list to be true.

**Assumption 8.3.**  $\{f^\alpha, \xi^\alpha, \alpha \in \mathbb{R}\}$  satisfies the assumptions 6.4 and 6.5, and for all  $\alpha \in \mathbb{R}$ , the assumption 7.12. Further, for all  $\alpha_0 \in \mathbb{R}$ , for  $\lambda \times \nu$ -a.a.  $(\theta, z)$ , the following has limit in  $H_T^2$ :

$$\begin{aligned} & \frac{1}{\alpha - \alpha_0} [f^\alpha(\cdot, X_t^{\alpha_0} + D_{\theta,z}^1 X_t^{\alpha_0}, Z_t^{\alpha_0} + D_{\theta,z}^1 Z_t^{\alpha_0}, K_t^{\alpha_0} + D_{\theta,z}^1 K_t^{\alpha_0}) - \\ & f^{\alpha_0}(\cdot, X_t^{\alpha_0} + D_{\theta,z}^1 X_t^{\alpha_0}, Z_t^{\alpha_0} + D_{\theta,z}^1 Z_t^{\alpha_0}, K_t^{\alpha_0} + D_{\theta,z}^1 K_t^{\alpha_0})], \end{aligned} \quad (45)$$

as  $\alpha \rightarrow \alpha_0$ . We call this limit

$$\partial_\alpha f^{\alpha_0}(t, X_t^{\alpha_0} + D_{\theta,z}^1 X_t^{\alpha_0}, Z_t^{\alpha_0} + D_{\theta,z}^1 Z_t^{\alpha_0}, K_t^{\alpha_0} + D_{\theta,z}^1 K_t^{\alpha_0}).$$

Further, for all  $\alpha_0 \in \mathbb{R}$  there exists  $\partial_\alpha D_{\theta,z}^1 \xi^{\alpha_0} \in L^2(P \times \nu \times \lambda)$  such that  $\Delta_\alpha D_{\theta,z}^1 \xi \rightarrow \partial_\alpha D_{\theta,z}^1 \xi^{\alpha_0}$  in  $L^2(P \times \nu \times \lambda)$  as  $\alpha \rightarrow \alpha_0$ . Finally,  $\partial_\alpha \xi^{\alpha_0} \in \mathbb{D}_{1,2}^1$  and  $D_{\theta,z}^1 \partial_\alpha \xi^{\alpha_0} = \partial_\alpha D_{\theta,z}^1 \xi^{\alpha_0}$ .

**Proposition 8.4.** *Under assumption 8.3 on  $\{f^\alpha, \xi^\alpha, \alpha \in \mathbb{R}\}$ , then for all  $\alpha_0 \in \mathbb{R}$ , for  $\lambda \times \nu$ -a.a.  $(\theta, z)$ , we have the following:*

1. *The BSDE (23) with  $(f^\alpha, \xi^\alpha)$  substituted by  $(\hat{f}^\alpha, D_{\theta,z}^1 \xi^\alpha)$  for  $\alpha = \alpha_0$  is well defined in  $\mathcal{V}$ . We call the solution  $(\partial_\alpha D_{\theta,z}^1 X^\alpha, \partial_\alpha D_{\theta,z}^1 Z^\alpha, \partial_\alpha D_{\theta,z}^1 K^\alpha)$ .*

2. *We have*

$$\begin{aligned} & (\Delta_\alpha D_{\theta,z}^1 X, \Delta_\alpha D_{\theta,z}^1 Z, \Delta_\alpha D_{\theta,z}^1 K) \rightarrow \\ & (\partial_\alpha D_{\theta,z}^1 X^{\alpha_0}, \partial_\alpha D_{\theta,z}^1 Z^{\alpha_0}, \partial_\alpha D_{\theta,z}^1 K^{\alpha_0}) \end{aligned}$$

in  $\mathcal{V}$  as  $\alpha \rightarrow \alpha_0$ .

*Proof.* To prove this, we need to show that  $(\hat{f}^\alpha, D_{\theta,z}^1 \xi^\alpha)$  satisfies the assumptions 6.4 and 6.5. For simplicity, we let  $\alpha_0 = 0$ .

Now, observe

$$\begin{aligned} \partial_x \hat{f}^\alpha(t, x, z, k) &= \partial_x f^\alpha(t, x + X_t^\alpha, z + Z_t^\alpha, k + K_t^\alpha) - \partial_x f^\alpha(t, X_t^\alpha, Z_t^\alpha, K_t^\alpha) = \\ & \quad \partial_x f^\alpha(t, x + X_t^\alpha, z + Z_t^\alpha, k + K_t^\alpha), \end{aligned}$$

and similarly

$$\begin{aligned} \partial_z \hat{f}^\alpha(t, x, z, k) &= \partial_y f^\alpha(t, x + X_t^\alpha, z + Z_t^\alpha, k + K_t^\alpha) \\ \nabla_k \hat{f}^\alpha(t, x, z, k) &= \nabla_k f^\alpha(t, x + X_t^\alpha, z + Z_t^\alpha, k + K_t^\alpha). \end{aligned}$$

Realizing this, we easily verify part 1 and 2 of assumption 6.4. Let us verify assumption 6.5. Here, the assumption about the end point follows from assumption 8.3. To verify the part about the generator, we must find a limit in  $\mathcal{V}$  for

$$\begin{aligned} & \frac{1}{\alpha} [\hat{f}^\alpha(\cdot, D_{\theta,z}^1 X^0, D_{\theta,z}^1 Z^0, D_{\theta,z}^1 K^0) - \hat{f}^0(\cdot, D_{\theta,z}^1 X^0, D_{\theta,z}^1 Z^0, D_{\theta,z}^1 K^0)] = \\ & \frac{1}{\alpha} [f^\alpha(\cdot, D_{\theta,z}^1 X^0 + X^\alpha, D_{\theta,z}^1 Z^0 + Z^\alpha, D_{\theta,z}^1 K^0 + K^\alpha) - \\ & \quad f^0(\cdot, D_{\theta,z}^1 X^0 + X^0, D_{\theta,z}^1 Z^0 + Z^0, D_{\theta,z}^1 K^0 + K^0)] - \\ & \frac{1}{\alpha} [f^\alpha(\cdot, X^\alpha, Z^\alpha, K^\alpha) - f^0(\cdot, X^0, Z^0, K^0)], \end{aligned} \quad (46)$$

as  $\alpha \rightarrow 0$ . We have already found the limit of the second term in (46) in the proof of theorem 6.6. That is, we proved

$$\begin{aligned} & \frac{1}{\alpha} [f^\alpha(\cdot, X^\alpha, Z^\alpha, K^\alpha) - f^0(\cdot, X^0, Z^0, K^0)] \rightarrow \\ & \quad \partial_x f^0(\cdot, X^0, Z^0, K^0) \partial_\alpha X^0 + \partial_z f^0(\cdot, X^0, Z^0, K^0) \partial_\alpha Z^0 + \\ & \quad \nabla_k f^0(\cdot, X^0, Z^0, K^0) (\partial_\alpha K^0) + \partial_\alpha f^0(\cdot, X^0, Z^0, K^0) \end{aligned}$$

in  $H_T^2$  as  $\alpha \rightarrow 0$ . To find the limit of the first term in (46), we split the expression with the notation  $f(t, \alpha_1, \alpha_2, \alpha_3, \alpha_4) = f^{\alpha_1}(t, D_{\theta,z}^1 X_t^0 + X_t^{\alpha_2}, D_{\theta,z}^1 Z_t^0 + Z_t^{\alpha_3}, D_{\theta,z}^1 K_t^0 + K_t^{\alpha_4})$ :

$$\begin{aligned} & \frac{1}{\alpha} [f^\alpha(t, D_{\theta,z}^1 X_t^0 + X_t^\alpha, D_{\theta,z}^1 Z_t^0 + Z_t^\alpha, D_{\theta,z}^1 K_t^0 + K_t^\alpha) - \\ & \quad f^0(t, D_{\theta,z}^1 X_t^0 + X_t^0, D_{\theta,z}^1 Z_t^0 + Z_t^0, D_{\theta,z}^1 K_t^0 + K_t^0)] = \\ & \frac{1}{\alpha} [f(t, \alpha, \alpha, \alpha, \alpha) - f(t, \alpha, 0, \alpha, \alpha)] + \frac{1}{\alpha} [f(t, \alpha, 0, \alpha, \alpha) - f(t, \alpha, 0, 0, \alpha)] + \\ & \frac{1}{\alpha} [f(t, \alpha, 0, 0, \alpha) - f(t, \alpha, 0, 0, 0)] + \frac{1}{\alpha} [f(t, \alpha, 0, 0, 0) - f(t, 0, 0, 0, 0)]. \end{aligned} \quad (47)$$

Note that by assumption 8.3, the last of the terms has limit  $\partial_\alpha f^0(t, D_{\theta,z}^1 X_t^0 + X_t^0, D_{\theta,z}^1 Z_t^0 + Z_t^0, D_{\theta,z}^1 K_t^0 + K_t^0)$ . To find the limit of the first term, we observe

$$\begin{aligned} & \frac{1}{\alpha} [f^\alpha(t, D_{\theta,z}^1 X_t^0 + X_t^\alpha, D_{\theta,z}^1 Z_t^0 + Z_t^\alpha, D_{\theta,z}^1 K_t^0 + K_t^\alpha) - \\ & \quad f^\alpha(t, D_{\theta,z}^1 X_t^0 + X_t^0, D_{\theta,z}^1 Z_t^0 + Z_t^\alpha, D_{\theta,z}^1 K_t^0 + K_t^\alpha)] = \\ & \frac{\mathbf{1}(X_t^\alpha \neq X_t^0)}{X_t^\alpha - X_t^0} [f^\alpha(t, D_{\theta,z}^1 X_t^0 + X_t^\alpha, D_{\theta,z}^1 Z_t^0 + Z_t^\alpha, D_{\theta,z}^1 K_t^0 + K_t^\alpha) - \\ & \quad f^\alpha(t, D_{\theta,z}^1 X_t^0 + X_t^0, D_{\theta,z}^1 Z_t^0 + Z_t^\alpha, D_{\theta,z}^1 K_t^0 + K_t^\alpha)] \Delta_\alpha X_t = \\ & \quad \left[ \int_0^1 \partial_x f^\alpha(t, D_{\theta,z}^1 X_t^0 + X_t^0 + \lambda(X_t^\alpha - X_t^0), \right. \\ & \quad \left. D_{\theta,z}^1 Z_t^0 + Z_t^\alpha, D_{\theta,z}^1 K_t^0 + K_t^\alpha) d\lambda \right] \Delta_\alpha X_t. \end{aligned}$$

Using the exact same argumentation as is used to show that (32) has zero limit, we can show that this has limit

$$\partial_x f^0(\cdot, D_{\theta,z}^1 X^0 + X^0, D_{\theta,z}^1 Z^0 + Z^0, D_{\theta,z}^1 K^0 + K^0) \partial_\alpha X^0$$

in  $H_T^2$  as  $\alpha \rightarrow 0$ .

Note that because  $\partial_x f^\alpha$  is uniformly bounded and uniformly continuous, the Malliavin derivatives in the arguments and the fact that we multiply by  $\Delta_\alpha X_t$  instead of  $\partial_\alpha X_t^0$  is easily taken care of.

Using the same argument for the second and third term in (47) we may conclude that

$$\begin{aligned} & \frac{1}{\alpha} [\hat{f}^\alpha(\cdot, D_{\theta,z}^1 X^0, D_{\theta,z}^1 Z^0, D_{\theta,z}^1 K^0) - \hat{f}^0(\cdot, D_{\theta,z}^1 X^0, D_{\theta,z}^1 Z^0, D_{\theta,z}^1 K^0)] \rightarrow \\ & [\partial_x f^0(\cdot, D_{\theta,z}^1 X^0 + X^0, D_{\theta,z}^1 Z^0 + Z^0, D_{\theta,z}^1 K^0 + K^0) - \\ & \quad \partial_x f^0(\cdot, X^0, Z^0, K^0)] \partial_\alpha X^0 + \\ & [\partial_z f^0(\cdot, D_{\theta,z}^1 X^0 + X^0, D_{\theta,z}^1 Z^0 + Z^0, D_{\theta,z}^1 K^0 + K^0) - \\ & \quad \partial_z f^0(\cdot, X^0, Z^0, K^0)] \partial_\alpha Z^0 + \\ & [\nabla_k f^0(\cdot, D_{\theta,z}^1 X^0 + X^0, D_{\theta,z}^1 Z^0 + Z^0, D_{\theta,z}^1 K^0 + K^0) - \\ & \quad \nabla_k f^0(\cdot, X^0, Z^0, K^0)] (\partial_\alpha K^0) + \\ & [\partial_\alpha f^0(\cdot, D_{\theta,z}^1 X^0 + X^0, D_{\theta,z}^1 Z^0 + Z^0, D_{\theta,z}^1 K^0 + K^0) - \\ & \quad \partial_\alpha f^0(\cdot, X^0, Z^0, K^0)] \end{aligned} \tag{48}$$

in  $\mathcal{V}$  as  $\alpha \rightarrow 0$ . Thus, we have verified assumption 6.5 and we call this limit

$$\partial_\alpha \hat{f}^0(t, D_{\theta,z}^1 X_t^0, D_{\theta,z}^1 Z_t^0, D_{\theta,z}^1 K_t^0).$$

We have also proved part 1 of the theorem. That is, all the terms in the generator are well defined and satisfies definition 2.1.

Now, the only point left to verify is part 3 of assumption 6.4. Because of remark 6.7, this assumption can be omitted, and instead we verify the condition equivalent to (31). That is, we must show

$$E \left[ \int_0^T [\partial_x \hat{f}^\alpha(t, D_{\theta,z}^1 X_t^0, D_{\theta,z}^1 Z_t^0, D_{\theta,z}^1 K_t^0) - \partial_x \hat{f}^0(t, D_{\theta,z}^1 X_t^0, D_{\theta,z}^1 Z_t^0, D_{\theta,z}^1 K_t^0)]^2 |\partial_\alpha D_{\theta,z}^1 X_t^0|^2 dt \right] \rightarrow 0 \quad (49)$$

and the equivalent expressions for  $\partial_z \hat{f}$  and  $\nabla_k \hat{f}$ . Because these are proven the same way, we only prove (49). To do this, we observe that with  $\partial_x f(t, \alpha_1, \alpha_2, \alpha_3, \alpha_4) = \partial_x f^{\alpha_1}(t, D_{\theta,z}^1 X_t^0 + X_t^{\alpha_2}, D_{\theta,z}^1 Z_t^0 + Z_t^{\alpha_3}, D_{\theta,z}^1 K_t^0 + K_t^{\alpha_4})$ , we have

$$\begin{aligned} & |\partial_x \hat{f}^\alpha(t, D_{\theta,z}^1 X_t^0, D_{\theta,z}^1 Z_t^0, D_{\theta,z}^1 K_t^0) - \partial_x \hat{f}^0(t, D_{\theta,z}^1 X_t^0, D_{\theta,z}^1 Z_t^0, D_{\theta,z}^1 K_t^0)|^2 = \\ & |\partial_x f(\alpha, \alpha, \alpha, \alpha, t) - \partial_x f(0, 0, 0, 0, t)|^2 \leq \\ & 4(|\partial_x f(\alpha, \alpha, \alpha, \alpha, t) - \partial_x f(\alpha, 0, \alpha, \alpha, t)|^2 + \\ & |\partial_x f(\alpha, 0, \alpha, \alpha, t) - \partial_x f(\alpha, 0, 0, \alpha, t)|^2 + \\ & |\partial_x f(\alpha, 0, 0, \alpha, t) - \partial_x f(\alpha, 0, 0, 0, t)|^2 + \\ & |\partial_x f(\alpha, 0, 0, 0, t) - \partial_x f(0, 0, 0, 0, t)|^2). \end{aligned}$$

Because  $\partial_\alpha D_{\theta,z}^1 X^0$  is well defined in the sense of part 1 of the lemma, the three first expressions substituted into (49) gives us the situation in (32) and because  $f^\alpha$  it self satisfies part 3 of assumption 6.4, the last term gives the same situation as in (31). So (49) is proven using known techniques. But then the lemma is proved.  $\square$

The problem now is that  $\partial_\alpha \hat{f}^\alpha(t, D_{\theta,z}^1 X_t^\alpha, D_{\theta,z}^1 Z_t^\alpha, D_{\theta,z}^1 K_t^\alpha)$  is not necessarily in  $L^2(P \times \lambda \times \nu \times \lambda)$ . For instance, we would need the following to be true:

$$E \left[ \int_0^T \int_{\mathbb{R}_0} \int_0^T [\partial_x f^\alpha(t, D_{\theta,z}^1 X_t^\alpha + X_t^\alpha, D_{\theta,z}^1 Z_t^\alpha + Z_t^\alpha, D_{\theta,z}^1 K_t^\alpha + K_t^\alpha) - \partial_x f^\alpha(t, X_t^\alpha, Z_t^\alpha, K_t^\alpha)]^2 |\partial_\alpha X_t^\alpha|^2 dt \nu(dz) d\theta \right] < \infty. \quad (50)$$

Other than assuming this is true, it is very hard to find any assumptions on  $\partial_x f^\alpha$  for this to work. One suggestion could be that  $\partial_x f^\alpha$  is bounded and Lipschitz. Then one may control this by terms on the form

$$E \left[ \int_0^T \int_{\mathbb{R}_0} \int_0^T [(1 \wedge |D_{\theta,z}^1 X_t^\alpha|^2) |\partial_\alpha X_t^\alpha|^2 dt \nu(dz) d\theta] \right].$$

Still it is not trivial to see what could be done to get this finite. So unfortunately, this must be left to later studies. However, the idea list 8.1 works in a very simple linear case. Here the problematic terms vanish.

**Theorem 8.5.** *Suppose the family of standard parameters  $\{f^\alpha, \xi^\alpha, \alpha \in \mathbb{R}\}$  is on the following form:*

$$f^\alpha(t, x, z, k) = \beta_t^\alpha x + \gamma_t^\alpha z + \int_{\mathbb{R}_0} \Psi_t^\alpha(y) k(y) \nu(dy) + \phi_t^\alpha,$$

where  $\beta, \gamma, \Psi$  and  $\phi$  satisfies the conditions in theorem 4.1 uniformly in  $\alpha$ . Suppose further that they are bounded, deterministic and differentiable in  $\alpha$  with bounded, continuous derivatives. Further, there exists constants  $D_1 \geq 0, D_2 \in (-1, 0]$  such that, for all  $(\alpha, t, z)$  we have

$$D_2(1 \wedge |y|^2) \leq \partial_\alpha \Psi_t^\alpha(y)^2 \leq D_1(1 \wedge |y|^2)$$

Then 1-4 in idea list 8.1 are true for  $\{f^\alpha, \xi^\alpha, \alpha \in \mathbb{R}\}$ .

*Proof.* We must check that  $(f^\alpha, \xi^\alpha)$  satisfies assumption 8.3. Assumption 7.12 is trivially satisfied, and by theorem 6.8, assumption 6.4 is satisfied. Also, (20) is true by assumption.

We are left to verify (21) and (45). For notational simplicity, we prove this in  $\alpha_0 = 0$ . Now, consider

$$\begin{aligned} & \frac{1}{\alpha} [f^\alpha(t, D_{\theta,z}^1 X_t^0 + X_t^0, D_{\theta,z}^1 Z_t^0 + Z_t^0, D_{\theta,z}^1 K_t^0 + K_t^0) - \\ & f^0(t, D_{\theta,z}^1 X_t^0 + X_t^0, D_{\theta,z}^1 Z_t^0 + Z_t^0, D_{\theta,z}^1 K_t^0 + K_t^0)] = \\ & \frac{\beta_t^\alpha - \beta_t^0}{\alpha} (D_{\theta,z}^1 X_t^0 + X_t^0) + \frac{\gamma_t^\alpha - \gamma_t^0}{\alpha} (D_{\theta,z}^1 Z_t^0 + Z_t^0) + \\ & \int_{\mathbb{R}_0} \frac{\Psi_t^\alpha(y) - \Psi_t^0(y)}{\alpha} (D_{\theta,z}^1 K_t^0(y) + K_t^0(y)) \nu(dy) + \frac{\phi_t^\alpha - \phi_t^0}{\alpha}. \end{aligned}$$

Because all the coefficients are bounded and deterministic with uniformly bounded derivatives, we see by the mean value inequality, that this is finite for all  $\alpha \in \mathbb{R}$ . Further, by the bounded convergence theorem we see that for  $\lambda \times \nu$ -a.a.  $(\theta, z)$ , it has limit

$$\begin{aligned} & \partial_\alpha \beta_t^0 (D_{\theta,z}^1 X_t^0 + X_t^0) + \partial_\alpha \gamma_t^0 (D_{\theta,z}^1 Z_t^0 + Z_t^0) + \\ & \int_{\mathbb{R}_0} \partial_\alpha \Psi_t^0(y) (D_{\theta,z}^1 K_t^0(y) + K_t^0(y)) \nu(dy) + \partial_\alpha \phi_t^0 \end{aligned}$$

in  $H_T^2$  as  $\alpha \rightarrow 0$ . Similarly

$$\begin{aligned} & \frac{1}{\alpha} [f^\alpha(t, X_t^0, Z_t^0, K_t^0) - f^0(t, X_t^0, Z_t^0, K_t^0)] \rightarrow \\ & \partial_\alpha \beta_t^0 X_t^0 + \partial_\alpha \gamma_t^0 Z_t^0 + \int_{\mathbb{R}_0} \partial_\alpha \Psi_t^0(y) K_t^0(y) \nu(dy) + \partial_\alpha \phi_t^0 \end{aligned}$$

in  $H_T^2$  as  $\alpha \rightarrow 0$ . This again implies that

$$\begin{aligned} \partial_\alpha \hat{f}^0(t, D_{\theta,z}^1 X_t^0, D_{\theta,z}^1 Z_t^0, D_{\theta,z}^1 K_t^0) = \\ \partial_\alpha \beta_t^0 D_{\theta,z}^1 X_t^0 + \partial_\alpha \gamma_t^0 D_{\theta,z}^1 Z_t^0 + \int_{\mathbb{R}_0} \partial_\alpha \Psi_t^0(y) D_{\theta,z}^1 K_t^0(y) \nu(dy), \end{aligned}$$

$\lambda \times \nu$ -a.e. So because  $\hat{f}^\alpha$  obtain the simple from

$$\hat{f}^\alpha(t, x, z, k) = \beta_t^\alpha x + \gamma_t^\alpha z + \int_{\mathbb{R}_0} \Psi_t^\alpha(y) k(y) \nu(dy),$$

we conclude that for a given  $\alpha \in \mathbb{R}$ , for  $\lambda \times \nu$ -a.a.  $(\theta, z)$ ,  $(\partial_\alpha D_{\theta,z}^1 X^\alpha, \partial_\alpha D_{\theta,z}^1 Z^\alpha, \partial_\alpha D_{\theta,z}^1 K^\alpha)$  is the solution of the BSDE with generator

$$\begin{aligned} \partial_\alpha f^{\alpha, \theta, z}(t, x, z_0, k) := \beta_t^\alpha x + \gamma_t^\alpha z_0 + \int_{\mathbb{R}_0} \Psi_t^\alpha(y) k(y) \nu(dy) + \\ \partial_\alpha \beta_t^\alpha D_{\theta,z}^1 X_t^\alpha + \partial_\alpha \gamma_t^\alpha D_{\theta,z}^1 Z_t^\alpha + \int_{\mathbb{R}_0} \partial_\alpha \Psi_t^\alpha(y) D_{\theta,z}^1 K_t^\alpha(y) \nu(dy). \end{aligned}$$

To verify part 3 of the idea list, we see from the assumptions on  $\xi^\alpha$  and the boundedness of the derivatives of the coefficients in  $\alpha$  that this is now trivial. This is because all the terms regarding  $\partial_x \hat{f}^\alpha$ ,  $\partial_z \hat{f}^\alpha$  and  $\nabla_k \hat{f}^\alpha$  are constant, and hence vanish in (48). Further,  $\partial_\alpha \hat{f}^\alpha$  does not depend on  $(X^\alpha, Z^\alpha, K^\alpha)$ . Thus we are left with an easy exercise with the a priori estimates.

To verify step 4 of the idea list, remember that because of the linearity of the Malliavin derivative, we have  $(D_{\theta,z}^1 \Delta_\alpha X, D_{\theta,z}^1 \Delta_\alpha Z, D_{\theta,z}^1 \Delta_\alpha K) = (\Delta_\alpha D_{\theta,z}^1 X, \Delta_\alpha D_{\theta,z}^1 Z, \Delta_\alpha D_{\theta,z}^1 K)$ , and we will work with the latter. We then observe that  $(\Delta_\alpha D_{\theta,z}^1 X, \Delta_\alpha D_{\theta,z}^1 Z, \Delta_\alpha D_{\theta,z}^1 K)$ , defined in (22), solves the BSDE, still with  $\alpha_0 = 0$ :

$$\begin{aligned} \Delta_\alpha D_{\theta,z}^1 X_t = \Delta_\alpha D_{\theta,z}^1 \xi + \\ \int_t^T \left[ \frac{1}{\alpha} (\beta_s^\alpha D_{\theta,z}^1 X_s^\alpha - \beta_s^0 D_{\theta,z}^1 X_s^0) + \frac{1}{\alpha} (\gamma_s^\alpha D_{\theta,z}^1 Z_s^\alpha - \gamma_s^0 D_{\theta,z}^1 Z_s^0) + \right. \\ \left. \frac{1}{\alpha} \int_{\mathbb{R}_0} (\Psi_s^\alpha(y) D_{\theta,z}^1 K_s^\alpha(y) - \Psi_s^0(y) D_{\theta,z}^1 K_s^0(y)) \nu(dy) \right] ds + \\ \int_t^T \Delta_\alpha D_{\theta,z}^1 Z_s dW(s) + \int_t^T \int_{\mathbb{R}_0} \Delta_\alpha D_{\theta,z}^1 K_s(y) \tilde{N}(ds, dy), \end{aligned}$$



$\lambda \times \nu$ -a.e. By splitting the generator, this is easily manipulated to be

$$\begin{aligned} \Delta_\alpha D_{\theta,z}^1 X_t &= \Delta_\alpha D_{\theta,z}^1 \xi + \\ &\int_t^T \left[ \beta_s^\alpha \Delta_\alpha D_{\theta,z}^1 X_s + \gamma_s^\alpha \Delta_\alpha D_{\theta,z}^1 Z_s + \int_{\mathbb{R}_0} \Psi_s^\alpha(y) \Delta_\alpha D_{\theta,z}^1 K_s(y) \nu(dy) + \right. \\ &\left. \frac{\beta_s^\alpha - \beta_s^0}{\alpha} D_{\theta,z}^1 X_s^0 + \frac{\gamma_s^\alpha - \gamma_s^0}{\alpha} D_{\theta,z}^1 Z_s^0 + \int_{\mathbb{R}_0} \frac{\Psi_s^\alpha(y) - \Psi_s^0(y)}{\alpha} D_{\theta,z}^1 K_s^0(y) \nu(dy) \right] ds \\ &+ \int_t^T \Delta_\alpha D_{\theta,z}^1 Z_s dW(s) + \int_t^T \int_{\mathbb{R}_0} \Delta_\alpha D_{\theta,z}^1 K_s(y) \tilde{N}(ds, dy). \end{aligned}$$

Thus,  $(\Delta_\alpha D_{\theta,z}^1 X, \Delta_\alpha D_{\theta,z}^1 Z, \Delta_\alpha D_{\theta,z}^1 K)$  solves the BSDE with generator

$$\begin{aligned} \Delta_\alpha f^{\theta,z}(s, x, z_0, k) &:= \beta_s^\alpha x + \gamma_s^\alpha z_0 + \int_{\mathbb{R}_0} \Psi_s^\alpha(y) k(y) \nu(dy) + \\ &\frac{\beta_s^\alpha - \beta_s^0}{\alpha} D_{\theta,z}^1 X_s^0 + \frac{\gamma_s^\alpha - \gamma_s^0}{\alpha} D_{\theta,z}^1 Z_s^0 + \int_{\mathbb{R}_0} \frac{\Psi_s^\alpha(y) - \Psi_s^0(y)}{\alpha} D_{\theta,z}^1 K_s^0(y) \nu(dy), \end{aligned}$$

$\lambda \times \nu$ -a.e. Finally, to prove that

$$(\Delta_\alpha D_{\theta,z}^1 X, \Delta_\alpha D_{\theta,z}^1 Z, \Delta_\alpha D_{\theta,z}^1 K) \rightarrow (\partial_\alpha D_{\theta,z}^1 X^0, \Delta_\alpha D_{\theta,z}^1 Z^0, \Delta_\alpha D_{\theta,z}^1 K^0)$$

in  $L^2(P \times \lambda \times \nu \times \lambda) \times L^2(P \times \lambda \times \nu \times \lambda) \times L^2(P \times \lambda \times \nu \times \lambda \times \nu)$  as  $\alpha \rightarrow 0$ , we use the a priori estimates with  $f^1 = \Delta_\alpha f^{\theta,z}$  and  $f^2 = \partial_\alpha f^{\alpha,\theta,z}$ . Then we have

$$\begin{aligned} \delta_2 f_s &= (\beta_s^0 - \beta_s^\alpha) \partial_\alpha D_{\theta,z}^1 X_s^0 + (\gamma_s^0 - \gamma_s^\alpha) \partial_\alpha D_{\theta,z}^1 Z_s^0 + \\ &\int_{\mathbb{R}_0} (\Psi_s^0(y) - \Psi_s^\alpha(y)) \partial_\alpha D_{\theta,z}^1 K_s^0(y) \nu(dy) + \\ &(\partial_\alpha \beta_s^0 - \frac{\beta_s^\alpha - \beta_s^0}{\alpha}) D_{\theta,z}^1 X_s^0 + (\partial_\alpha \gamma_s^0 - \frac{\gamma_s^\alpha - \gamma_s^0}{\alpha}) D_{\theta,z}^1 Z_s^0 + \\ &\int_{\mathbb{R}_0} (\partial_\alpha \Psi_s^0(y) - \frac{\Psi_s^\alpha(y) - \Psi_s^0(y)}{\alpha}) D_{\theta,z}^1 K_s^0(y) \nu(dy). \end{aligned}$$

Thus, for a constant  $L$  independent of  $\alpha$ , we obtain

$$\begin{aligned}
& \int_0^T \int_{\mathbb{R}} [ \| \partial_\alpha D_{\theta,z}^1 X^0 - \Delta_\alpha D_{\theta,z}^1 X \|^2 + \| \partial_\alpha D_{\theta,z}^1 Z^0 - \Delta_\alpha D_{\theta,z}^1 Z \|^2 + \\
& \quad \| \partial_\alpha D_{\theta,z}^1 K^0 - \Delta_\alpha D_{\theta,z}^1 K \|^2 ] \nu(dz) d\theta \leq \\
L & \int_0^T \int_{\mathbb{R}_0} [ \| \partial_\alpha D_{\theta,z}^1 \xi^0 - \Delta_\alpha D_{\theta,z}^1 \xi \|^2_{L^2_T} ] \nu(dz) d\theta + \\
L & \int_0^T \int_{\mathbb{R}_0} [ \| (\partial_\alpha \beta^0 - \frac{\beta^\alpha - \beta^0}{\alpha}) D_{\theta,z}^1 X^0 \|^2_{H^2_T} + \| (\beta^0 - \beta^\alpha) \partial_\alpha D_{\theta,z}^1 X^0 \|^2_{H^2_T} + \\
& \quad \| (\partial_\alpha \gamma^0 - \frac{\gamma^\alpha - \gamma^0}{\alpha}) D_{\theta,z}^1 Z^0 \|^2_{H^2_T} + \| (\gamma^0 - \gamma^\alpha) \partial_\alpha D_{\theta,z}^1 Z^0 \|^2_{H^2_T} + \\
& \quad \| (\partial_\alpha \Psi^0 - \frac{\Psi^\alpha - \Psi^0}{\alpha}) D_{\theta,z}^1 K^0 \|^2_{\dot{H}^2_T} + \\
& \quad \| (\Psi^0 - \Psi^\alpha) \partial_\alpha D_{\theta,z}^1 K \|^2_{\dot{H}^2_T} ] \nu(dz) d\theta.
\end{aligned}$$

Because  $(D_{\theta,z}^1 X^0, D_{\theta,z}^1 Z^0, D_{\theta,z}^1 K^0)$  and  $(\partial_\alpha D_{\theta,z}^1 X^0, \partial_\alpha D_{\theta,z}^1 Z^0, \partial_\alpha D_{\theta,z}^1 K^0)$  both have finite integrals and because  $\beta, \gamma$  and  $\Psi$  are deterministic, bounded and differentiable in  $\alpha$  with bounded, continuous derivatives, we get from the bounded convergence theorem that this has limit zero. Hence the theorem is proved.  $\square$

As was mentioned in the discussion after theorem 7.13, it would be of high interest to study the case where  $\beta^\alpha, \gamma^\alpha$  and  $\Psi^\alpha$  are stochastic. We would not have the problematic terms like (50). However, the chain rule would give us terms on the form

$$D_{\theta,z}^1 (X_t^\alpha \beta_t^\alpha) = \beta_t^\alpha D_{\theta,z}^1 X_t^\alpha + X_t^\alpha D_{\theta,z}^1 \beta_t^\alpha + D_{\theta,z}^1 \beta_t^\alpha D_{\theta,z}^1 X_t^\alpha.$$

Proving this to be contained in  $L^2(P \times \lambda \times \nu \times \lambda)$  would lead anyway to the same problem as in (50). So there are much left to study in this section.

By now, I do not directly see any applications for this, but I see potential for applications. For instance, when calculating the greeks,  $\Delta$  may be calculated as

$$\Delta = \frac{\partial}{\partial \alpha} E[\varphi(X_T^\alpha)],$$

with  $\alpha$  representing the initial price. See chapter 4.4. in [10]. Controlling the differential of  $X^\alpha$  in the parameter, could be a tool. Also, when controlling  $D_{\theta,z}^1 \partial_\alpha X^\alpha$ , one could use the Clark-Ocone formula on the  $\alpha$ -derivatives as a tool. This is of course speculation, and certainly much more study is needed.

## References

- [1] D. Applebaum. Martingale-valued measures, Ornstein-Uhlenbeck processes with jumps and operator self-decomposability in Hilbert space. *Lecture Notes in Math.*, 1874:171–196, 2006.
- [2] D. Applebaum. *Lévy processes and stochastic calculus*. Cambridge University Press, second edition, 2009.
- [3] G. Barles, R. Buckdahn, and E. Pardoux. Backward stochastic differential equations and integral-partial differential equations. *Stochastics An International Journal of Probability and Stochastic Processes*, 60(1):57–83, 1997.
- [4] R. G. Bartle. *The elements of integration and Lebesgue measure*. Wiley Classics Library, 1995.
- [5] C. Bender and M. Kohlmann. Optimal superhedging under nonconvex constraints. *Preprint*, 2004.
- [6] M. S. Berger. *Nonlinearity and functional analysis*, volume 74 of *Pure and Applied Mathematics*. Academic Press, 1977.
- [7] R. Buckdahn and Y. Hu. Pricing of american contingent claims with jump stock price and constrained portfolios. *Mathematics of Operations Research*, 23(1):177–203, 1998.
- [8] R. Cont and P. Tankov. *Financial modelling with jump processes*. Chapman & Hall/CRC, 2004.
- [9] Ł. Delong and P. Imkeller. On malliavin’s differentiability of BSDE with time delayed generators driven by Brownian motions and Poisson random measures. *Preprint*, 2009.
- [10] G. Di Nunno, B. Øksendal, and F. Proske. *Malliavin calculus for Lévy processes with applications to finance*. Springer-Verlag, 2009.
- [11] N. El Karoui, S. Peng, and M. C. Quenez. Backward stochastic differential equations in finance. *Math. Finance*, 7(1):1–71, 1997.
- [12] R. Elie and I. Kharroubi. Constrained backward SDEs with jumps: Application to optimal switching. *Preprint*, 2009.
- [13] Y. Hu, P. Imkeller, and M. Müller. Utility maximization in incomplete markets. *The Annals of Applied Probability*, 15(3):1691–1712, 2005.
- [14] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes*. North-Holland, second edition, 1989.

- [15] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*. Springer-Verlag, second edition, 1998.
- [16] B. Øksendal and A. Sulem. *Applied stochastic control of jump diffusions*. Springer-Verlag, third edition, 2009.
- [17] H. Kunita. Representation of martingales with jumps and applications to mathematical finance. *Advanced Studies in Pure Mathematics*, 41:209–232, 2004.
- [18] Q. Q. Lin. Nonlinear Doob-Meyer decomposition with jumps. *Acta Mathematica Sinica, English Series*, 19(1):69–78, 2003.
- [19] A. Løkka. Martingale representation of functionals of Lévy processes. *Stochastic Anal. Appl.*, 22(4):867–892, 2004.
- [20] M.-A. Moralis. An extended existence result for quadratic BSDEs with jumps with application to the utility maximization problem. *Preprint*, 2008.
- [21] M.-A. Moralis. Utility maximization in a jump market model. *Preprint*, 2008.
- [22] S. Peng. Monotonic limit theorem of BSDE and nonlinear decomposition theorem of Doob-Meyer’s type. *Probab. Theory Relat. Fields*, 113:473–499, 1999.
- [23] E. Petrou. Malliavin calculus in Lévy spaces and applications to finance. *Electronic Journal of Probability*, 13:852–879, 2008.
- [24] P. E. Protter. *Stochastic integration and differential equations*. Springer-Verlag, second edition, 2004. Version 2.1.
- [25] M. Royer. Backward stochastic differential equations with jumps and related non-linear expectations. *Stochastic processes and their applications*, 116:1358–1376, 2006.
- [26] S. Tang and X. Li. Necessary conditions for optimal control of stochastic systems with random jumps. *SIAM J. Control Optim.*, 32(5):1447–1475, 1994.