# Malliavin Calculus for Lévy Processes with Applications to Finance

by

Martin Petter Johansson

A thesis presented for the degree of
Doctor of Philosophy of the University of London
and the
Diploma of Imperial College

October 2004

Department of Mathematics
Imperial College London
180 Queen's Gate
London SW7 2BZ

# ACKNOWLEDGEMENTS

First and foremost I would like to thank my supervisor Prof Mark Davis, who has been a role model not only scientifically but also, more generally, in life. His scientific knowledge is well documented, but it is his ability to inspire and encourage that makes him the ideal supervisor.

I have also benefited greatly from weekly reading seminars on stochastic calculus and mathematical finance. I would like to thank all participants and in particular Dr Dan Crisan for organising and leading these seminars.

If Imperial College London supplied scientific support, Citigroup has been the institution supplying financial support, without which this project would have been impossible. I would like to thank Dr Eduardo Epperlein for making this arrangement possible, for his guidance and genuine concern.

Finally I must thank my family—Emma, Thomas, Gunilla, Anna and Lina—for support, motivation and meaning.

# ABSTRACT

Extending Gaussian Malliavin derivatives, to a finite moments Lévy process framework, using chaos expansions has proven to be a successful approach. In this work the theory is extended by the introduction of the Skorohod integral, and its properties are investigated. From this rather general case the scope is narrowed down–step by step–to a class of jump diffusion processes for which we have a Malliavin derivative chain rule, which is important for the application of Malliavin calculus to variance reduction of Monte Carlo simulation of contingent claim sensitivities. Stochastic weights for these simulations are derived and examples with numerical experiments are presented. These stochastic weights are well known in the continuous case, but the introduction of discontinuous jumps allows for extensions to other asset classes such as credit derivatives. Monte Carlo methods are widely used in particular for credit derivatives, and therefore alternative methods for faster convergence of sensitivities have previously been developed. One such technique is the likelihood ratio method, which is closely linked with the Malliavin weighted method. This discussion is also formalised.

# CONTENTS

**1. Introduction** . . . . . 8

1.1 Background . . . . . 8

1.2 Outline of thesis . . . . . 10

**Part I Malliavin calculus for Lévy processes** 12

**2. Malliavin calculus for Lévy processes satisfying a moment condition** 13

2.1 The chaotic representation property . . . . . 13

2.1.1 Setup and strongly orthogonal martingales . . . . . 13

2.1.2 The chaotic and the predictable representation property . . 15

2.2 Malliavin calculus . . . . . 16

2.2.1 Malliavin derivative . . . . . 16

2.2.2 Skorohod integral . . . . . 18

**3. Malliavin Calculus for simple Lévy processes** . . . . . 24

3.1 Setup and representation theorems . . . . . 24

3.2 Malliavin calculus . . . . . 26

3.2.1 Malliavin derivative . . . . . 26

3.2.2 Skorohod integral . . . . . 28

3.3 Separable jump diffusions . . . . . 34

**Part II Stochastic weights for derivative sensitivities** 38

**4. Malliavin Monte Carlo Greeks for separable jump diffusions** . . . . . 39

4.1 Sensitivity weights . . . . . 39

4.1.1 Variations in the drift coefficient . . . . . 41
4.1.2 Variations in the initial condition . . . . . 41
4.1.3 Variations in the diffusion coefficient . . . . . 43
4.1.4 Variations in the jump intensity . . . . . 45
4.1.5 Variations in the jump amplitude . . . . . 47
4.2 Examples . . . . . 48
4.3 Numerical experiments . . . . . 55

**5. Stochastic weights for credit derivative Greeks** . . . . . 59
5.1 Intensity sensitivity for single name claims . . . . . 59
5.1.1 Constant intensity . . . . . 59
5.1.2 Cox processes . . . . . 60
5.1.3 Example . . . . . 62
5.2 Correlation sensitivity for baskets of credits . . . . . 64
5.2.1 Correlated default events . . . . . 64
5.2.2 Correlation sensitivity . . . . . 65
5.2.3 Examples and extensions . . . . . 68
5.3 Alternative methods . . . . . 70

**6. The likelihood ratio and Malliavin methods** . . . . . 73
6.1 Preliminaries . . . . . 73
6.1.1 Likelihood ratio weights . . . . . 73
6.1.2 General Malliavin weight expressions . . . . . 74
6.2 Differentiability of a density and the Likelihood ratio method . . . 75

**References** . . . . . 79

# LIST OF FIGURES


4.1 Monte Carlo simulation of $\text{Delta}_{\text{BS}}$ for a call option using finite difference approximation and Malliavin weighting. The model parameters are $x = 100$, Strike $= 100$, $r = 0.05$, $\sigma = 0.3$, $\alpha = 0.5$ and $\lambda = 0.1$. . . . . 56

4.2 Monte Carlo simulation of $\text{Delta}_{\text{BS}}$ for a digital option using finite difference approximation and Malliavin weighting. The model parameters are $x = 100$, Strike $= 100$, $r = 0.05$, $\sigma = 0.3$, $\alpha = 0.5$ and $\lambda = 0.1$. . . . . 57

4.3 Monte Carlo simulation of $\text{Delta}_{\text{H}}$ for a call option using finite difference approximation and Malliavin weighting. The model parameters are $x_1 = 100$, Strike $= 100$, $x_2 = 0.04$, $r = 0.05$, $\kappa = 1$, $\theta = 0.04$, $\sigma = 0.04$, $\rho = -0.8$, $\alpha = 0.5$ and $\lambda = 0.1$. . . . . 57

4.4 Monte Carlo simulation of $\text{Delta}_{\text{H}}$ for a digital option using finite difference approximation and Malliavin weighting. The model parameters are $x_1 = 100$, Strike $= 100$, $x_2 = 0.04$, $r = 0.05$, $\kappa = 1$, $\theta = 0.04$, $\sigma = 0.04$, $\rho = -0.8$, $\alpha = 0.5$ and $\lambda = 0.1$. . . . . 58

5.1 CIR fit to Libor curve, Bombardier credit spread curve and parallel shifted Bombardier credit spread curve as of 23 April 2004. Sources: Mark-It partners (BOMB SNRFOR USD MR) and Citigroup. . . . . 63

5.2 Monte Carlo simulation of the sensitivity to parallel shifts in credit spread curve for a 5 year digital credit default swap on Bombardier as of 23 April 2004. . . . . 64

5.3 Monte Carlo simulation of the correlation sensitivity for a one year digital second to default credit swap on a basket of two credits. Both jump intensities were 0.3 and the correlation was 0.3. . . . . 69

5.4 Monte Carlo simulation of the correlation sensitivity for a one year digital second to default credit swap on a basket of two credits. A $t$–copula with 2 degrees of freedom was used, both jump intensities were 0.3 and the correlation was 0.3. . . . . 70

5.5 Monte Carlo simulation of the intensity sensitivity for a one year digital credit default swap. The constant jump intensity was 0.3. . . 72

# 1. INTRODUCTION

## 1.1 Background

Malliavin calculus has recently come into fashion in quantitative finance. The reason is its application to variance reduction techniques when the sensitivity of a derivative contract is evaluated by Monte Carlo simulation. Using simulation techniques, one obvious method to obtain the sensitivities is to evaluate the derivative contract at different parameter inputs and approximate the sensitivity by finite difference approximation. It is easily understood that the difference between two noisy evaluations yields an even noisier evaluation, and the convergence of derivative sensitivities–Greeks–is therefore very slow.

Glasserman and Yao (1992) show that the convergence using independent simulation runs in the finite difference approximation yields a convergence rate of $n^{-1/3}$, but by using the same random number sequences for all simulation runs a convergence rate of $n^{-1/2}$ can be achieved. In practical applications however, it is found that the convergence is still painfully slow especially for discontinuous pay off functions. Broadie and Glasserman (1996) find a partial cure for the slow convergence of the Greeks. They develop the *likelihood ratio method* and the *pathwise method* in which finite difference approximations of separate simulations can be avoided. Both methods have their pros and cons. The pathwise method is fast but not very flexible as different payoff functions call for different derivations and care needs to be taken if the payoff function is not differentiable. The likelihood ratio method is pay off independent, but its use is limited to the cases where we know the density of the underlying process explicitly.

To overcome the deficiency of knowing the density explicitly Fournié et al. (1999) introduced stochastic weights derived using Malliavin calculus rather than

the explicit expression of the density. The underlying price process is assumed to be a Markov diffusion in $\mathbb{R}^n$,

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t,$$

where $\{W_t, t \geq 0\}$ is an $n$–dimensional Brownian motion. They use the Malliavin integration by parts formula to transform the problem of calculating derivatives by finite difference approximations to calculating expectations of the form

$$E\left[\phi(X_T)\pi\right],$$

where $\pi$ is a random variable.

The objective of this work is to derive stochastic weights for calculating the Greeks in a jump diffusion setting where the jump amplitude is deterministic. In particular we consider jump diffusions, $X_t$, which can be written $X_t = f(X_t^c, X_t^d)$, where $X_t^c$ is a Markov diffusion and $X_t^d$ is driven by the Poisson process and not dependent on the the initial value $X_0$. Though it might seem restrictive, this case includes some interesting price processes such as the jump diffusion version of Black–Scholes and various stochastic volatility models with added Poisson jumps.

Recently El-Khatib and Privault (2004) considered a market driven by jumps alone. Their setup allows for random jump sizes, and by imposing a regularity condition on the payoff they use Malliavin calculus on Poisson space to derive weights for Asian options. This treatment is completely orthogonal to the one presented in this thesis since the focus here is on the Wiener process rather than the Poisson process.

To derive stochastic wights for derivative sensitivities several results from Malliavin calculus are needed. An exhaustive treatment of the theory in the Gaussian case is given in Nualart (1995), and the first task of the thesis is to extend this theory to Malliavin calculus for general Lévy processes. The approach to extension of Malliavin calculus followed here was initiated in Nualart and Schoutens (2000), where chaotic and predictable representation properties for a class of Lévy processes satisfying a moment condition were derived. In the Gaussian Malliavin calculus one choice of starting point is the analogous representation properties, and in León et al. (2002) the Malliavin derivative is defined in a similar way but

with the more general representation properties of Nualart and Schoutens (2000) as starting point.

The work in Nualart and Schoutens (2000) and León et al. (2002) goes a long way towards a Malliavin calculus for Lévy processes, but stops short of defining the Skorohod integral. The contribution of this thesis is to start from that definition and follow through to stochastic weights for more general price processes and more general financial products.

## 1.2 Outline of thesis

Whereas the main aim of this thesis is to extend results regarding Monte Carlo simulation of contingent claim sensitivities, the theory leading up to the main chapter is interesting in its own right. The thesis is therefore split into two parts. The first part of the thesis is concerned with extending the existing literature on Malliavin calculus by considering more general processes. In Chapter 2 we start with general Lévy processes on which we put restrictions with aim to ensure existence of moments of all orders. The existing theory of Malliavin calculus in this setting, as presented in Nualart and Schoutens (2000) and León et al. (2002), is revisited before the Skorohod integral is defined. Some properties of the Skorohod integral are investigated and in particular it is shown how to explicitly integrate a dense set of non–adapted stochastic processes.

In Chapter 3 the scope is narrowed from the rather general Lévy processes of the previous chapter to jump diffusions with constant jump intensity and jump amplitude. The structure is adopted from the previous chapter, and hence, the extension in terms of the introduction of the Skorohod integral is preceded by a summary of results previously published in León et al. (2002).

The second part of the thesis is where the theory developed earlier is applied to the derivation of stochastic weights for calculating Greeks using Monte Carlo simulation. The first sections of Chapter 4 follows the same steps as in Fournié et al. (1999) to derive weights for what is known as *rho*, *delta* and *vega* in our generalised setting. This setting gives rise to two additional Greeks, for which stochastic weights also are derived, namely the sensitivities to jump intensity and jump am-

plitude. Examples and numerical experiments are also presented.

The more generalised setting does not only give rise to new Greeks for products considered earlier. The jump diffusion setting also allows us to look at the ever growing credit derivatives market. In Chapter 5 stochastic weights for single name and basket credit derivative Greeks are derived, and examples are compared to alternative variance reduction techniques.

Finally, in Chapter 6, the link between Malliavin weights and likelihood ratio weights, as discussed briefly above, is investigated. Sufficient conditions, in terms of properties of the price process, are derived for the existence of likelihood ratio weights. This brings light to the close relationship between the two methods.

# Part I

# MALLIAVIN CALCULUS FOR LÉVY PROCESSES

# 2. MALLIAVIN CALCULUS FOR LÉVY PROCESSES SATISFYING A MOMENT CONDITION

## 2.1 The chaotic representation property

### 2.1.1 Setup and strongly orthogonal martingales

Consider a probability space $(\Omega, \mathcal{F}, P)$ endowed with a filtration $\{\mathcal{F}_t, t \geq 0\}$, which satisfies the usual conditions. On this probability space we define a right continuous stochastic process with left limits $X$ to be a *Lévy process* if $X_0 = 0$ and for every $s, t \geq 0$ the increment $X_{t+s} - X_t$ is independent of the process $\{X_r, 0 \leq r \leq t\}$ and has the same law as $X_s$. We say that a Lévy process has stationary independent increments. A Lévy process is not necessarily continuous and we denote the discontinuous jumps by $\Delta X_t = X_t - X_{t-}$ where $X_{t-}$ is the left limit of the process.

The law of a Lévy process has an *infinitely divisible* characteristic function in that

$$E\left[e^{i\theta X_t}\right] = (\phi(\theta))^t .$$

Further, the *characteristic exponent* defined by $\psi = \log\phi(\theta)$ satisfies the *Lévy–Khintchine* formula:

$$\psi(\theta) = ia\theta - \frac{\sigma^2}{2}\theta^2 + \int_{\mathbb{R}} \left(e^{i\theta x} - 1 - i\theta x \mathbb{1}_{\{|x|<1\}}\right) \nu(dx),$$

where $a \in \mathbb{R}$, $\sigma^2 \geq 0$ and $\nu$ is the *Lévy measure* satisfying $\int_{\mathbb{R}}(1 \wedge x^2)\nu(dx) < \infty$. For a more detailed description of Lévy processes see for example Bertoin (1996).

The strategy in Nualart and Schoutens (2000) is to decompose the Lévy process $X_t$ into components, which can be used to construct pairwise strongly orthogonal

martingales. Recall that two martingales $M$ and $N$ are strongly orthogonal if their product is a uniformly integrable martingale. It can be shown that this is the case if and only if their quadratic covariation $[M, N]$ is a uniformly integrable martingale. See for example Rogers and Williams (2000) section IV26.

The chaotic representation will be given in terms of iterated Itô integrals with respect to the constructed strongly orthogonal martingales. The end product, Theorem 2.1.2, has similarities to the purely Gaussian setup where the only martingale under consideration is Brownian motion. See Nualart (1995) for an exhaustive treatment of Gaussian Malliavin calculus.

A necessary condition for the construction of the strongly orthogonal martingales is the following assumption, which we apply for the rest of the chapter.

Assumption 2.1.1: The Lévy measure $\nu$ of $X$ satisfies that there exist $\epsilon > 0$ and $\delta > 0$ such that

$$\int_{(-\epsilon,\epsilon)^c} e^{\delta|x|}\nu(dx) < \infty.$$

The assumption implies that

$$\int_{-\infty}^{\infty} |x|^n \nu(dx) < \infty, \quad n \geq 2, \tag{2.1}$$

and that the characteristic function $E[\exp(i\theta X_t)]$ is analytic in a neighbourhood of 0. This, in turn, implies that $X_t$ has moments of all orders and polynomials are dense in $L^2(\mathbb{R}, P \circ X_1^{-1})$. See Nualart and Schoutens (2000) for details.

Nualart and Schoutens continue by decomposing the process $X_t$ into components, which will be used to construct the orthogonal martingales. Define

$$\begin{aligned} X_t^{(1)} &= X_t \\ X_t^{(n)} &= \sum_{0<s\leq t} (\Delta X_s)^n, \quad n \geq 2. \end{aligned}$$

The processes $X_t^{(n)}$, $n \geq 1$ are referred to as *power jump processes* and are themselves Lévy processes. They jump at the same time as $X_t$, and further, we have that

$$\begin{aligned} E[X_t^{(1)}] &= E[X_t] =: m_1 t < \infty \\ E[X_t^{(n)}] &= t\int_{-\infty}^{\infty} x^n \nu(dx) =: m_n t < \infty, \quad n \geq 2, \end{aligned}$$

where the claim of finite moments is given by (2.1). As a consequence we can construct the compensated power jump processes $Y_t^{(n)}$, which Nualart and Schoutens refer to as *Teugels martingales*:

$$Y_t^{(n)} = X_t^{(n)} - m_n t, \quad n \geq 1.$$

The Teugels martingales are the building blocks for our pairwise strongly orthogonal martingales. Since,

$$[Y^{(k)}, Y^{(l)}]_t = X_t^{(k+l)} + \mathbf{1}_{\{k=l=1\}}\sigma^2 t, \quad k, l \geq 1,$$

we can construct processes $H_t^{(n)}$ as linear combinations of $Y^{(k)}, 1 \leq k \leq n$ such that $H^{(n)}$, $n \geq 1$ are pairwise strongly orthogonal. In particular we can choose $a_k^{(n)}, 1 \leq k \leq n-1$ such that for

$$H_t^{(n)} = Y_t^{(n)} + a_{n-1}^{(n)} Y_t^{(n-1)} + \cdots + a_1^{(n)} Y^{(1)},$$

we have $E\left[[H^{(k)}, H^{(l)}]_1\right] = 0$ for $k \neq l$. In Nualart and Schoutens (2000) it is shown how this Gram-Schmidt type orthogonalisation can be done by identifying an isometry between the space of linear combinations of Teugels martingales and the space of real polynomials. Examples are worked out where the coefficients $a_k^{(n)}$ for certain Lévy processes can be found as coefficients of well known orthogonal polynomials.

### 2.1.2 The chaotic and the predictable representation property

As mentioned above the chaotic representation property will be given in terms of iterated integrals. We define the iterated integral of $n$–th order of some square integrable deterministic function $f$ as

$$L_n^{(i_1,\ldots,i_n)}(f) = \int_0^\infty \left( \int_0^{t_n-} \cdots \left( \int_0^{t_2-} f(t_1, \ldots, t_n) dH_{t_1}^{(i_1)} \right) \ldots dH_{t_{n-1}}^{(i_{n-1})} \right) dH_{t_n}^{(i_n)}.$$

It will also be useful to introduce the notation

$$\Sigma_n = \{(t_1, \ldots, t_n) \in \mathbb{R}_+^n : 0 < t_1 < t_2 < \ldots < t_n\},$$

and

$$\begin{aligned} \Sigma_n^{(k)}(t) \quad = \quad & \{(t_1, \ldots, t_{k-1}, t_{k+1}, \ldots, t_n) \in \Sigma_{n-1} : \\ & 0 < t_1 < \ldots < t_{k-1} < t < t_{k+1} < \ldots < t_n\}. \end{aligned}$$

We are now in a position to state the main result of Nualart and Schoutens (2000), which is the following *Chaotic representation property*.

Theorem 2.1.2 (Nualart and Schoutens): Let $F \in L^2(\Omega)$, then $F$ has a representation of the form

$$F = E[F] + \sum_{n=1}^{\infty} \sum_{i_1,\ldots,i_n \geq 1} L_n^{(i_1,\ldots,i_n)}(f_{i_1,\ldots,i_n}),$$

where $f_{i_1,\ldots,i_n} \in L^2(\Sigma_n)$.

Once the chaotic representation property is established the following *predictable representation property* follows as a direct consequence.

Theorem 2.1.3 (Nualart and Schoutens): Let $F \in L^2(\Omega)$, then $F$ has a representation of the form

$$F = E[F] + \sum_{n=1}^{\infty} \int_0^{\infty} \phi_t^{(n)} dH_t^{(n)},$$

where $\phi_t^{(n)}$, $n \geq 1$ are predictable processes.

## 2.2 Malliavin calculus

### 2.2.1 Malliavin derivative

In León et al. (2002) the authors build on the chaotic and the predictable representation properties of Nualart and Schoutens (2000) and define a Malliavin derivative. The predictable representation property suggests the existence of a Clark–Ocone formula, where the predictable process in Theorem 2.1.3 is replaced by the previsible projection of the Malliavin derivative of the random variable in question. This Clark–Ocone formula is one of the main results of León et al. (2002).

As in the Gaussian case, not all square integrable random variables are Malliavin differentiable. For the definition of the Malliavin derivative we must first define the spaces of random variables, which are differentiable in the $l$–th direction, $l \geq 1$. In the definition below we write $q_i t = E\left[[H^{(i)}]_t\right]$.

Definition 2.2.1 (Differentiability in the $l$–th direction): We say that $F$ is differentiable in the $l$–th direction ($l \geq 1$) if $F \in \mathbb{D}^{(l)}$ where

$$\begin{aligned}\mathbb{D}^{(l)} = \{F \in L^2(\Omega),\, F = E[F] + \sum_{n=1}^{\infty} \sum_{i_1,\dots,i_n \geq 1} L_n^{(i_1,\dots,i_n)}(f_{i_1,\dots,i_n}) : \\ \sum_{n=1}^{\infty} \sum_{i_1,\dots,i_n \geq 1} \sum_{k=1}^{n} \mathbf{1}_{\{i_k = l\}} q_{i_1} \cdots q_{i_{k-1}} q_{i_{k+1}} \cdots q_{i_n} \\ \int_0^{\infty} \| f_{i_1,\dots,i_n}(\dots,t,\dots) \mathbf{1}_{\{\Sigma_n^{(k)}(t)\}} \|^2_{L^2([0,\infty)^{n-1})} dt < \infty\}.\end{aligned}$$

This definition ensures that the Malliavin derivative defined below is in $L^2(\Omega \times \mathbb{R}_+)$. In fact, the spaces $\mathbb{D}^{(l)}$ are all dense in $L^2(\Omega)$ since random variables with finite chaos expansion are in $\mathbb{D}^{(l)}$. It can also be shown that if we remove the discontinuous part of $X_t$ the space $\mathbb{D}^{(1)}$ collapses to the classical Gaussian space $\mathbb{D}^{1,2}$ discussed e.g. in Nualart (1995).

Definition 2.2.2 (Malliavin derivative): For $F \in \mathbb{D}^{(l)}$ such that

$$F = E[F] + \sum_{n=1}^{\infty} \sum_{i_1,\dots,i_n \geq 1} L_n^{(i_1,\dots,i_n)}(f_{i_1,\dots,i_n}),$$

we define the derivative of $F$ in the $l$–th direction as the element of $L^2(\Omega \times \mathbb{R}_+)$ given by

$$D_t^{(l)} F = \sum_{n=1}^{\infty} \sum_{i_1,\dots,i_n \geq 1} \sum_{k=1}^{n} \mathbf{1}_{\{i_k = l\}} L_{n-1}^{(i_1,\dots,i_{k-1},i_{k+1},\dots,i_n)} \left( f_{i_1,\dots,i_n}(\dots,t,\dots) \mathbf{1}_{\{\Sigma_n^{(k)}(t)\}} \right).$$

With this definition of the Malliavin derivative we can state the advertised *Clark–Ocone formula*. The proof follows from the chaotic representation property, Theorem 2.1.2.

Theorem 2.2.3 (León, Solé, Utzet and Vives): If $F \in \bigcap_{j=0}^{\infty} \mathbb{D}^{(j)}$ then

$$F = E[F] + \sum_{j=1}^{\infty} \int_0^{\infty} {}^{p}(D_t^{(j)} F) dH_t^{(j)}.$$

We end this section with a list of properties of the Malliavin derivative, which follow directly from the definition.

- The Malliavin derivative is a linear operator;
- $D_t^{(l)} F = 0$ for all $l$ and $t$ if and only if $F$ is deterministic;
- If $F \in \mathcal{F}_t$, then $D_s^{(l)} F = 0$ for all $s > t$ and $l \geq 1$.

### 2.2.2 Skorohod integral

Now we continue where León et al. (2002) stopped. Since the Malliavin derivative is a densely defined operator we can define its adjoint, which as usual will be called the Skorohod integral.

Definition 2.2.4 (Skorohod integral): Let $u_t$ be a stochastic process in $L^2(\Omega \times \mathbb{R}^+)$ not necessarily adapted such that

$$\left| E\left[ \int_0^\infty (D_t^{(l)} F) u_t q_l dt \right] \right| < c \|F\|_{L^2(\Omega)},$$

for some constant $c$ depending on $u$ and any $F \in \mathbb{D}^{(l)}$. We say that $u$ is Skorohod integrable in the $l$–th direction or $u \in \operatorname{Dom} \delta^{(l)}$. We define the Skorohod integral in the $l$–th direction, $\delta^{(l)}(u)$, as the operator mapping $L^2(\Omega \times \mathbb{R}^+)$ to $L^2(\Omega)$ for which

$$E\left[ \int_0^\infty (D_t^{(l)} F) u_t q_l dt \right] = E\left[ F \delta^{(l)}(u) \right],$$

for any $F \in \mathbb{D}^{(l)}$.

Due to the linearity of the Malliavin derivative it follows from the definition that the Skorohod integral is a linear operator. It is also immediate that the Skorohod integral is a closed operator since adjoints to densely defined operators always are.[1] To further investigate the properties of the Skorohod integral the next result gives an explicit expression for the Skorohod integral in a simple case.

Lemma 2.2.5: Let $h(t) \in L^2(\mathbb{R}^+)$ and $G \in L^2(\Omega)$ with finite chaos expansion:

$$G = E[G] + \sum_{n=1}^{N} \sum_{0 < i_1, \ldots, i_n \leq N} L_n^{(i_1, \ldots, i_n)}(g_{i_1, \ldots, i_n}).$$

Then $Gh \in \operatorname{Dom} \delta^{(l)}$ for $1 \leq l \leq m$ and

$$\begin{aligned} \delta^{(l)}(Gh) = {} & \int_0^\infty E[G] h(t_1) dH_{t_1}^{(l)} + \sum_{n=1}^{N} \sum_{0 < i_1, \ldots, i_n \leq N} \sum_{k=1}^{n+1} \int_0^\infty \\ & \cdots \int_0^{t_{k+2}-} \int_0^{t_{k+1}-} \int_0^{t_k-} \cdots \int_0^{t_2-} g_{i_1, \ldots, i_n}(t_1, \ldots, t_{k-1}, t_{k+1}, \ldots, t_{n+1}) \\ & h(t_k) dH_{t_1}^{(i_1)} \ldots dH_{t_{k-1}}^{(i_{k-1})} dH_{t_k}^{(l)} dH_{t_{k+1}}^{(i_{k+1})} \ldots dH_{t_{n+1}}^{(i_n)}. \end{aligned}$$

[1] See for example Debnath and Mikusiński (1999) Theorem 4.12.6.

*Proof.* For any $F \in \mathbb{D}^{(l)}$ we have

$$E\left[\int_0^\infty (D_t^{(l)}F)Gh(t)q_l dt\right]$$
$$= q_l \int_0^\infty E\left[\left(\sum_{n=1}^\infty \sum_{i_1,\dots,i_n\geq 1} \sum_{k=1}^n \mathbf{1}_{\{i_k=l\}} L_{n-1}^{(i_1,\dots,i_{k-1},i_{k+1},\dots,i_n)}\left(f_{i_1,\dots,i_n}(\dots,t,\dots)\mathbf{1}_{\{\Sigma_n^{(k)}(t)\}}\right)\right)\right.$$
$$\left.\left(E[G] + \sum_{n=1}^N \sum_{0<i_1,\dots,i_n\leq N} L_n^{(i_1,\dots,i_n)}(g_{i_1,\dots,i_n})\right)h(t)\right]dt.$$

We have (by isometry as shown in León et al. (2002)) that

$$E[L_n^{(i_1,\dots,i_n)}(f)L_m^{(j_1,\dots,j_m)}(\tilde{f})] = \begin{cases} q_{i_1}\dots q_{i_n}\int_{\Sigma_n} f\tilde{f}\,dt_1\dots dt_n, \\ \qquad\qquad \text{if multi–indices identical,} \\ 0, \text{ otherwise,} \end{cases}$$

so we get (note how the infinite sum for $F$ is capped due to the above fact)

$$E\left[\int_0^\infty (D_t^{(l)}F)Gh(t)q_l dt\right] = q_l \int_0^\infty \left(f_l(t)h(t)E[G]\right.$$
$$+ \sum_{n=2}^{N+1} \sum_{0<i_1,\dots,i_n\leq N} \sum_{k=1}^n \mathbf{1}_{\{i_k=l\}} q_{i_1}\dots q_{i_{k-1}}q_{i_{k+1}}\dots q_{i_n}$$
$$\int_0^\infty \dots \int_0^{t_{k+2}} \int_0^{t_{k+1}} \dots \int_0^{t_2} f_{i_1,\dots,i_n}(\dots,t,\dots)\mathbf{1}_{\{\Sigma_n^{(k)}(t)\}}$$
$$\left. g_{i_1,\dots,i_{k-1},i_{k+1},\dots,i_n}(t_1,\dots,t_{k-1},t_{k+1},\dots,t_n)h(t)dt_1\dots dt_{k-1}dt_{k+1}\dots dt_n\right)dt.$$

We can 'bring the $dt$ integration in to the $k$th position' and transform back into stochastic integrals:

$$E\left[\int_0^\infty (D_t^{(l)}F)Ghq_l dt\right] = E\left[\left(E[F] + \sum_{n=1}^\infty \sum_{i_1,\dots,i_n\geq 1} \int_0^\infty \int_0^{t_n-}\right.\right.$$
$$\left.\dots \int_0^{t_2-} f_{i_1,\dots,i_n}(t_1,\dots,t_n)dH_{t_1}^{(i_1)}\dots dH_{t_{n-1}}^{(i_{n-1})}dH_{t_n}^{(i_n)}\right)$$
$$\times \left(\int_0^\infty E[G]h(t_1)dH_{t_1}^{(l)} + \sum_{n=1}^N \sum_{0<i_1,\dots,i_n\leq N} \sum_{k=1}^{n+1} \int_0^\infty \dots \int_0^{t_{k+2}-} \int_0^{t_{k+1}-} \int_0^{t_k-}\right.$$
$$\dots \int_0^{t_2-} g_{i_1,\dots,i_n}(t_1,\dots,t_{k-1},t_{k+1},\dots,t_{n+1})h(t_k)dH_{t_1}^{(i_1)}$$
$$\left.\left.\dots dH_{t_{k-1}}^{(i_{k-1})}dH_{t_k}^{(l)}dH_{t_{k+1}}^{(i_{k+1})}\dots dH_{t_{n+1}}^{(i_n)}\right)\right].$$

An application of the Schwarz inequality ensures that the condition in Definition 2.2.4 is satisfied, and therefore, completes the proof. □

Corollary 2.2.6: The Skorohod integral is a densely defined operator in $L^2(\Omega \times \mathbb{R}^+)$.

*Proof.* Processes of the form $Hg(t)$ where $H \in L^2(\Omega)$ with finite chaos expansion and $g \in L^2(\mathbb{R}^+)$ are dense in $L^2(\Omega \times \mathbb{R}^+)$. □

We see that Skorohod integration increases the order of the chaos expansion with a $dH^{(l)}$ integral mixed in at every possible position. This is inline with Skorohod integration in the Gaussian case as considered for example in Øksendal (1996).

The result in Lemma 2.2.5 can be extended to processes where the stochastic part does not have a finite chaos expansion. In this more general case a condition for Skorohod integrability is present as shown in the following Proposition.

Proposition 2.2.7: Let $h(t) \in L^2(\mathbb{R}^+)$ and $G \in L^2(\Omega)$ with chaos expansion

$$G = E[G] + \sum_{n=1}^{\infty} \sum_{i_1,\ldots,i_n \geq 1} L_n^{(i_1,\ldots,i_n)}(g_{i_1,\ldots,i_n}).$$

Then

$$\begin{aligned}\delta^{(l)}(Gh) = \int_0^\infty E[G]h(t_1)dH_{t_1}^{(l)} + \sum_{n=1}^{\infty} \sum_{i_1,\ldots,i_n \geq 1} \sum_{k=1}^{n+1} \int_0^\infty \\ \cdots \int_0^{t_{k+2}-} \int_0^{t_{k+1}-} \int_0^{t_k-} \cdots \int_0^{t_2-} g_{i_1,\ldots,i_n}(t_1,\ldots,t_{k-1},t_{k+1},\ldots,t_{n+1}) \\ h(t_k)dH_{t_1}^{(i_1)} \ldots dH_{t_{k-1}}^{(i_{k-1})} dH_{t_k}^{(l)} dH_{t_{k+1}}^{(i_{k+1})} \ldots dH_{t_{n+1}}^{(i_n)}, \quad (2.2)\end{aligned}$$

in the sense that $Gh \in \operatorname{Dom} \delta^{(l)}$ if the infinite sum converges in $L^2(\Omega)$.

*Proof.* We can approximate $G$ by $G^N$ defined by

$$G^N = E[G] + \sum_{n=1}^{N} \sum_{0 < i_1,\ldots,i_n \leq N} L_n^{(i_1,\ldots,i_n)}(g_{i_1,\ldots,i_n}).$$

In fact, in Nualart and Schoutens (2000) it is shown that with capped chaos expansions, as the one above, we can represent all polynomials with order less than or equal to $N$. Hence, when $N$ goes to infinity $G^N h$ converges to $Gh$.

Using Lemma 2.2.5 we can integrate $G^N$ for all $N$, and $\delta^{(l)}(G^N h)$ converges as $N$ goes to infinity if and only if the infinite sum in (2.2) converges. Further, since $\delta^{(l)}$ is a closed operator $\delta^{(l)}(Gh) = \lim_{N\to\infty} \delta^{(l)}(G^N h)$ if the limit exists. □

Proposition 2.2.7 can be seen as the foundation for Skorohod integration of Lévy processes satisfying Assumption 2.1.1. We now have a way of integrating dense building blocks of $L^2(\Omega \times \mathbb{R}^+)$ in the same way as *elementary processes* are dense building blocks for Itô integration. In fact, since the Skorohod integral is a densely defined closed operator we can integrate any Skorohod integrable process by approximating it by sums of $Hg(t)$ processes as in Proposition 2.2.7.

The next task is to show that the classical relationship between Itô and Skorohod integration holds in our setting. That is, when previsible processes are integrated the two integrals coincide. We follow the route set out in Øksendal (1996) and start with investigating what adaptedness means in terms of iterated integrals.

Lemma 2.2.8: Let $h(t) \in L^2(\mathbb{R}^+)$ and $G \in L^2(\Omega)$ with chaos expansion

$$G = E[G] + \sum_{n=1}^{\infty} \sum_{i_1,\dots,i_n \geq 1} L_n^{(i_1,\dots,i_n)}(g_{i_1,\dots,i_n}).$$

The stochastic process $Gh(t)$ is $\mathcal{F}_{t-}$–adapted if and only if for all $(i_1,\dots,i_n)$

$$g_{i_1,\dots,i_n}(t_1,\dots,t_n)h(t) = 0 \text{ for } \max_{j=1,\dots,n} t_j > t.$$

That is, if and only if $g_{i_1,\dots,i_n}(t_1,\dots,t_n) = 0$ when the first $t_i$ passes the first $t$ when $h(t) \neq 0$.

*Proof.* For each term in $Gh(t)$ we have

$$\begin{aligned}
&E\left[L_n^{(i_1,\dots,i_n)}(g_{i_1,\dots,i_n})h(t)|\mathcal{F}_{t-}\right] \\
&\quad = E\left[\int_0^\infty \int_0^{t_n-} \cdots \int_0^{t_2-} g_{i_1,\dots,i_n}(t_1,\dots,t_n) dH_{t_1}^{(i_1)} \dots dH_{t_{n-1}}^{(i_{n-1})} dH_{t_n}^{(i_n)} h(t) \Big| \mathcal{F}_{t-}\right] \\
&\quad\quad = \int_0^{t-} \int_0^{t_n-} \cdots \int_0^{t_2-} g_{i_1,\dots,i_n}(t_1,\dots,t_n) dH_{t_1}^{(i_1)} \dots dH_{t_{n-1}}^{(i_{n-1})} dH_{t_n}^{(i_n)} h(t) \\
&\quad\quad\quad = L_n^{(i_1,\dots,i_n)}\left(g_{i_1,\dots,i_n}(t_1,\dots,t_n)\mathbb{1}_{\{\max_{j=1,\dots,n} t_j < t\}} h(t)\right).
\end{aligned}$$

□

An example of a $g_{i_1,\ldots,i_n}h$–pair satisfying the condition in Lemma 2.2.8 is

$$\tilde{g}_{i_1,\ldots,i_n}(t_1,\ldots,t_n)\mathbb{1}_{\{\max_{j=1,\ldots,n} t_j<t_0\}}\tilde{h}(t)\mathbb{1}_{\{t>t_0\}},$$

for some $t_0 > 0$. As usual we define a previsible process to be a square integrable process adapted to the left limit of the filtration under consideration. The following proposition make use of the previous lemma and shows that the Itô and Skorohod integral coincides when previsible processes are integrated.

Proposition 2.2.9: If $E\left[\int_0^\infty (Gh(t))^2\,dt\right] < \infty$ and $Gh(t)$ is $\mathcal{F}_{t-}$–adapted, then $Gh \in \operatorname{Dom}\delta^{(l)}$ for $l \geq 1$ and

$$\delta^{(l)}(Gh) = \int_0^\infty Gh(t)dH_t^{(l)}.$$

*Proof.* By Proposition 2.2.7 we have that $Gh \in \operatorname{Dom}\delta^{(l)}$ if the following sum converges in $L^2(\Omega)$:

$$\int_0^\infty E[G]h(t_1)dH_{t_1}^{(l)} + \sum_{n=1}^\infty \sum_{i_1,\ldots,i_n\geq 1}\sum_{k=1}^{n+1}\int_0^\infty \cdots \int_0^{t_{k+2}-}\int_0^{t_{k+1}-}\int_0^{t_k-}\cdots\int_0^{t_2-} g_{i_1,\ldots,i_n}(t_1,\ldots,t_{k-1},t_{k+1},\ldots,t_{n+1}) h(t_k)dH_{t_1}^{(i_1)}\ldots dH_{t_{k-1}}^{(i_{k-1})}dH_{t_k}^{(l)}dH_{t_{k+1}}^{(i_{k+1})}\ldots dH_{t_{n+1}}^{(i_n)}. \quad (2.3)$$

Because $Gh(t)$ is $\mathcal{F}_{t-}$–adapted we have from Lemma 2.2.8 that all but one term in the innermost sum vanishes so that (2.3) can be simplified as

$$\int_0^\infty E[G]h(t_1)dH_{t_1}^{(l)} + \sum_{n=1}^\infty \sum_{i_1,\ldots,i_n\geq 1}\int_0^\infty \int_0^{t_{n+1}-}\cdots\int_0^{t_2-} g_{i_1,\ldots,i_n}(t_1,\ldots,t_n) h(t_{n+1})dH_{t_1}^{(i_1)}\ldots dH_{t_n}^{(i_n)}dH_{t_{n+1}}^{(l)}. \quad (2.4)$$

By the same isometry argument as in the proof of Lemma 2.2.5 we have that the $L^2(\Omega)$ norm of (2.4) is equal to

$$q_l \int_0^\infty (E[G])^2 h^2(t_1)dt_1 + q_{i_1}\dots q_{i_n}q_l \sum_{n=1}^\infty \sum_{i_1,\dots,i_n\geq 1} \int_0^\infty \int_0^{t_{n+1}} \dots \int_0^{t_2} g^2_{i_1,\dots,i_n}(t_1,\dots,t_n)h^2(t_{n+1})dt_1\dots dt_n dt_{n+1}$$

$$= q_l \int_0^\infty (E[G])^2 h^2(t_1)dt_1 + q_{i_1}\dots q_{i_n}q_l \sum_{n=1}^\infty \sum_{i_1,\dots,i_n\geq 1} \int_0^\infty \int_0^\infty \dots \int_0^{t_2} g^2_{i_1,\dots,i_n}(t_1,\dots,t_n)h^2(t_{n+1})dt_1\dots dt_n dt_{n+1}$$

$$= q_l \int_0^\infty h^2(t_1)dt_1 (E[G])^2 + q_{i_1}\dots q_{i_n}q_l \int_0^\infty h^2(t_{n+1})dt_{n+1} \sum_{n=1}^\infty \sum_{i_1,\dots,i_n\geq 1} \int_0^\infty \int_0^{t_n} \dots \int_0^{t_2} g^2_{i_1,\dots,i_n}(t_1,\dots,t_n)dt_1\dots dt_{n-1}dt_n$$

$$= q_l \int_0^\infty h^2(t)dt E[G^2] < \infty,$$

by assumption. Further, $\delta^{(l)}(Gh)$ is explicitly given by (2.4), so

$$\delta^{(l)}(Gh) = \int_0^\infty E[G]h(t_1)dH^{(l)}_{t_1} + \sum_{n=1}^\infty \sum_{i_1,\dots,i_n\geq 1} \int_0^\infty \int_0^\infty \dots \int_0^{t_2-} g_{i_1,\dots,i_n}(t_1,\dots,t_n) h(t_{n+1})dH^{(i_1)}_{t_1}\dots dH^{(i_n)}_{t_n}dH^{(l)}_{t_{n+1}}$$

$$= \int_0^\infty h(t_1)dH^{(l)}_{t_1}E[G] + \int_0^\infty h(t_{n+1}) \left( \sum_{n=1}^\infty \sum_{i_1,\dots,i_n\geq 1} \int_0^\infty \int_0^{t_n-} \dots \int_0^{t_2-} g_{i_1,\dots,i_n}(t_1,\dots,t_n)dH^{(i_1)}_{t_1}\dots dH^{(i_{n-1})}_{t_{n-1}}dH^{(i_n)}_{t_n} \right) dH^{(l)}_{t_{n+1}}$$

$$= \int_0^\infty Gh(t)dH^{(l)}_t. \qquad \square$$

**Corollary 2.2.10:** If $E\left[\int_0^\infty u_t^2 dt\right] < \infty$ and $u_t$ is $\mathcal{F}_{t-}$–adapted, then $u \in \mathrm{Dom}\,\delta^{(l)}$ for $l \geq 1$ and

$$\delta^{(l)}(u_t) = \int_0^\infty u_t dH^{(l)}_t.$$

*Proof.* The processes considered in Proposition 2.2.9 include elementary processes. By approximating $u_t$ by elementary processes we get

$$\delta^{(l)}\left(\sum_{i=1}^{M_n} G^{(i)}_n h^{(i)}_n\right) = \sum_{i=1}^{M_n} \int_0^\infty G^{(i)}_n h^{(i)}_n(t)dH^{(l)}_t \to \int_0^\infty u_t dH^{(l)}_t \text{ as } n \to \infty,$$

and the result follows by the fact that the Skorohod integral is a closed operator. $\square$

# 3. MALLIAVIN CALCULUS FOR SIMPLE LÉVY PROCESSES

## 3.1 Setup and representation theorems

Following on from the previous chapter we will now consider a special case of Lévy process satisfying Assumption 2.1.1. On a filtered probability space $(\Omega, \mathcal{F}, P)$, define a *simple Lévy process* as given by

$$X_t = \sigma W_t + \alpha_1 N_t^{(1)} + \cdots + \alpha_m N_t^{(m)}, \quad t \geq 0,$$

where $\{W_t,\, t \geq 0\}$ is a standard Brownian motion and $\{N_t^{(j)},\, t \geq 0\}$, $j = 1, \ldots, m$, are mutually independent Poisson processes with intensities $\lambda_1, \ldots, \lambda_m$, which are also independent of the Brownian motion. The jump amplitudes $\alpha_j$, $j = 1, \ldots, m$ are different non-null constants. Also, let the filtration $\mathcal{F}_t$ be the one generated by the simple Lévy process $X_t$.

This chapter follows the previous in the way the Malliavin calculus is developed. Hence, the idea is to represent random variables on the probability space $(\Omega, \mathcal{F}, P)$ by iterated integrals, and from there define the Malliavin derivative and the Skorohod integral. The difference is in the way the representation is achieved. In this much simpler case León et al. (2002) derive a representation not in terms of power jump processes as in Theorem 2.1.2, but in terms of the Brownian motion and the compensated Poisson processes instead. We use the notation

$$\begin{aligned} G_t^{(0)} &= W_t \\ G_t^{(j)} &= N_t^{(j)} - \lambda_j t, \quad j = 1, \ldots, m, \end{aligned}$$

and define $L_n^{(i_1,\ldots,i_n)}(f)$ to be the iterated integral of some deterministic square in-

tegrable function $f$ with respect to $G_t^{(i_1)}, \ldots, G_t^{(i_n)}$:

$$L_n^{(i_1,\ldots,i_n)}(f) = \int_0^\infty \left( \int_0^{t_n^-} \ldots \left( \int_0^{t_2^-} f(t_1,\ldots,t_n) dG_{t_1}^{(i_1)} \right) \ldots dG_{t_{n-1}}^{(i_{n-1})} \right) dG_{t_n}^{(i_n)}.$$

Recall that

$$\Sigma_n = \{(t_1,\ldots,t_n) \in \mathbb{R}_+^n : 0 < t_1 < t_2 < \ldots < t_n\},$$

and

$$\begin{aligned} \Sigma_n^{(k)}(t) &= \{(t_1,\ldots,t_{k-1},t_{k+1},\ldots,t_n) \in \Sigma_{n-1} : \\ &\qquad 0 < t_1 < \ldots < t_{k-1} < t < t_{k+1} < \ldots < t_n\}. \end{aligned}$$

With these modified definitions of iterated integrals León et al. (2002) follow Nualart and Schoutens (2000) and formulate the following chaotic representation property, which is a modification of Theorem 2.1.2.

Theorem 3.1.1 (León, Solé, Utzet and Vives): Let $F \in L^2(\Omega, \mathcal{F}, P)$, then $F$ has a representation of the form

$$F = E[F] + \sum_{n=1}^{\infty} \sum_{0 \le i_1,\ldots,i_n \le m} L_n^{(i_1,\ldots,i_n)}(f_{i_1,\ldots,i_n}),$$

where $f_{i_1,\ldots,i_n} \in L^2(\Sigma_n)$.

Note that in the simple Lévy process case the number of orthogonal martingales is finite and not infinite as in the general case, and hence the second sum in Theorem 3.1.1 is finite given $n$. As was the case in the previous chapter the predictable representation property follows directly from the chaotic representation.

Theorem 3.1.2 (León, Solé, Utzet and Vives): Let $F \in L^2(\Omega)$, then $F$ has a representation of the form

$$F = E[F] + \sum_{n=1}^{m} \int_0^\infty \phi_t^{(n)} dG_t^{(n)},$$

where $\phi_t^{(n)}$, $1 \le n \le m$ are predictable processes.

## 3.2 Malliavin calculus

### 3.2.1 Malliavin derivative

We continue to follow the path set out in Chapter 2. The following definitions of differentiability and derivative are the analogues of Definition 2.2.1 and Definition 2.2.2.

Definition 3.2.1 (Differentiability in the $l$–th direction): We say that $F$ is differentiable in the $l$–th direction ($l = 0, \dots, m$) if $F \in \mathbb{D}^{(l)}$ where

$$\begin{aligned}\mathbb{D}^{(l)} \;=\; \{F \in L^2(\Omega),\, F = E[F] + \sum_{n=1}^{\infty} \sum_{0 \le i_1,\dots,i_n \le m} L_n^{(i_1,\dots,i_n)}(f_{i_1,\dots,i_n}) :\\ \sum_{n=1}^{\infty} \sum_{0 \le i_1,\dots,i_n \le m} \sum_{k=1}^{n} \mathbf{1}_{\{i_k = l\}} \lambda_{i_1} \dots \lambda_{i_{k-1}} \lambda_{i_{k+1}} \dots \lambda_{i_n}\\ \int_0^{\infty} \| f_{i_1,\dots,i_n}(\dots, t, \dots) \mathbf{1}_{\{\Sigma_n^{(k)}(t)\}} \|^2_{L^2([0,\infty)^{n-1})} dt < \infty\},\end{aligned}$$

and $\lambda_0 = 1$.

Again, the spaces $\mathbb{D}^{(l)}$ are all dense in $L^2(\Omega)$ and the above definition ensures that the Malliavin derivative defined below is in $L^2(\Omega \times \mathbb{R}_+)$.

Definition 3.2.2 (Malliavin derivative): For $F \in \mathbb{D}^{(l)}$ such that

$$F = E[F] + \sum_{n=1}^{\infty} \sum_{0 \le i_1,\dots,i_n \le m} L_n^{(i_1,\dots,i_n)}(f_{i_1,\dots,i_n}),$$

we define the derivative of $F$ in the $l$–th direction as the element of $L^2(\Omega \times \mathbb{R}_+)$ given by

$$D_t^{(l)} F = \sum_{n=1}^{\infty} \sum_{0 \le i_1,\dots,i_n \le m} \sum_{k=1}^{n} \mathbf{1}_{\{i_k = l\}} L_{n-1}^{(i_1,\dots,i_{k-1},i_{k+1},\dots,i_n)} \left( f_{i_1,\dots,i_n}(\dots, t, \dots) \mathbf{1}_{\{\Sigma_n^{(k)}(t)\}} \right).$$

We can now state the following *Clark–Ocone formula* and the *chain rule*.

Theorem 3.2.3 (León, Solé, Utzet and Vives): If $F \in \bigcap_{j=0}^{m} \mathbb{D}^{(j)}$ then

$$F = E[F] + \int_0^{\infty} {}^{p}(D_t^{(0)} F) dW_t + \sum_{j=1}^{m} \int_0^{\infty} {}^{p}(D_t^{(j)} F) d(N_t^{(j)} - \lambda_j t)$$

Theorem 3.2.4 (León, Solé, Utzet and Vives): Let $F = f(Z, Z') \in L^2(\Omega)$, where $Z$ only depends on the Brownian motion $W$, and $Z'$ only depends on the Poisson processes $N^{(1)}, \ldots, N^{(m)}$. Assume that $f(x,y)$ is a continously differentiable function with bounded partial derivatives in the variable x, and that $Z \in \mathbb{D}^{(0)}$. Then $F \in \mathbb{D}^{(0)}$ and

$$D^{(0)}F = \frac{\partial f}{\partial x}(Z, Z')D^{(0)}Z$$

where $D^{(0)}Z$ is the usual Gaussian Malliavin derivative.

The chain rule is one of the main results in León et al. (2002) and it will play a central role in the next chapter. It allows us to calculate the Malliavin derivative in $W$ direction using the classical rules. This, together with the notion of Skorohod integration, is what allows us to derive the stochastic weights for calculating the Greeks using Monte Carlo simulation.

For the theory to work in a multidimensional setting we add Brownian motions as

$$\begin{aligned} G_t^{(j)} &= W_t^{(j)}, \quad j = 1, \ldots, d \\ G_t^{(j)} &= (N_t^{(j-d)} - \lambda_{j-d}t), \quad j = d+1, \ldots, d+m. \end{aligned}$$

Definition 3.2.1 with $l = 1, \ldots, d$ defines spaces of random variables differentiable with respect to the respective Brownian motion and we group the $d$ Brownian motions together in one $d$-dimensional column vector denoted simply by $W_t$. For a random variable $F \in \bigcap_{i=1}^{d} \mathbb{D}^{(i)}$ we write $D_t^{(0)}F$—the Malliavin derivative with respect to Brownian motion—as a row vector where each component is the Malliavin derivative as defined in Definition 3.2.2.

For functions $F \in \mathbb{D}^{(l+d)}$, $l = 1, \ldots, m$ we will denote by $D_t^{(l)}F$ the Malliavin derivative defined as

$$D_t^{(l)}F = \sum_{n=1}^{\infty} \sum_{0 \le i_1, \ldots, i_n \le m} \sum_{k=1}^{n} \mathbf{1}_{\{i_k = l+d\}} L_{n-1}^{(i_1, \ldots, i_{k-1}, i_{k+1}, \ldots, i_n)} \Big( f_{i_1, \ldots, i_n}(\ldots, t, \ldots) \mathbf{1}_{\{\Sigma_n^{(k)}(t)\}} \Big).$$

Note that with these new definitions $D_t^{(0)}F$ will be a $d$-dimensional row vector, whereas $D_t^{(i)}F$, $i = 1, \ldots, m$ will be scalars, each denoting the Malliavin derivative with respect to the $i$th Poisson process. Now the above theory still holds true with $d$-dimensional Brownian noise.

### 3.2.2 Skorohod integral

We now continue the development of a Malliavin calculus for simple Lévy processes by defining a Skorohod integral as the adjoint of the derivative operator. We can do this since $D$ is a densely defined operator.

Definition 3.2.5 (Skorohod integral): Let $u_t$ be a stochastic process in $L^2(\Omega \times \mathbb{R}^+)$ not necessarily adapted such that

$$\left| E\left[ \int_0^\infty (D_t^{(l)} F) u_t \lambda_l dt \right] \right| < c \|F\|_{L^2(\Omega)},$$

for some constant $c$ depending on $u$ and any $F \in \mathbb{D}^{(l)}$. We say that $u$ is Skorohod integrable in the $l$–th direction or $u \in \mathrm{Dom}\, \delta^{(l)}$. We define the Skorohod integral in the $l$–th direction, $\delta^{(l)}(u)$, as the operator mapping $L^2(\Omega \times \mathbb{R}^+)$ to $L^2(\Omega)$ for which

$$E\left[ \int_0^\infty (D_t^{(l)} F) u_t \lambda_l dt \right] = E\left[ F \delta^{(l)}(u) \right],$$

for any $F \in \mathbb{D}^{(l)}$.

In the case of multidimensional Brownian motion we will denote by $\delta^{(0)}(u)$ the Skorohod integral of a vector process $u$. Analogous to the Itô integral the Skorohod integral is here the sum of the Skorohod integral of the components of $u$ integrated with respect to its respective Brownian motion. In the remainder of this section the main results will be proven for one–dimensional Brownian motion with multidimensional extensions given as remarks.

Below it will be shown that, as was the case in the previous chapter, the Skorohod integral is a densely defined linear operator as is the Gaussian Skorohod integral. The Skorohod integral in Definition 3.2.5 has other properties in common with its Gaussian analogue. One such similarity that will be discussed here is the result presented in Proposition 3.2.10, which provides a good tool for calculating the Skorohod integral with respect to Brownian motion. First, however, some results parallel to the ones in the previous chapter will be presented, starting with the an explicit expression for the Skorohod integral for certain processes.

**Proposition 3.2.6:** Let $g(t) \in L^2(\mathbb{R}^+)$ and $F \in L^2(\Omega)$ with chaos expansion

$$F = E[F] + \sum_{n=1}^{\infty} \sum_{0 \le i_1, \ldots, i_n \le m} L_n^{(i_1, \ldots, i_n)}(f_{i_1, \ldots, i_n}).$$

Then

$$\delta^{(l)}(Fg) = \int_0^\infty E[F]g(t_1)dG_{t_1}^{(l)} + \sum_{n=1}^{\infty} \sum_{0 \le i_1, \ldots, i_n \le m} \sum_{k=1}^{n+1} \int_0^\infty \cdots \int_0^{t_{k+2}-} \int_0^{t_{k+1}-} \int_0^{t_k-} \cdots \int_0^{t_2-} f_{i_1, \ldots, i_n}(t_1, \ldots, t_{k-1}, t_{k+1}, \ldots, t_{n+1}) g(t_k) dG_{t_1}^{(i_1)} \ldots dG_{t_{k-1}}^{(i_{k-1})} dG_{t_k}^{(l)} dG_{t_{k+1}}^{(i_{k+1})} \ldots dG_{t_{n+1}}^{(i_n)},$$

in the sense that $Fg \in \operatorname{Dom} \delta^{(l)}$ if the infinite sum converges in $L^2(\Omega)$.

*Proof.* The first step is to derive an expression for the Skorohod integral for a process with finite chaos expansion. This can be done in an identical fashion to the general case, Lemma 2.2.5. In addition, the approximation of general random variables in $L^2(\Omega)$ with their capped sum equivalents works as well here as it did in the proof of Proposition 2.2.7. □

**Corollary 3.2.7:** The Skorohod integral is a densely defined operator in $L^2(\Omega \times \mathbb{R}^+)$.

*Proof.* Processes of the form $Fg(t)$ where $F \in L^2(\Omega)$ and $g \in L^2(\mathbb{R}^+)$ are dense in $L^2(\Omega \times \mathbb{R}^+)$. □

Skorohod integration increases the order of the chaos expansion with a $dG_l$ integral mixed in at every possible position. A point worth to be made again is that this is inline with Skorohod integration in the Gaussian case.

With Proposition 3.2.6 it is trivial to follow exactly the same steps as in the previous chapter to conclude that when previsible processes are integrated the Itô and Skorohod integral coincide.

**Proposition 3.2.8:** If $E\left[\int_0^\infty u_t^2 dt\right] < \infty$ and $u_t$ is $\mathcal{F}_{t-}$–adapted, then $u_t \in \operatorname{Dom} \delta^{(l)}$ for $l \ge 1$ and

$$\delta^{(l)}(u) = \int_0^\infty u_t dG_t^{(l)}.$$

*Remark.* When $u_t$ is a matrix–process and $W_t$ is multidimensional we have (from the definition of the Skorohod integral) the integration by parts formula:

$$E\left[\left(F\int_0^\infty u_t^* dW_t\right)^*\right] = E\left[\int_0^\infty (D_t^{(0)}F)u_t dt\right],$$

where the integration is done component–wise and $*$ denotes transpose.

Next we set out to prove Proposition 3.2.10, which provides a nice decomposition of the Skorohod integral for certain integrands. Proposition 3.2.10 is the equivalent of Property (4) Section 1.3.1 in Nualart (1995), but since we do not have a powerful enough chain rule in the present setting we have to go through a different route. We start with a proposition, which will provide us with an explicit expression for the Skorohod integral when $Fg(t)$ as above is replaced by $Fu_t$, where $u_t$ is stochastic but previsible.

Proposition 3.2.9: For $u_t$ previsible and $F \in L^2(\Omega)$ with chaos expansion

$$F = E[F] + \sum_{n=1}^{\infty} \sum_{0\le i_1,\dots,i_n\le m} L_n^{(i_1,\dots,i_n)}(f_{i_1,\dots,i_n}),$$

such that $Fu_t$ in $L^2(\Omega \times \mathbb{R}^+)$ we have

$$\begin{aligned}\delta^{(l)}(Fu) = &\int_0^\infty E[F]u_{t_1} dG_{t_1}^{(l)} + \sum_{n=1}^{\infty}\sum_{0\le i_1,\dots,i_n\le m}\sum_{k=1}^{n+1}\int_0^\infty \\ &\cdots\int_0^{t_{k-1}-}\int_0^{t_k-}\int_0^{t_{k+1}-}\cdots\int_0^{t_2-} f_{i_1,\dots,i_n}(t_1,\dots,t_{k-1},t_{k+1},\dots,t_{n+1})dG_{t_1}^{(i_1)} \\ &\dots dG_{t_{k-1}}^{(i_{k-1})}u_{t_k}dG_{t_k}^{(l)}dG_{t_{k+1}}^{(i_k)}\dots dG_{t_{n+1}}^{(i_n)}, \quad (3.1)\end{aligned}$$

in the sense that $Fu \in \mathrm{Dom}\,\delta^{(l)}$ if the right hand side is in $L^2(\Omega \times \mathbb{R}^+)$.

*Proof.* Take any $H \in \mathbb{D}^{(l)}$.

$$E\left[\int_0^\infty (D_t^{(l)}H)Fu_t dt\right] = \int_0^\infty E\Bigg[\Bigg(\sum_{n=1}^\infty \sum_{0\le i_1,\dots,i_n\le m} \sum_{k=1}^n \mathbf{1}_{\{i_k=l\}} \int_0^\infty \cdots \int_0^{t_{k+2}-} \int_0^{t_{k+1}-} \cdots \int_0^{t_2-} h_{i_1,\dots,i_n}(\dots,t,\dots)\mathbf{1}_{\{\Sigma_n^{(k)}(t)\}} dG_{t_1}^{(i_1)} \dots dG_{t_{k-1}}^{(i_{k-1})} dG_{t_{k+1}}^{(i_{k+1})} \dots dG_{t_n}^{(i_n)}\Bigg)$$
$$\times \left(E[F] + \sum_{q=1}^\infty \sum_{0\le j_1,\dots,j_q\le m} \int_0^\infty \int_0^{t_q-} \cdots \int_0^{t_2-} f_{j_1,\dots,j_q}(\dots) dG_{t_1}^{(j_1)} \dots dG_{t_{q-1}}^{(j_{q-1})} dG_{t_q}^{(j_q)}\right) u_t\Bigg] dt$$

$$= E\Bigg[\sum_{n=1}^\infty \sum_{0\le i_1,\dots,i_n\le m} \sum_{q=n-k}^\infty \sum_{0\le j_1,\dots,j_q\le m} \sum_{k=1}^n \mathbf{1}_{\{i_k=l\}} \Bigg(\int_0^\infty \cdots \int_0^{t_{k+2}-} \int_0^{t_{k+1}-} \int_0^{t-} \cdots \int_0^{t_2-} h_{i_1,\dots,i_n}(\dots,t,\dots) dG_{t_1}^{(i_1)} \dots dG_{t_{k-1}}^{(i_{k-1})} u_t dt\, dG_{t_{k+1}}^{(i_{k+1})} \dots dG_{t_n}^{(i_n)}\Bigg)$$
$$\times \left(\int_0^\infty \int_0^{t_q-} \cdots \int_0^{t_2-} f_{j_1,\dots,j_q}(\dots) dG_{t_1}^{(j_1)} \dots dG_{t_{q-1}}^{(j_{q-1})} dG_{t_q}^{(j_q)}\right)\Bigg],$$

where it is understood that the $q = 0$ term equals $E[F]$. For the $D_t^{(l)}H$ contribution, the integrals outside the $dt$ integral can be matched up with the outer integrals in the $F$–contribution with an isometry argument:

$$E\left[\int_0^\infty (D_t^{(l)}H)Fu_t dt\right]$$
$$= E\Bigg[\sum_{n=1}^\infty \sum_{0\le i_1,\dots,i_n\le m} \sum_{k=1}^n \sum_{q=n-k}^\infty \sum_{0\le j_1,\dots,j_q\le m} \mathbf{1}_{\{i_k=l\}} \mathbf{1}_{\{(i_{k+1},\dots,i_n)=(j_{q-n+k+1},\dots,j_q)\}}$$
$$\lambda_{i_{k+1}} \dots \lambda_{i_n} \int_0^\infty \int_0^{t_n-} \cdots \int_0^{t_{k+1}-} \left(\int_0^{t_k-} \cdots \int_0^{t_2-} h_{i_1,\dots,i_n}(\dots) dG_{t_1}^{(i_1)} \dots dG_{t_{k-1}}^{(i_{k-1})}\right)$$
$$\times \left(\int_0^{t_{q-n+k+1}-} \cdots \int_0^{t_2-} f_{j_1,\dots,j_q}(\dots) dG_{t_1}^{(j_1)} \dots dG_{t_{q-n+k}}^{(j_{q-n+k})}\right) u_{t_k} dt_k \dots dt_{n-1} dt_n\Bigg].$$

By the same isometry argument we can now transform back into stochastic integrals.

$$E\left[\int_0^\infty (D_t^{(l)}H)Fu_t dt\right] = E\Bigg[\sum_{n=1}^{\infty}\sum_{0\le i_1,\dots,i_n\le m}\sum_{k=1}^{n}\sum_{q=n-k}^{\infty}\sum_{0\le j_1,\dots,j_q\le m}$$
$$\left(\int_0^\infty\cdots\int_0^{t_{k+1}^-}\cdots\int_0^{t_2^-} h_{i_1,\dots,i_n}(\dots)dG_{t_1}^{(i_1)}\dots dG_{t_k}^{(l)}\dots dG_{t_n}^{(i_n)}\right)$$
$$\times\left(\int_0^\infty\cdots\int_0^{t_{q-n+k+2}^-}\cdots\int_0^{t_2^-} f_{j_1,\dots,j_q}(\dots)dG_{t_1}^{(j_1)}\dots u_{t_{q-n+k+1}}dG_{t_{q-n+k+1}}^{(l)}\dots dG_{t_{q+1}}^{(j_q)}\right)\Bigg]$$

$$= E\Bigg[\sum_{n=1}^{\infty}\sum_{0\le i_1,\dots,i_n\le m}\left(\int_0^\infty\int_0^{t_n^-}\cdots\int_0^{t_2^-} h_{i_1,\dots,i_n}(\dots)dG_{t_1}^{(i_1)}\dots dG_{t_{n-1}}^{(i_{n-1})}dG_{t_n}^{(i_n)}\right.$$
$$\sum_{k=1}^{n}\sum_{q=n-k}^{\infty}\sum_{0\le j_1,\dots,j_q\le m}\int_0^\infty\cdots\int_0^{t_{q-n+k+2}^-}$$
$$\left.\left.\cdots\int_0^{t_2^-} f_{j_1,\dots,j_q}(\dots)dG_{t_1}^{(j_1)}\dots u_{t_{q-n+k+1}}dG_{t_{q-n+k+1}}^{(l)}\dots dG_{t_{q+1}}^{(j_q)}\right)\right].$$

We can safely rewrite the last set of sums on the same form as in (3.1). Further, adding $E[H]$ to complete $H$ completes the proof. □

Proposition 3.2.10: For $Fu$ in $L^2(\Omega\times\mathbb{R}^+)$, where $F\in\mathbb{D}^{(0)}$ and $u_t$ previsible, we have

$$\delta^{(0)}(Fu) = F\int_0^\infty u_t dW_t - \int_0^\infty (D_t^{(0)}F)u_t dt.$$

in the sense that $Fu\in\operatorname{Dom}\delta^{(0)}$ if and only if the right hand side is in $L^2(\Omega)$.

*Proof.* We focus on the expression $F\int_0^\infty u_t dG_t^{(0)}$, where 0 denotes the Brownian motion. By the chaos expansion of $F$ we see that the product will be given by terms of the form

$$\int_0^\infty\int_0^{t_n^-}\cdots\int_0^{t_2^-} f_{i_1,\dots,i_n}(t_1,\dots,t_n)dG_{t_1}^{(i_1)}\dots dG_{t_{n-1}}^{(i_{n-1})}dG_{t_n}^{(i_n)}\int_0^\infty u_t dG_t^{(0)},$$

and the strategy is to prove that

$$\int_0^\infty \int_0^{t_n^-} \cdots \int_0^{t_2^-} f_{i_1,\ldots,i_n}(t_1,\ldots,t_n) dG_{t_1}^{(i_1)} \ldots dG_{t_{n-1}}^{(i_{n-1})} dG_{t_n}^{(i_n)} \int_0^\infty u_t dG_t^{(0)}$$

$$= \sum_{k=1}^{n+1} \int_0^\infty \cdots \int_0^{t_{k+2}^-} \int_0^{t_{k+1}^-} \int_0^{t_k^-} \cdots \int_0^{t_2^-} f_{i_1,\ldots,i_n}(t_1,\ldots,t_{k-1},t_{k+1},\ldots,t_{n+1})$$

$$dG_{t_1}^{(i_1)} \ldots dG_{t_{k-1}}^{(i_{k-1})} u_{t_k} dG_{t_k}^{(0)} dG_{t_{k+1}}^{(i_k)} \ldots dG_{t_{n+1}}^{(i_n)}$$

$$+ \sum_{k=1}^{n} \int_0^\infty \cdots \int_0^{t_k^-} \cdots \int_0^{t_2^-} f_{i_1,\ldots,i_n}(t_1,\ldots,t_n) dG_{t_1}^{(i_1)} \ldots u_{t_k} d[G^{(i_k)}, G^{(0)}]_{t_k} \ldots dG_{t_n}^{(i_n)}. \quad (3.2)$$

The first sum of the right hand side of (3.2) is the contribution from the chaos expansion of $\delta^{(l)}(Fu)$ and the second sum is the contribution from the chaos expansion of $\int_0^\infty (D_t^{(0)} F) u_t dt$. We prove (3.2) by induction on $n$. For $n = 1$ we write

$$Z_t = \int_0^t f_{i_1}(t_1) dG_{t_1}^{(i_1)}, \quad Y_t = \int_0^t u_{t_1} dG_{t_1}^{(0)}.$$

By Itô's formula[1] we have

$$d(Z_t Y_t) = Y_{t-} dZ_t + Z_{t-} dY_t + d[Z^{cm}, Y^{cm}]_t,$$

So,

$$d(Z_t Y_t) = \int_0^{t-} u_{t_1} dG_{t_1}^{(0)} f_{i_1}(t) dG_t^{(i_1)} + \int_0^{t-} f_{i_1}(t_1) dG_{t_1}^{(i_1)} u_t dG_t^{(0)} + f_{i_1}(t) u_t d[G^{(i_1)}, G^{(0)}]_t,$$

and the claim holds true for $n = 1$. Now, assume that the claim holds for $n - 1$ and consider

$$Z_t = \int_0^t \int_0^{t_n^-} \cdots \int_0^{t_2^-} f_{i_1,\ldots,i_n}(t_1,\ldots,t_n) dG_{t_1}^{(i_1)} \ldots dG_{t_{n-1}}^{(i_{n-1})} dG_{t_n}^{(i_n)}$$

$$Y_t = \int_0^t u_{t_1} dG_{t_1}^{(0)}.$$

[1] See for example Rogers and Williams (2000) section VI39.

By Itô's formula again

$$d(Z_tY_t) = \left(\int_0^{t-} u_{t_1} dG_{t_1}^{(0)} \int_0^{t-}\int_0^{t_{n-1}-}\cdots\int_0^{t_1-} f_{i_1,\ldots,i_n}(t_1,\ldots,t_{n-1},t)\right.$$
$$\left. dG_{t_1}^{(i_1)}\ldots dG_{t_{n-2}}^{(i_{n-2})} dG_{t_{n-1}}^{(i_{n-1})}\right) dG_t^{(i_n)}$$
$$+\left(\int_0^{t-}\int_0^{t_n^-}\cdots\int_0^{t_2^-} f_{i_1,\ldots,i_n}(t_1,\ldots,t_{n-1},t) dG_{t_1}^{(i_1)}\ldots dG_{t_{n-1}}^{(i_{n-1})} dG_{t_n}^{(i_n)}\right) u_t dG_t^{(0)}$$
$$+\int_0^{t-}\int_0^{t_{n-1}^-}\cdots\int_0^{t_2^-} f_{i_1,\ldots,i_n}(t_1,\ldots,t_{n-1},t) dG_{t_1}^{(i_1)}\ldots dG_{t_{n-2}}^{(i_{n-2})} dG_{t_{n-1}}^{(i_{n-1})} u_t d[G^{(i_n)},G^{(0)}]_t,$$

and by substituting the expression in brackets on the first row by the induction assumption, the result follows. □

*Remark.* In the case when $F$ is a $d$-dimensional random column–vector and $u_t$ is a $(d \times d)$ matrix–process Proposition 3.2.10 translates to

$$\delta^{(0)}(uF) = F^* \int_0^\infty u_t dW_t - \int_0^\infty \mathrm{Tr}\left((D_t^{(0)}F)u_t\right) dt,$$

with the convention that the Itô integral for a matrix process is a column–vector.

## 3.3 Separable jump diffusions

We started out in the previous chapter by briefly considering general Lévy processes. We have since constantly narrowed down the scope by considering Lévy processes satisfying Assumption 2.1.1 and then simple Lévy processes. Now we will make the final specialisation and consider a class of jump diffusions for which we can use the chain rule Theorem 3.2.4.

Definition 3.3.1 (Separability): Let $X_t \in \bigcap_{j=1}^{d+m} \mathbb{D}^{(j)}$ be a Markov jump diffusion , of the form

$$dX_t = b(X_{t-})dt + \sigma(X_{t-})dW_t + \sum_{k=1}^m \alpha_k(X_{t-})(dN_t^{(k)} - \lambda_k dt), \quad X_0 = x,$$

for which the we have a continuously differentiable function $f$ with bounded derivative in the first argument such that

$$X_t = f(X_t^c, X_t^d), \quad X_0^c = x,$$

where $X_t^c$ only is driven by the Brownian motion and $X_t^d$ is not dependent on $x$ and only driven by the Poisson processes. We say that the jump diffusion process $X_t$ is separable.

Note that in the definition of separability there are no conditions imposed on the drift, diffusion and jump coefficients. Relevant conditions will be discussed when the definition is used. It is obvious that $X_t^c$ is a Markov diffusion and we write

$$dX_t^c = b_c(X_t^c)dt + \sigma_c(X_t^c)dW_t, \quad X_0^c = x. \tag{3.3}$$

An important property of the class of separable jump diffusions is that they are differentiable with respect to Brownian motion if their first variation process (defined below) is well defined. The Malliavin derivative can then be calculated using the chain rule given in Theorem 3.2.4 . This, together with the notion of Skorohod integration, allows us to derive stochastic weights for calculating the Greeks using Monte Carlo simulation by more or less mimicking the work of Fournié et al. (1999). More on this in the next chapter.

The separability condition might seem restrictive at first, but it can be shown that many interesting processes in Mathematical Finance are indeed separable. To that end consider the stochastic volatility model:

$$\begin{aligned} dX_t^{(1)} &= b_1\left(X_t^{(2)},\ldots,X_t^{(d)}\right)X_{t-}^{(1)}dt + \sigma_1\left(X_t^{(2)},\ldots,X_t^{(d)}\right)X_{t-}^{(1)}dW_t \\ &\quad + \textstyle\sum_{k=1}^m(\alpha_k-1)X_{t-}^{(1)}(dN_t^{(k)}-\lambda_k dt), \quad X_0^{(1)} = x_1 \\ dX_t^{(i)} &= b_i\left(X_t^{(2)},\ldots,X_t^{(d)}\right)dt + \sigma_i\left(X_t^{(2)},\ldots,X_t^{(d)}\right)dW_t, \\ &\quad X_0^{(i)} = x_i, \quad i = 2,\ldots,d, \end{aligned} \tag{3.4}$$

where $W$ is $d$–dimensional Brownian motion, $b_i$ as well as $\sigma_i$ have bounded continuous derivatives for $i = 1,\ldots,d$ and $\alpha_k$, $k = 1,\ldots,m$ are deterministic constants. We introduce the $d$-dimensional continuous process $X_t^c$ defined by

$$\begin{aligned} dX_t^{c\,(1)} &= \left(b_1\left(X_t^{c\,(2)},\ldots,X_t^{c\,(d)}\right) + \textstyle\sum_{k=1}^m\lambda_k(1-\alpha_k)\right)X_t^{c\,(1)}dt \\ &\quad +\sigma_1\left(X_t^{c\,(2)},\ldots,X_t^{c\,(d)}\right)X_{t-}^{c\,(1)}dW_t \quad X_0^{c\,(1)} = x_1 \\ dX_t^{c\,(i)} &= b_i\left(X_t^{c\,(2)},\ldots,X_t^{c\,(d)}\right)dt + \sigma_i\left(X_t^{c\,(2)},\ldots,X_t^{c\,(d)}\right)dW_t, \\ &\quad X_0^{c\,(i)} = x_i, \quad i = 2,\ldots,d, \end{aligned} \tag{3.5}$$

and the $d$-dimensional discontinuous process

$$\begin{aligned} \hat{X}_t^{(1)} &= \alpha_1^{N_t^{(1)}} \dots \alpha_m^{N_t^{(m)}} X_t^{c\,(1)} \\ \hat{X}_t^{(i)} &= X_t^{c\,(i)}, \quad i = 2, \dots, d. \end{aligned} \tag{3.6}$$

It is clear that $(X^{(2)}, \dots, X^{(d)})$ and $(\hat{X}^{(2)}, \dots, \hat{X}^{(d)})$ are indistinguishable due to pathwise uniqueness of the solutions to the SDEs, and by applying the Itô formula to $\hat{X}_t^{(1)}$ we get

$$\begin{aligned} d\hat{X}_t^{(1)} &= \frac{\partial \hat{X}^{(1)}}{\partial X^{c\,(1)}}(t-) dX_t^{c\,(1)} + \frac{1}{2}\frac{\partial^2 \hat{X}^{(1)}}{\partial X^{c\,(1)^2}}(t-) d[X^{c\,(1)}]_t + \Delta \hat{X}_t^{(1)} \\ &= \alpha_1^{N_{t-}^{(1)}} \dots \alpha_m^{N_{t-}^{(m)}} \left[ \left( b_1 \left( X_t^{c\,(2)}, \dots, X_t^{c\,(d)} \right) + \sum_{k=1}^{m} \lambda_k (1 - \alpha_k) \right) X_t^{c\,(1)} dt \right. \\ &\quad \left. + \sigma_1 \left( X_t^{c\,(2)}, \dots, X_t^{c\,(d)} \right) X_t^{c\,(1)} dW_t^{(1)} \right] \\ &\quad + \sum_{k=1}^{m} (\alpha_k - 1) \alpha_1^{N_{t-}^{(1)}} \dots \alpha_m^{N_{t-}^{(m)}} X_t^{c\,(1)} dN_t^{(k)} \\ &= \left( b_1 \left( \hat{X}_t^{(2)}, \dots, \hat{X}_t^{(d)} \right) + \sum_{k=1}^{m} \lambda_k (1 - \alpha_k) \right) \hat{X}_{t-}^{(1)} dt \\ &\quad + \sigma_1 \left( \hat{X}_t^{(2)}, \dots, \hat{X}_t^{(d)} \right) \hat{X}_{t-}^{(1)} dW_t^{(1)} + \sum_{k=1}^{m} (\alpha_k - 1) \hat{X}_{t-}^{(1)} dN_t^{(k)}, \end{aligned}$$

so that $\hat{X}_t^{(1)} = X_t^{(1)}$.

We summarise the result in the following Lemma.

**Lemma 3.3.2:** Let $X_t$, $X_t^c$ and $\hat{X}_t$ be as defined in (3.4), (3.5) and (3.6) respectively. Then $\hat{X}_t = X_t$ a.s.

For the diffusion process $X_t^c$ in (3.3) we define the *first variation process* as

$$dY_t = b_c'(X_t^c) Y_t dt + \sum_{i=1}^{d} \sigma_{ci}'(X_t^c) Y_t dW_t^{(i)}, \quad Y_0 = I,$$

where $I$ is the identity matrix, $\sigma_{ci}$ is the $i$-th column vector of $\sigma_c$ and prime denotes derivatives. It is true that

$$Y_t = \nabla_x X_t^c.$$

If $b_c$ and $\sigma_c$ have continuous bounded derivatives the first variation process is well defined, and then the Malliavin derivative of $X_t^c$ can be written as[2]

$$D_s^{(0)} X_t^c = Y_t Y_s^{-1} \sigma_c(X_s^c) \mathbf{1}_{\{s \le t\}}.$$

[2] See Nualart (1995) section 2.3.1.

Note that boundedness of $b_c$ and $\sigma_c$ is not a necessary condition for the existence of the first variation process. In deed, the stochastic volatility setup (3.4) also has a well defined first variation process since $X_t^{(1)}$ can be seen as a stand–alone linear stochastic differential equation with random coefficients.

# Part II

# STOCHASTIC WEIGHTS FOR DERIVATIVE SENSITIVITIES

# 4. MALLIAVIN MONTE CARLO GREEKS FOR SEPARABLE JUMP DIFFUSIONS

## 4.1 Sensitivity weights

In this chapter we will apply the theory of the previous chapters to the problem of Monte Carlo convergence of derivative sensitivities. The treatment follows that of Fournié et al. (1999), but in our more general setting of separable jump diffusions.

On a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ we consider a separable price process of the form

$$dX_t = b(X_{t-})dt + \sigma(X_{t-})dW_t + \sum_{k=1}^{m} \alpha_k(X_{t-})(dN_t^{(k)} - \lambda_k dt), \quad X_0 = x.$$

As before the continuous part of $X_t$ will be denoted by $X_t^c$ and we write

$$dX_t^c = b_c(X_t^c)dt + \sigma_c(X_t^c)dW_t, \quad X_0^c = x.$$

An important role for the derivation of the stochastic weights will be played by the first variation process for $X_t^c$ defined by

$$dY_t = b_c'(X_t^c)Y_t dt + \sum_{i=1}^{d} \sigma_{ci}'(X_t^c)Y_t dW_t^{(i)}, \quad Y_0 = I,$$

with notation as in Section 3.3. A sufficient (but not necessary) condition for existence of $Y_t$ is that $b_c$ and $\sigma_c$ have bounded continuous derivatives.[1] It is true that

$$Y_t = \nabla_x X_t^c,$$

and it is also known that the Malliavin derivative of $X_t^c$ can be written as[2]

$$D_s^{(0)} X_t^c = Y_t Y_s^{-1} \sigma_c(X_s^c) \mathbf{1}_{\{s \le t\}}. \tag{4.1}$$

[1] See Fournié et al. (1999).

[2] See Nualart (1995) section 2.3.1.

We define the payoff function

$$\phi = \phi(X_{t_1}, \ldots, X_{t_n}), \tag{4.2}$$

to be a square integrable function discounted from maturity $T$ and evaluated at the times $t_1, \ldots, t_n$. The price of a contingent claim is then expressed as

$$v(x) = E\left[\phi(X_{t_1}, \ldots, X_{t_n})\right].$$

The objective is to differentiate $v$ with respect to the model parameters and for the proofs we will need to assume the diffusion matrix to be *uniformly elliptic*, that is

$$\exists \eta > 0, \quad \xi^* \sigma^*(x)\sigma(x)\xi \geq \eta|\xi|^2 \quad \text{for any } \xi,\ x \in \mathbb{R}^n.$$

We will also need the following technical lemma:

Lemma 4.1.1: Let $g : \mathbb{R} \to \mathbb{R}$ be a continuously differentiable function with bounded derivative and let $G$ be a random variable in $\bigcap_{j=1}^{d+m} \mathbb{D}^{(j)}$ satisfying the conditions for the chain rule (Theorem 3.2.4). Then

$$g(G) \in \bigcap_{j=1}^{d+m} \mathbb{D}^{(j)}.$$

*Proof.* The chain rule (Theorem 3.2.4) implies that $g(G) \in \bigcap_{j=1}^{d} \mathbb{D}^{(j)}$. For $k \in [d+1, d+m]$ we can write the Malliavin derivative as[3]

$$D_t^{(k)} G = G(\omega + \theta_t^{(k)}) - G(\omega),$$

where $\omega + \theta_t^{(k)}$ means 'an extra jump in the $k$th Poisson process at time $t$'. We have,

$$\begin{aligned} E\left[\int_0^\infty \left(D_t^{(k)} g(G)\right)^2 dt\right] &= E\left[\int_0^\infty \left(g(G(\omega + \theta_t^{(k)})) - g(G(\omega))\right)^2 dt\right] \\ &\leq E\left[M^2 \int_0^\infty \left(G(\omega + \theta_t^{(k)}) - G(\omega)\right)^2 dt\right] < \infty, \end{aligned}$$

where $M$ is a bound on the derivative of $g$. □

[3] See León et al. (2002) Proposition 2.4(b).

### 4.1.1 Variations in the drift coefficient

In order to assess the sensitivity of the price of the contingent claim $v$ to changes in the drift coefficient we introduce the perturbed process

$$\begin{aligned} dX_t^\epsilon &= (b(X_{t-}^\epsilon) + \epsilon\gamma(X_{t-}^\epsilon))\,dt + \sigma(X_{t-}^\epsilon)dW_t + \sum_{k=1}^{m} \alpha_k(X_{t-}^\epsilon)(dN_t^{(k)} - \lambda_k dt) \\ X_0^\epsilon &= x, \end{aligned}$$

where $\epsilon$ is a scalar and $\gamma$ is a bounded function. The following Proposition tells us how sensitive the price of a claim on the perturbed process is to $\epsilon$ in the point $\epsilon = 0$.

Proposition 4.1.2: Suppose that the diffusion matrix $\sigma$ is uniformly elliptic. For $v^\epsilon(x)$ defined as

$$v^\epsilon(x) = E\left[\phi(X_T^\epsilon)\right],$$

we have

$$\frac{\partial}{\partial\epsilon} v^\epsilon(x)\Big|_{\epsilon=0} = E\left[\phi(X_T)\int_0^T (\sigma^{-1}(X_{t-})\gamma(X_{t-}))^* dW_t\right]$$

*Proof.* The proof builds on an application of the Girsanov Theorem which holds true even in the presence of Poisson jumps. See e.g. the proof of Theorem E2 in Karatzas and Shreve (1998). □

*Remark.* The result holds true even for path-dependent claims $\phi = \phi(X(.))$ and also for processes which are not separable as in Definition 3.3.1.

### 4.1.2 Variations in the initial condition

This is where we will make full use of the theory developed in Chapter 3 for the first time. We rely heavily on the chain rule and the Skorohod integral, which allow us to repeat the steps in Fournié et al. (1999).

First define the set $\Gamma$, of square integrable functions taking values in $\mathbb{R}$, as

$$\Gamma = \left\{ a \in L^2([0,T)) : \int_0^{t_i} a(t)dt = 1, \quad \forall i = 1,\ldots,n \right\},$$

where $t_i$ are as defined in (4.2).

**Proposition 4.1.3:** Assume that the diffusion matrix $\sigma_c$ is uniformly elliptic. Then for any $a(t) \in \Gamma$

$$(\nabla v(x))^* = E\left[\phi(X_{t_1}, \ldots, X_{t_n}) \int_0^T a(t) \left(\sigma_c^{-1}(X_t^c) Y_t\right)^* dW_t\right].$$

*Proof.* First assume that $\phi$ is continuously differentiable with bounded gradient. In this case it is possible to 'differentiate inside the expectation'.[4] We will denote by $\nabla_i \phi(X_{t_1}, \ldots, X_{t_n})$ the gradient of $\phi$ with respect to $X_{t_i}$, and by $\frac{\partial X_{t_i}}{\partial x}$ the $(d \times d)$ matrix of derivatives of the $d$–dimensional random variable $X_{t_i}$ with respect to its initial condition.

Remember that from the separability assumption $X_T^d$ does not depend on $x$ so that

$$\begin{aligned} \nabla v(x) &= E\left[\sum_{i=1}^n \nabla_i \phi(X_{t_1}, \ldots, X_{t_n}) \frac{\partial X_{t_i}}{\partial x}\right] \\ &= E\left[\sum_{i=1}^n \nabla_i \phi(X_{t_1}, \ldots, X_{t_n}) \frac{\partial X_{t_i}}{\partial X_{t_i}^c} Y_{t_i}\right]. \end{aligned} \tag{4.3}$$

We want to rewrite $Y_{t_i}$ in terms of the Malliavin derivative for $X_{t_i}$ using (4.1), but to do that we must also use the chain rule (Theorem 3.2.4). We have

$$D_t^{(0)} X_{t_i} = \frac{\partial X_{t_i}}{\partial X_{t_i}^c} D_t^{(0)} X_{t_i}^c = \frac{\partial X_{t_i}}{\partial X_{t_i}^c} Y_{t_i} Y_t^{-1} \sigma_c(X_t^c) \mathbf{1}_{\{t \le t_i\}}.$$

For any $a(t) \in \Gamma$ we can express $\frac{\partial X_{t_i}}{\partial X_{t_i}^c} Y_{t_i}$ as

$$\frac{\partial X_{t_i}}{\partial X_{t_i}^c} Y_{t_i} = \int_0^T a(t) (D_t^{(0)} X_{t_i}) \sigma_c^{-1}(X_t^c) Y_t dt.$$

Inserting in (4.3) yields

$$\nabla v(x) = E\left[\int_0^T \sum_{i=1}^n \nabla_i \phi(X_{t_1}, \ldots, X_{t_n}) (D_t^{(0)} X_{t_i}) a(t) \sigma_c^{-1}(X_t^c) Y_t dt\right],$$

We know from Lemma 4.1.1 that $\phi(X_{t_1}, \ldots, X_{t_n}) \in \bigcap_{j=1}^{d+m} \mathbb{D}^{(j)}$ since $\phi$ is continuously differentiable with bounded gradient and $X_t \in \bigcap_{j=1}^{d+m} \mathbb{D}^{(j)}$ is separable. We

[4] The proof of this claim is in Fournié et al. (1999) built on the fact that almost surely convergence together with uniform integrability implies convergence in $L^1$ norm. The almost surely convergence holds true in the present treatment exactly as stated in the continuous case. The uniform integrability follows from the the boundedness of $\frac{\partial X_T}{\partial X_T^c}$ and the fact that the map $x \mapsto X_t$ is a.s. continuous even in the presence of jumps.

can thus use the chain rule again to get

$$\nabla v(x) = E\left[\int_0^T \left(D_t^{(0)}\phi(X_{t_1},\ldots,X_{t_n})\right) a(t)\sigma_c^{-1}(X_t^c)Y_t dt\right].$$

Since the diffusion matrix $\sigma_c$ is uniformly elliptic by assumption we can deduce that $a(t)\sigma_c^{-1}(X_t^c)Y_t \in L^2(\Omega \times [0,T])$. We can therefore use the integration by parts formula implied by the definition of the Skorohod integral and Proposition 3.2.8 to get

$$(\nabla v(x))^* = E\left[\phi(X_{t_1},\ldots,X_{t_n})\int_0^T a(t)\left(\sigma_c^{-1}(X_t^c)Y_t\right)^* dW_t\right].$$

The fact that the class of continuously differentiable functions with bounded gradient is dense in $L^2$ can be used, exactly as in Fournié et al. (1999), to prove the general result for $\phi \in L^2$. □

*Remark.* The above argument can easily be repeated to get derivatives of higher order.

*Remark.* Proposition 4.1.3 is known as the Bismut–Elworthy formula. See Elworthy and Li (1994) for an alternative proof in the continuous diffusion case.

We see that the discontinuities do not appear in the stochastic weight. However, the fact that the payoff function is evaluated for the full price process ensures that the sensitivity commonly known as *delta* does depend on the jump parameters.

### 4.1.3 Variations in the diffusion coefficient

Calculating the stochastic weight for what is commonly known as *vega* is a bit more involved than the previous Greeks. It is here the need for a Skorohod integral arises and we use Proposition 3.2.10 to interpret the result.

As in Section 4.1.1 we need to define the perturbed process with respect to the property under investigation; in this case the diffusion coefficient:

$$\begin{aligned} dX_t^\epsilon &= b(X_{t-}^\epsilon)dt + (\sigma(X_{t-}^\epsilon) + \epsilon\gamma(X_{t-}^\epsilon))\,dW_t + \sum_{k=1}^m \alpha_k(X_{t-}^\epsilon)(dN_t^{(k)} - \lambda_k dt) \\ X_0^\epsilon &= x, \end{aligned}$$

where again $\epsilon$ is a scalar and $\gamma$ is a continuously differentiable function with bounded derivative. We will also need to introduce the variation process with respect to the parameter $\epsilon$

$$\begin{aligned} dZ_t^\epsilon &= b'(X_{t-}^\epsilon)Z_{t-}^\epsilon dt + \sum_{i=1}^{d} \left(\sigma_i'(X_{t-}^\epsilon) + \epsilon\gamma_i'(X_{t-}^\epsilon)\right) Z_{t-}^\epsilon dW_t^{(i)} + \gamma(X_{t-}^\epsilon)dW_t \\ &\quad + \sum_{k=1}^{m} \alpha_k'(X_{t-}^\epsilon)Z_{t-}^\epsilon(dN_t^{(k)} - \lambda_k dt), \quad Z_0^\epsilon = 0_n, \end{aligned}$$

so that $\frac{\partial X_t^\epsilon}{\partial \epsilon} = Z_t^\epsilon$. Further, we define the set $\Gamma_n$, of square integrable functions taking values in $\mathbb{R}$, as

$$\Gamma_n = \left\{ a \in L^2([0,T)) : \int_{t_{i-1}}^{t_i} a(t)dt = 1, \quad \forall i = 1,\ldots,n \right\}.$$

In this setting Proposition 4.1.4 is the analog of Proposition 4.1.2.

Proposition 4.1.4: Assume that the diffusion matrix $\sigma_c$ is uniformly elliptic and that for $\beta_{t_i} = \left(\frac{\partial X_{t_i}}{\partial X_{t_i}^c} Y_{t_i}\right)^{-1} Z_{t_i}$, $i = 1,\ldots,n$ we have $\sigma_c^{-1}(X^c)Y\beta \in \operatorname{Dom}\delta^{(0)}$. For $v^\epsilon(x)$ defined as

$$v^\epsilon(x) = E\left[\phi(X_{t_1}^\epsilon,\ldots,X_{t_n}^\epsilon)\right],$$

we have for any $a \in \Gamma_n$

$$\frac{\partial}{\partial\epsilon} v^\epsilon(x)\Big|_{\epsilon=0} = E\left[\phi(X_{t_1},\ldots,X_{t_n})\delta^{(0)}\left(\sigma_c^{-1}(X_\cdot^c)Y_\cdot\tilde{\beta}_\cdot\right)\right],$$

where

$$\tilde{\beta}_t = \sum_{i=1}^{n} a(t)\left(\beta_{t_i} - \beta_{t_{i-1}}\right)\mathbf{1}_{\{t\in[t_{i-1},t_i)\}},$$

for $t_0 = 0$. Furthermore, if $\beta \in \mathbb{D}^{(0)}$ the Skorohod integral can be calculated according to

$$\begin{aligned} \delta^{(0)}\left(\sigma_c^{-1}(X_\cdot^c)Y_\cdot\tilde{\beta}_\cdot\right) &= \sum_{i=1}^{n}\Bigg\{\beta_{t_i}^* \int_{t_{i-1}}^{t_i} a(t)\left(\sigma_c^{-1}(X_t^c)Y_t\right)^* dW_t \\ &\quad - \int_{t_{i-1}}^{t_i} a(t)\mathrm{Tr}\left((D_t^{(0)}\beta_{t_i})\sigma_c^{-1}(X_t^c)Y_t\right)dt \\ &\quad - \int_{t_{i-1}}^{t_i} a(t)\left(\sigma_c^{-1}(X_t^c)Y_t\beta_{t_{i-1}}\right)^* dW_t\Bigg\}. \end{aligned}$$

*Proof.* As in the proof of Proposition 4.1.3 we may assume $\phi$ to be continuously differentiable with bounded gradient, and we can take the differential under the expectation to get

$$\frac{\partial}{\partial \epsilon} v^{\epsilon}(x) = E\left[\sum_{i=1}^{n} \nabla_i \phi(X^{\epsilon}_{t_1}, \ldots, X^{\epsilon}_{t_n}) Z^{\epsilon}_{t_i}\right]. \tag{4.4}$$

Next we define $\beta_t = \left(\frac{\partial X_t}{\partial X^c_t} Y_t\right)^{-1} Z_t = \left(\frac{\partial X_t}{\partial X^c_t} Y_t\right)^{-1} Z^{\epsilon=0}_t$, and proceed as in the proof of Proposition 4.1.3. Because of (4.1) and since $a \in \Gamma_n$ we can write

$$\begin{aligned}
\int_0^T \frac{\partial X_{t_i}}{\partial X^c_{t_i}} (D^{(0)}_t X^c_{t_i}) \sigma_c^{-1}(X^c_t) Y_t \tilde{\beta}_t dt &= \int_0^{t_i} \frac{\partial X_{t_i}}{\partial X^c_{t_i}} Y_{t_i} \tilde{\beta}_t dt \\
&= \frac{\partial X_{t_i}}{\partial X^c_{t_i}} Y_{t_i} \sum_{j=1}^{i} \int_{t_{j-1}}^{t_j} a(t) \left(\beta_{t_j} - \beta_{t_{j-1}}\right) dt \\
&= \frac{\partial X_{t_i}}{\partial X^c_{t_i}} Y_{t_i} \beta_{t_i} = Z_{t_i}.
\end{aligned}$$

Inserting into (4.4) we get, by using the chain rule and Lemma 4.1.1,

$$\begin{aligned}
\frac{\partial}{\partial \epsilon} v^{\epsilon}(x)\Big|_{\epsilon=0} &= E\left[\int_0^{T} \sum_{i=1}^{n} \nabla_i \phi(X^{\epsilon}_{t_1}, \ldots, X^{\epsilon}_{t_n}) \frac{\partial X_{t_i}}{\partial X^c_{t_i}} (D^{(0)}_t X^c_{t_i}) \sigma_c^{-1}(X^c_t) Y_t \tilde{\beta}_t dt\right] \\
&= E\left[\int_0^T (D^{(0)}_t \phi(X^{\epsilon}_{t_1}, \ldots, X^{\epsilon}_{t_n})) \sigma_c^{-1}(X^c_t) Y_t \tilde{\beta}_t dt\right].
\end{aligned}$$

By assumption and the linearity of the Skorohod integral $\sigma_c^{-1}(X^c) Y \tilde{\beta} \in \operatorname{Dom} \delta^{(0)}$. So,

$$\frac{\partial}{\partial \epsilon} v^{\epsilon}(x)\Big|_{\epsilon=0} = E\left[\phi(X^{\epsilon}_{t_1}, \ldots, X^{\epsilon}_{t_n})\, \delta^{(0)}\left(\sigma_c^{-1}(X^c_{\cdot}) Y_{\cdot} \tilde{\beta}_{\cdot}\right)\right],$$

and Proposition 3.2.10 can be used to calculate the Skorohod integral. □

*Remark.* The assumption of $\sigma_c^{-1}(X^c) Y \beta \in \operatorname{Dom} \delta^{(0)}$ might seem restrictive at first, but it can be shown that important examples do satisfy the assumption. See the Appendix for the explicit calculations in the case of the Heston stochastic volatility model.

### 4.1.4 Variations in the jump intensity

The stochastic weight for variations in jump intensity is derived using the same technique as for variations in drift coefficient and it is given by the following result.

**Proposition 4.1.5:** Suppose that the diffusion matrix $\sigma$ is uniformly elliptic. For $j = 1, \ldots, m$ we have

$$\frac{\partial v}{\partial \lambda_j} = E\left[\phi(X_T)\left(\frac{N_T^{(j)}}{\lambda_j} - T - \int_0^T (\sigma^{-1}(X_{t-})\alpha_j(X_{t-}))^* dW_t\right)\right].$$

*Proof.* The argument will be carried out for the first of the $m$ Poisson processes to simplify the notation. Consider the perturbed process

$$\begin{aligned} dX_t^\epsilon &= b(X_{t-}^\epsilon)dt + \sigma(X_{t-}^\epsilon)dW_t + \alpha_1(X_{t-}^\epsilon)(dN_t^\epsilon - (\lambda_1 + \epsilon)dt) \\ &\quad + \sum_{k=2}^m \alpha_k(X_{t-}^\epsilon)(dN_t^{(k)} - \lambda_k dt), \quad X_0^\epsilon = x, \end{aligned}$$

where $\epsilon$ is a positive deterministic parameter and $N_t^\epsilon$ is a Poisson process with intensity $\lambda_1 + \epsilon$. As in the proof of Proposition 4.1.2 we note that by changing measure we can write

$$E\left[\phi(X_T^\epsilon)\right] = E\left[M_T^\epsilon \phi(X_T)\right],$$

where $M_t^\epsilon = M_t^W M_t^N$ is the Radon-Nikodym derivative for which the altered drift due to the perturbation is controlled for by $M_t^W$ and the increased jump intensity is governed by $M_t^N$. Explicitly,

$$\begin{aligned} M_t^W &= \exp\left(-\epsilon \int_0^T (\sigma^{-1}(X_{t-})\alpha_1(X_{t-}))^* dW_t - \frac{\epsilon^2}{2}\int_0^T \|\sigma^{-1}(X_{t-})\alpha_1(X_{t-})\|^2 ds\right) \\ M_t^N &= \left(1 + \frac{\epsilon}{\lambda_1}\right)^{N_t^{(1)}} \exp(-\epsilon t). \end{aligned}$$

Note that we can write

$$M_T^\epsilon = 1 - \epsilon \int_0^T M_t^W (\sigma^{-1}(X_{t-})\alpha_1(X_{t-}))^* dW_t + \frac{\epsilon}{\lambda_1}\int_0^T M_t^N (dN_t^{(1)} - \lambda_1 dt),$$

which implies that

$$\begin{aligned} \frac{\partial v}{\partial \lambda_j} &= \lim_{\epsilon \to 0} E\left[\frac{\phi(X_T^\epsilon) - \phi(X_T)}{\epsilon}\right] = \lim_{\epsilon \to 0} E\left[\phi(X_T)\left(\frac{M_T^\epsilon - 1}{\epsilon}\right)\right] \\ &= E\left[\phi(X_T)\left(\int_0^T \frac{1}{\lambda_1}(dN_t^{(1)} - \lambda_1 dt) - \int_0^T (\sigma^{-1}(X_{t-})\alpha_1(X_{t-}))^* dW_t\right)\right]. \end{aligned}$$

□

*Remark.* As was the case for variations in the drift coefficient the above result holds true even for path–dependent claims and non–separable price dynamics.

### 4.1.5 Variations in the jump amplitude

To derive a stochastic weight for the sensitivity to the amplitude parameter $\alpha$ we adopt the same technique as in the proof of Proposition 4.1.4—the diffusion coefficient result. To that end, consider the perturbed process

$$\begin{aligned} dX_t^\epsilon &= b(X_{t-}^\epsilon)dt + \sigma(X_{t-}^\epsilon)dW_t + (\alpha_1(X_{t-}^\epsilon) + \epsilon\gamma(X_{t-}^\epsilon))(dN_t^{(1)} - \lambda_1 dt) \\ &\quad + \sum_{k=2}^{m} \alpha_k(X_{t-}^\epsilon)(dN_t^{(k)} - \lambda_k dt), \quad X_0^\epsilon = x, \end{aligned}$$

where $\epsilon$ is a deterministic parameter and $\gamma$ is a continuously differentiable function with bounded derivative. Again, for notational simplicity, we consider variations in the first of the $m$ jump amplitudes. The variation process with respect to the parameter $\epsilon$ becomes

$$\begin{aligned} dZ_t^\epsilon &= b'(X_{t-}^\epsilon)Z_{t-}^\epsilon dt + \sum_{i=1}^{d} \sigma_i'(X_{t-}^\epsilon)Z_{t-}^\epsilon dW_t^{(i)} \\ &\quad + (\alpha_1'(X_{t-}^\epsilon) + \epsilon\gamma'(X_{t-}^\epsilon))Z_{t-}^\epsilon(dN_t^{(1)} - \lambda_1 dt) \\ &\quad + \sum_{k=2}^{m} \alpha_k'(X_{t-}^\epsilon)Z_{t-}^\epsilon(dN_t^{(k)} - \lambda_k dt) + \gamma(X_{t-}^\epsilon)(dN_t^{(1)} - \lambda_1 dt), \; Z_0^\epsilon = 0. \end{aligned}$$

The stochastic weight in this context is obtained similarly to the vega weight. The statement of the following proposition is therefore almost identical to Proposition 4.1.4.

Proposition 4.1.6: Assume that the diffusion matrix $\sigma_c$ is uniformly elliptic and that for $\beta_{t_i} = \left(\frac{\partial X_{t_i}}{\partial X_{t_i}^c} Y_{t_i}\right)^{-1} Z_{t_i}$, $i = 1, \ldots, n$ we have $\sigma_c^{-1}(X^c)Y\beta \in \operatorname{Dom}\delta^{(0)}$. For $v^\epsilon(x)$ defined as

$$v^\epsilon(x) = E\left[\phi(X_{t_1}^\epsilon, \ldots, X_{t_n}^\epsilon)\right],$$

we have for any $a \in \Gamma_n$

$$\frac{\partial}{\partial\epsilon} v^\epsilon(x)\Big|_{\epsilon=0} = E\left[\phi(X_{t_1}, \ldots, X_{t_n})\delta^{(0)}\left(\sigma_c^{-1}(X_\cdot^c)Y_\cdot\tilde{\beta}_\cdot\right)\right],$$

where

$$\tilde{\beta}_t = \sum_{i=1}^{n} a(t)\left(\beta_{t_i} - \beta_{t_{i-1}}\right)\mathbf{1}_{\{t\in[t_{i-1},t_i)\}},$$

for $t_0 = 0$. Furthermore, if $\beta \in \mathbb{D}^{(0)}$ the Skorohod integral can be calculated according to

$$\begin{aligned}\delta^{(0)}\left(\sigma_c^{-1}(X_\cdot^c)Y_\cdot\tilde{\beta}_\cdot\right) &= \sum_{i=1}^{n}\Bigg\{\beta_{t_i}^*\int_{t_{i-1}}^{t_i}a(t)\left(\sigma_c^{-1}(X_t^c)Y_t\right)^* dW_t \\ &\quad - \int_{t_{i-1}}^{t_i}a(t)\mathrm{Tr}\left((D_t^{(0)}\beta_{t_i})\sigma_c^{-1}(X_t^c)Y_t\right)dt \\ &\quad - \int_{t_{i-1}}^{t_i}a(t)\left(\sigma_c^{-1}(X_t^c)Y_t\beta_{t_{i-1}}\right)^* dW_t\Bigg\}\end{aligned}$$

*Proof.* Exactly the same proof as for Proposition 4.1.4 with redefined $\beta_t$ and $Z_t$.

*Remark.* As long as $\hat{\alpha}_k \neq 0$ for $k = 1,\ldots,m$ we have $\beta_{t_i} \in \bigcap_{j=1}^{d+m}\mathbb{D}^{(j)}$ for the stochastic volatility models as presented earlier.

## 4.2 Examples

In this section we will explicitly derive examples of stochastic weights for a couple of different pricing models. In particular we will look at jump diffusion versions of the Black–Scholes model and the Heston model.

In the Black–Scholes model the price process follows

$$dX_t = rX_{t-}dt + \sigma X_{t-}dW_t + (\alpha - 1)X_{t-}(dN_t - \lambda dt), \quad X_0 = x, \tag{4.5}$$

where $\mu$, $\sigma$ and $\alpha$ are constants and the intensity of the Poisson process is $\lambda$. The price process is clearly separable by Lemma 3.3.2 since it is a special case of a stochastic volatility model. In particular we see that the continuous price process contribution $X_t^c$ is identical to the Black–Scholes price process with modified drift.

Proposition 4.2.1: In the jump diffusion version of the Black–Scholes model defined in (4.5) we have for the following contingent claim

$$v = E\left[e^{-rT}\phi(X_T)\right],$$

the sensitivity weights:

$$\text{Rho}_{\text{BS}} = \frac{\partial v}{\partial r} = E\left[e^{-rT}\phi(X_T)\frac{W_T}{\sigma}\right] - E\left[Te^{-rT}\phi(X_T)\right]$$

$$\text{Delta}_{\text{BS}} = \frac{\partial v}{\partial x} = E\left[e^{-rT}\phi(X_T)\frac{W_T}{x\sigma T}\right]$$

$$\text{Gamma}_{\text{BS}} = \frac{\partial^2 v}{\partial x^2} = E\left[e^{-rT}\phi(X_T)\frac{1}{x^2\sigma T}\left(\frac{W_T^2}{\sigma T} - W_T - \frac{1}{\sigma}\right)\right]$$

$$\text{Vega}_{\text{BS}} = \frac{\partial v}{\partial \sigma} = E\left[e^{-rT}\phi(X_T)\left(\frac{W_T^2}{\sigma T} - W_T - \frac{1}{\sigma}\right)\right]$$

$$\text{Lambda}_{\text{BS}} = \frac{\partial v}{\partial \lambda} = E\left[e^{-rT}\phi(X_T)\left(\frac{N_T - \lambda T}{\lambda} - \frac{(\alpha-1)W_T}{\sigma}\right)\right]$$

$$\text{Alpha}_{\text{BS}} = \frac{\partial v}{\partial \alpha} = E\left[e^{-rT}\phi(X_T)\left(\frac{N_T}{\alpha} - \lambda T\right)\frac{W_T}{\sigma T}\right].$$

*Proof.* Since $X_t^c$ is identical to the Black–Scholes price process with modified drift the result follows by the same arguments as can be found in Fournié et al. (1999). □

Rho might need some further explanation. Rho is defined as the sensitivity of the claim with respect to interest rate, and here the interest rate appears both in the discount factor and in the drift. Analogous to section 4.1.1 we define the Rho as the sensitivity of the claim with respect to the factor $\epsilon$, as in the perturbed interest rate $r + \epsilon\gamma$, in the point $\epsilon = 0$.

The jump diffusion version of the Heston model

$$\begin{aligned} dX_t^{(1)} &= rX_{t-}^{(1)}dt + \sqrt{X_t^{(2)}}X_{t-}^{(1)}dW_t^{(1)} + (\alpha-1)X_{t-}^{(1)}(dN_t - \lambda dt), \quad X_0^{(1)} = x_1 \\ dX_t^{(2)} &= \kappa(\theta - X_t^{(2)})dt + \sigma\sqrt{X^{(2)}}dW_t^{(2)}, \quad X_0^{(2)} = x_2, \end{aligned} \tag{4.6}$$

with $E[dW_t^{(1)}dW_t^{(1)}] = \rho dt$ is also a special case of the general stochastic volatility model of Section 3.3 Chapter 3 and the separability assumption therefore holds by Lemma 3.3.2. $X_t^{(1)}$ is the process for the security on which the contingent claim is written, and $X_t^{(2)}$ is a process governing the volatility. The price of the claim in this setting is

$$v = E\left[e^{-rT}\phi(X_T^{(1)})\right],$$

and the continuous process $X_t^c$ is identical to the price process in the original Heston model, but again with modified drift.

**Proposition 4.2.2:** The stochastic weights for calculating the Greeks for the jump diffusion version of the Heston model (4.6) are

$$\text{Rho}_{\text{H}} = E\left[e^{-rT}\phi(X_T^{(1)})\left(\int_0^T \frac{dW_t^{(1)}}{\sqrt{x_t^{(2)}}} - \frac{\rho}{\sqrt{1-\rho^2}}\int_0^T \frac{dW_t^{(2)}}{\sqrt{x_t^{(2)}}}\right)\right] - E\left[Te^{-rT}\phi(X_T^{(1)})\right]$$

$$\text{Delta}_{\text{H}} = E\left[e^{-rT}\phi(X_T^{(1)})\frac{1}{x_1 T}\left(\int_0^T \frac{dW_t^{(1)}}{\sqrt{x_t^{(2)}}} - \frac{\rho}{\sqrt{1-\rho^2}}\int_0^T \frac{dW_t^{(2)}}{\sqrt{x_t^{(2)}}}\right)\right]$$

$$\text{Gamma}_{\text{H}} = E\left[e^{-rT}\phi(X_T^{(1)})\frac{1}{x_1^2 T^2}\left(\left(\int_0^T \frac{dW_t^{(1)}}{\sqrt{x_t^{(2)}}} - \frac{\rho}{\sqrt{1-\rho^2}}\int_0^T \frac{dW_t^{(2)}}{\sqrt{x_t^{(2)}}}\right)^2 - \frac{1}{1-\rho^2}\int_0^T \frac{dt}{x_t^{(2)}} - T\left(\int_0^T \frac{dW_t^{(1)}}{\sqrt{x_t^{(2)}}} - \frac{\rho}{\sqrt{1-\rho^2}}\int_0^T \frac{dW_t^{(2)}}{\sqrt{x_t^{(2)}}}\right)\right)\right]$$

$$\text{Vega}_{\text{H}}^{x_2} = E\left[e^{-rT}\phi(X_T^{(1)})\frac{1}{T}\left(\int_0^T \frac{Y_t^{(1,2)}}{x_t^{c(1)}\sqrt{x_t^{(2)}}}dW_t^{(1)} + \int_0^T \frac{T\frac{1}{\sigma}Y_t^{(2,2)} - \rho X_t^{c(1)}Y_t^{(1,2)}}{\sqrt{1-\rho^2}x_t^{c(1)}\sqrt{x_t^{(2)}}}dW_t^{(2)}\right)\right]$$

$$\text{Vega}_{\text{H}}^{\text{P}} = E\left[e^{-rT}\phi(X_T^{(1)})\frac{1}{T}\left(\left(W_T^{(1)} - \int_0^T \sqrt{X_t^{(2)}}dt\right)\left(\int_0^T \frac{dW_t^{(1)}}{\sqrt{x_t^{(2)}}} - \frac{\rho}{\sqrt{1-\rho^2}}\int_0^T \frac{dW_t^{(2)}}{\sqrt{x_t^{(2)}}}\right) - \int_0^T \frac{dt}{\sqrt{x_t^{(2)}}}\right)\right]$$

$$\text{Lambda}_{\text{H}} = E\left[e^{-rT}\phi(X_T^{(1)})\left(\frac{N_T}{\lambda} - T - (\alpha-1)\left(\int_0^T \frac{dW_t^{(1)}}{\sqrt{x_t^{(2)}}} - \frac{\rho}{\sqrt{1-\rho^2}}\int_0^T \frac{dW_t^{(2)}}{\sqrt{x_t^{(2)}}}\right)\right)\right]$$

$$\text{Alpha}_{\text{H}} = E\left[e^{-rT}\phi(X_T^{(1)})\left(\left(\frac{N_T}{\alpha T} - \lambda\right)\left(\int_0^T \frac{dW_t^{(1)}}{\sqrt{x_t^{(2)}}} - \frac{\rho}{\sqrt{1-\rho^2}}\int_0^T \frac{dW_t^{(2)}}{\sqrt{x_t^{(2)}}}\right)\right)\right].$$

*Remark.* The two vegas are different in the sense that $\text{Vega}_{\text{H}}^{x_2}$ is the sensitivity to changes in the initial value of the volatility process and $\text{Vega}_{\text{H}}^{\text{P}}$ is the sensitivity to a perturbation as in section 4.1.3. Since the initial value of the volatility process is just another parameter coming from the calibration $\text{Vega}_{\text{H}}^{x_2}$ is perhaps not as interesting as $\text{Vega}_{\text{H}}^{\text{P}}$ which is the analogue of $\text{Vega}_{\text{BS}}$.

*Proof.* We write the pricing process in matrix form as

$$dX_t = b(X_{t-})\,dt + a(X_{t-})\,dW_t + c(X_{t-})\,(dN_t - \lambda dt)$$

where

$$b(X_{t-}) = \begin{pmatrix} rX_{t-}^{(1)} \\ \kappa(\theta - X_t^{(2)}) \end{pmatrix}, \quad c(X_{t-}) = \begin{pmatrix} (\alpha - 1)X_{t-}^{(1)} \\ 0 \end{pmatrix} \quad \text{and}$$

$$a(X_{t-}) = \begin{pmatrix} \sqrt{X_t^{(2)}}X_{t-}^{(1)} & 0 \\ \sigma\sqrt{X_t^{(2)}}\rho & \sigma\sqrt{X_t^{(2)}}\sqrt{1-\rho^2} \end{pmatrix}.$$

The inverse of the diffusion matrix becomes

$$a(X_{t-})^{-1} = \frac{1}{\sigma X_{t-}^{(1)} X_t^{(2)} \sqrt{1-\rho^2}} \begin{pmatrix} \sigma\sqrt{X_t^{(2)}}\sqrt{1-\rho^2} & 0 \\ -\sigma\sqrt{X_t^{(2)}}\rho & \sqrt{X_t^{(2)}}X_{t-}^{(1)} \end{pmatrix}.$$

***Stochastic weight for* Rho *in the Heston model***

We perturb the original drift with $\gamma(x) = (x_1, 0)^*$ to get the perturbed process

$$dX_t^\epsilon = (b(X_{t-}^\epsilon) + \epsilon\gamma(X_{t-}^\epsilon))\,dt + a(X_{t-}^\epsilon))\,dW_t + c(X_{t-}^\epsilon)\,(dN_t - \lambda dt).$$

Now it is clear that the row–vector that should be integrated with respect to the Brownian motion is

$$\left( \frac{1}{\sqrt{X_t^{(2)}}}, \quad -\frac{\rho}{\sqrt{X_t^{(2)}}\sqrt{1-\rho^2}} \right),$$

and we get the expression in the proposition.

***Stochastic weight for* Delta *in the Heston model***

The first variation process is in this case a matrix process

$$dY_t = b'(X_t^c)Y_t dt + a_1'(X_t^c)Y_t dW_t^{(1)} + a_2'(X_t^c)Y_t dW^{(2)}, \quad Y_0 = I,$$

with

$$b'(X^c_{t-}) = \begin{pmatrix} r+\lambda(1-\alpha) & 0 \\ 0 & -\kappa \end{pmatrix}$$

$$a'_1(X^c_{t-}) = \begin{pmatrix} \sqrt{X^{c(2)}_t} & \frac{1}{2}\frac{X^{c(1)}_t}{\sqrt{X^{c(2)}_t}} \\ 0 & \frac{1}{2}\frac{\sigma\rho}{\sqrt{X^{c(2)}_t}} \end{pmatrix}$$

$$a'_2(X^c_{t-}) = \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{2}\frac{\sigma\sqrt{1-\rho^2}}{\sqrt{X^{c(2)}_t}} \end{pmatrix}.$$

For the Delta we are interested in the first row of the matrix $(a^{-1}(X^c_t)Y_t)^*$ integrated over the 2-dimensional Brownian motion. From the above we can deduce that the row vector of interest is

$$\left( \frac{1}{X^{c(1)}_t\sqrt{X^{c(2)}_t}} Y^{(1,1)}_t, \quad -\frac{\rho}{\sqrt{1-\rho^2}}\frac{1}{X^{c(1)}_t\sqrt{X^{c(2)}_t}} Y^{(1,1)}_t \right).$$

The first–row first–column element in the first variation process can be seen to be

$$Y^{(1,1)}_t = \frac{1}{x} X^{c(1)}_t,$$

since

$$dY^{(1,1)}_t = (r+\lambda(1-\alpha))Y^{(1,1)}_t dt + \sqrt{X^{c(2)}_t} Y^{(1,1)}_t dW^{(1)}_t + \frac{1}{2}\frac{X^{c(1)}_t}{\sqrt{X^{c(2)}_t}} Y^{(2,1)}_t dW^{(1)}_t$$

$$Y^{(1,1)}_0 = 1,$$

and $Y^{(2,1)}_t = 0$ due to the fact that $X^{c(2)}_t$ does not depend explicitly on $X^{c(1)}_t$ and in particular not its initial value. This leaves us with the expression for $\text{Delta}_H$ presented in the proposition.

### *Stochastic weight for* Gamma *in the Heston model*

Differentiating the expression for **Delta** we get two contributions. One contribution comes from the derivative of the $\frac{1}{x_1 T}$ factor and the other comes from repeating

the steps in the proof of Proposition 4.1.3. In the latter process we end up with the following Skorohod integral

$$\delta^{(0)}\left(\frac{1}{x_1 T}\left(\int_0^T \frac{dW_t^{(1)}}{\sqrt{x_t^{(2)}}} - \int_0^T \frac{\rho}{\sqrt{1-\rho^2}}\frac{dW_t^{(2)}}{\sqrt{x_t^{(2)}}}\right)\frac{1}{T}a^{-1}(X^c)Y\right),$$

which can be evaluated using Proposition 3.2.10 to get (we are interested in the first element of the Skorohod integral vector)

$$\frac{1}{(x_1 T)^2}\left(\int_0^T \frac{dW_t^{(1)}}{\sqrt{x_t^{(2)}}} - \int_0^T \frac{\rho}{\sqrt{1-\rho^2}}\frac{dW_t^{(2)}}{\sqrt{x_t^{(2)}}}\right)^2 - \frac{1}{(x_1 T)^2}\frac{1}{1-\rho^2}\int_0^T \frac{1}{x_t^{(2)}}dt.$$

This, together with the contribution from the derivative of $\frac{1}{x_1 T}$, leaves us with the expression in the Proposition.

### *Stochastic weight for* $\mathbf{Vega}^{x_2}$ *in the Heston model*

The interpretation of Vega as the sensitivity with respect to initial volatility level follows directly from the discussion above. In this case the row–vector to be integrated over the Brownian motion is

$$\left(\frac{1}{X_t^{c(1)}\sqrt{X_t^{c(2)}}}Y_t^{(1,2)}, \quad \frac{1}{\sqrt{1-\rho^2}}\frac{1}{X_t^{c(1)}\sqrt{X_t^{c(2)}}}\left[\frac{1}{\sigma}Y_t^{(2,2)} - \rho X_t^{c(1)}Y_t^{(1,2)}\right]\right).$$

### *Stochastic weight for* $\mathbf{Vega}^{\mathbf{P}}$ *in the Heston model*

We perturb the original diffusion matrix with $\gamma$ to get the perturbed process

$$dX_t^\epsilon = b(X_{t-}^\epsilon)\,dt + (a(X_{t-}^\epsilon) + \epsilon\gamma(X_{t-}^\epsilon))\,dW_t + c(X_{t-}^\epsilon)\,(dN_t - \lambda dt),$$

where

$$\gamma(x) = \begin{pmatrix} x_1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Again using the fact that $X_t^{(2)}$ does not depend explicitly on $X_t^{(1)}$ we can deduce that the variation process with respect to $\epsilon$, $Z_t^\epsilon$, has a vanishing second component.

We write $Z_t^{(2)} = 0$. Furthermore,

$$\begin{aligned}\beta_t &= Y_t^{-1}\frac{\partial X_t^c}{\partial X_t}Z_t \\ &= \frac{1}{Y_t^{(1,1)}Y_t^{(2,2)}}\begin{pmatrix} Y_t^{(2,2)} & -Y_t^{(1,2)} \\ 0 & Y_t^{(1,1)}\end{pmatrix}\begin{pmatrix} \frac{\partial X_t^{c(1)}}{\partial X_t^{(1)}} & \frac{\partial X_t^{c(1)}}{\partial X_t^{(2)}} \\ \frac{\partial X_t^{c(2)}}{\partial X_t^{(1)}} & \frac{\partial X_t^{c(2)}}{\partial X_t^{(2)}}\end{pmatrix}\begin{pmatrix} Z_t^{(1)} \\ 0\end{pmatrix} \\ &= \begin{pmatrix} Z_t^{(1)}\frac{\partial X_t^{c(1)}}{\partial X_t^{(1)}} / \ Y_t^{(1,1)} \\ 0\end{pmatrix}.\end{aligned}$$

Note that by the Itô formula

$$Z_t^{(1)} = X_t^{(1)}\left(W_t^{(1)} - \int_o^t \sqrt{X_s^{(2)}}ds\right),$$

so that

$$\beta_t^{(1)} = x_1\left(W_t^{(1)} - \int_0^t \sqrt{X_s^{(2)}}ds\right).$$

Except for expressions already derived, the only thing we need in order to calculate the weight for Vega$^P$ is the Malliavin derivative of $\beta_t$ in the direction of the Brownian motion. Using the chain rule (Theorem 3.2.4) on a sequence of continuously differentiable functions with bounded derivatives approximating $\sqrt{X_s^{(2)}}$, together with (4.1) we get

$$\begin{aligned}D_t^{(0)}\beta_T^{(1)} &= x_1\left((1,0) - \frac{1}{2}\int_0^T \frac{1}{\sqrt{X_s^{(2)}}}D_t^{(0)}X_s^{(2)}ds\right) \\ &= x_1\left((1,0) - \frac{1}{2}\int_t^T \sigma\frac{\sqrt{X_t^{(2)}}}{\sqrt{X_s^{(2)}}}\frac{Y_s^{(2,2)}}{Y_t^{(2,2)}}(\rho,\sqrt{1-\rho^2})ds\right).\end{aligned}$$

To arrive at the expression in the proposition we note that

$$\mathrm{Tr}\left((D_t^{(0)}\beta_T)a^{-1}Y_t\right) = \frac{1}{\sqrt{X_t^{(2)}}}.$$

*Stochastic weight for* **Lambda** *in the Heston model*

The quantity we need to expand in this case is $a^{-1}(X_{t-})c(X_{t-})$. From what was stated in the beginning of this section it is clear that the vector to be integrated over Brownian motion is

$$(\alpha - 1)\left(\frac{1}{\sqrt{X_t^{(2)}}}, \quad -\frac{\rho}{\sqrt{X_t^{(2)}}\sqrt{1-\rho^2}}\right).$$

*Stochastic weight for* **Alpha** *in the Heston model*

The perturbed process is chosen as

$$dX_t^\epsilon = b(X_{t-}^\epsilon)\,dt + a(X_{t-}^\epsilon)\,dW_t + (c(X_{t-}^\epsilon) + \epsilon\gamma(X_{t-}^\epsilon))\,(dN_t - \lambda dt),$$

with

$$\gamma(x) = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}.$$

As was the case for Vega$^{\mathrm{P}}$ $X_t^{(2)}$ does not depend explicitly on $X_t^{(1)}$ so we get

$$\begin{aligned}
\beta_t &= Y_t^{-1}\frac{\partial X_t^c}{\partial X_t} Z_t \\
&= \frac{1}{Y_t^{(1,1)}Y_t^{(2,2)}} \begin{pmatrix} Y_t^{(2,2)} & -Y_t^{(1,2)} \\ 0 & Y_t^{(1,1)} \end{pmatrix} \begin{pmatrix} \frac{\partial X_t^{c(1)}}{\partial X_t^{(1)}} & \frac{\partial X_t^{c(1)}}{\partial X_t^{(2)}} \\ \frac{\partial X_t^{c(2)}}{\partial X_t^{(1)}} & \frac{\partial X_t^{c(2)}}{\partial X_t^{(2)}} \end{pmatrix} \begin{pmatrix} Z_t^{(1)} \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} Z_t^{(1)}\frac{\partial X_t^{c(1)}}{\partial X_t^{(1)}} / \; Y_t^{(1,1)} \\ 0 \end{pmatrix}.
\end{aligned}$$

Note that by the Itô formula

$$Z_t^{(1)} = X_t^{(1)}\left(\frac{N_t}{\alpha} - \lambda t\right),$$

so that

$$\beta_t^{(1)} = x_1\left(\frac{N_t}{\alpha} - \lambda t\right).$$

The Skorohod integral expansion in Proposition 3.2.10 is easy to compute in this case since $D_t^{(0)}\beta_t = 0$. Adding up the results we arrive at the formula in the proposition. □

## 4.3 Numerical experiments

To show the power of the stochastic weight technique a number of convergence plots are presented in this section. In Figures 4.1 to 4.4 below the value of the Delta is captured at different iterations of the simulation.

Looking at the convergence plots it is evident that the finite difference approximation performs better for the call option, but the Malliavin weight approach performs better for the digital. The reason for this is that the stochastic weight adds some randomness to the expression to be simulated, and for the call option this effect worsens the convergence. However, for the discontinuous payoff function of the digital option, the finite difference approximation produces large errors since the contributions to be averaged are either zero or one. In this case the Malliavin weight and payoff is still smooth and the effect of added randomness through the stochastic weight is not prominent.

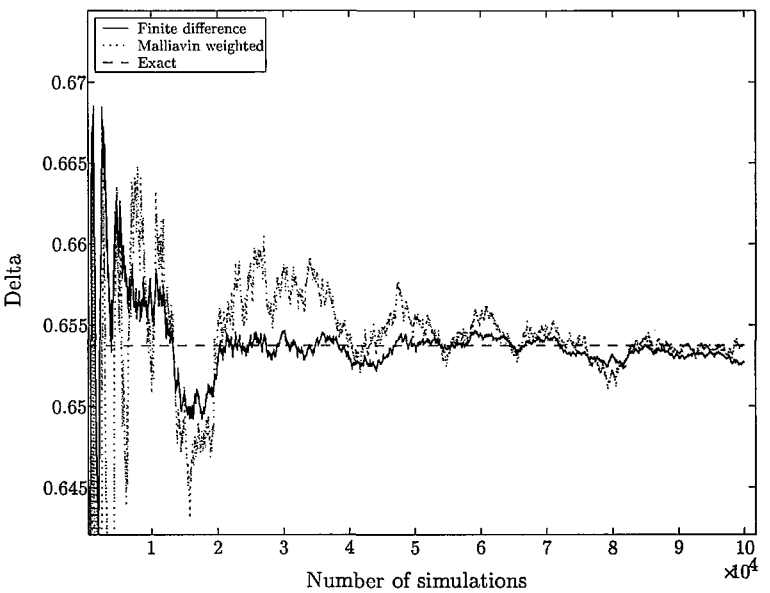


**Figure 4.1:** Monte Carlo simulation of $\text{Delta}_{BS}$ for a call option using finite difference approximation and Malliavin weighting. The model parameters are $x = 100$, Strike $= 100$, $r = 0.05$, $\sigma = 0.3$, $\alpha = 0.5$ and $\lambda = 0.1$.

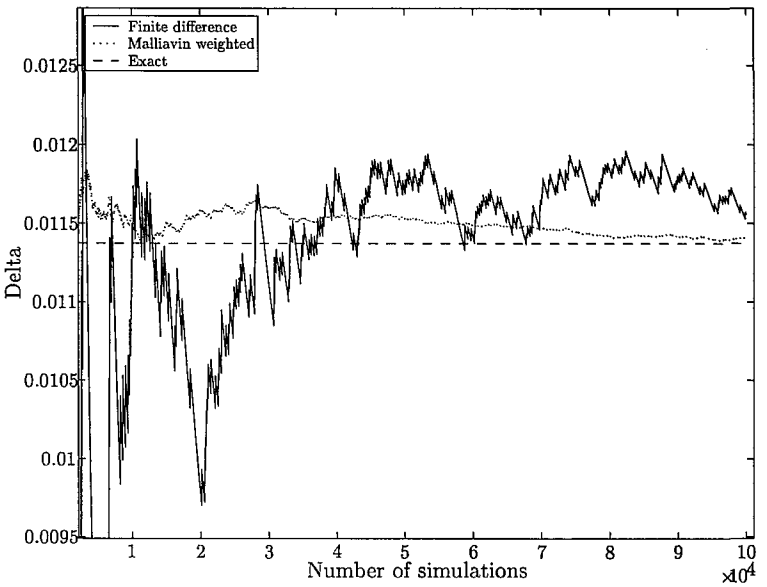


**Figure 4.2:** Monte Carlo simulation of $\text{Delta}_{\text{BS}}$ for a digital option using finite difference approximation and Malliavin weighting. The model parameters are $x = 100$, Strike $= 100$, $r = 0.05$, $\sigma = 0.3$, $\alpha = 0.5$ and $\lambda = 0.1$.

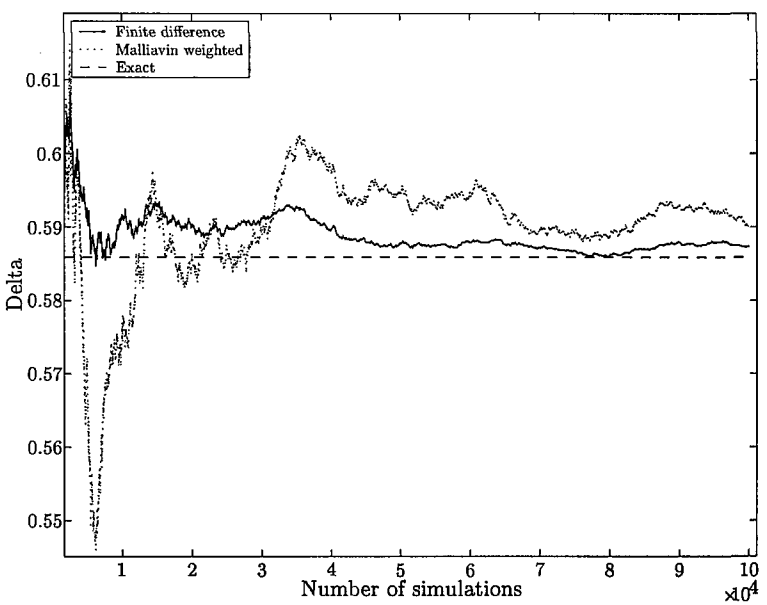


**Figure 4.3:** Monte Carlo simulation of $\text{Delta}_{\text{H}}$ for a call option using finite difference approximation and Malliavin weighting. The model parameters are $x_1 = 100$, Strike $= 100$, $x_2 = 0.04$, $r = 0.05$, $\kappa = 1$, $\theta = 0.04$, $\sigma = 0.04$, $\rho = -0.8$, $\alpha = 0.5$ and $\lambda = 0.1$.

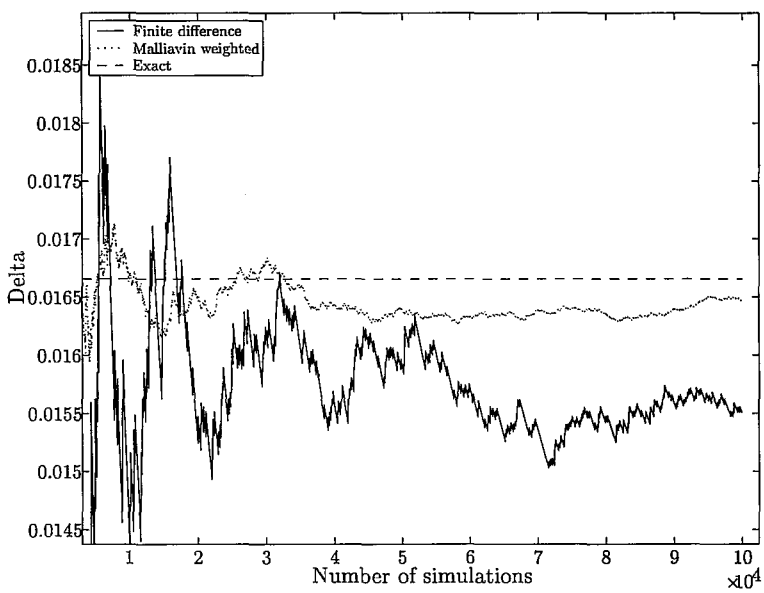


**Figure 4.4:** Monte Carlo simulation of $\text{Delta}_{\text{H}}$ for a digital option using finite difference approximation and Malliavin weighting. The model parameters are $x_1 = 100$, Strike $= 100$, $x_2 = 0.04$, $r = 0.05$, $\kappa = 1$, $\theta = 0.04$, $\sigma = 0.04$, $\rho = -0.8$, $\alpha = 0.5$ and $\lambda = 0.1$.

# 5. STOCHASTIC WEIGHTS FOR CREDIT DERIVATIVE GREEKS

## 5.1 Intensity sensitivity for single name claims

### 5.1.1 Constant intensity

Remember that, as shown in Proposition 4.1.5, for a price process of the form

$$dX_t = b(X_{t-})dt + a(X_{t-})dW_t + c(X_{t-})\left(dN_t - \lambda dt\right), \quad X_0 = x,$$

we have for $v(\lambda) = E\left[\Phi(X_T)\right]$:

$$\frac{\partial v}{\partial \lambda} = E\left[\Phi(X_T)\left(\frac{N_T}{\lambda} - T - \int_0^T a^{-1}(X_{t-})c(X_{t-})dW_t\right)\right].$$

For credit derivatives, depending only on one reference credit, the situation is much simpler. The payoff function depends on the first jump time of some counting process—the time when the reference company defaults on its debt. In the above framework the process under consideration is $dX_t = dN_t$ and the contingent claim is expressed as

$$v = E\left[\Phi(\tau)\right], \tag{5.1}$$

where $\tau = \inf\{t > 0 : N_t = 1\}$. Also, remember that the above sensitivity result is true for path–dependent payoffs such as payoffs depending on the first jump time. Typically the payoff function $\Phi$ contains an indicator function flagging for the event that the default time occurs before maturity of the contract. This discontinuity suggests that the stochastic weight approach should be successful in reducing the variance of the Monte Carlo estimator of the intensity sensitivity as we saw in Section 4.3 Chapter 4.

In the same spirit as in Section 4.1.4 Chapter 4 we introduce the perturbed process

$$X_t^\epsilon = N_t^\epsilon,$$

where $\epsilon$ is a deterministic parameter and $N_t^\epsilon$ is a Poisson process with intensity $\lambda + \epsilon$. We further define $\tau^\epsilon$ to be the first jump time of this perturbed process.

Proposition 5.1.1: The sensitivity of the claim (5.1) with respect to jump intensity can be written as

$$\frac{\partial v}{\partial \lambda} = E\left[\Phi(\tau)\left(\frac{1}{\lambda} - \tau\right)\right].$$

*Proof.* The proof is basically the same as for the case including a continuous part, Proposition 4.1.5. Changing measure we can write

$$E\left[\Phi(\tau^\epsilon)\right] = E\left[M_\tau^\epsilon \Phi(\tau)\right],$$

where

$$M_t^\epsilon = \left(1 + \frac{\epsilon}{\lambda}\right)^{N_t} \exp(-\epsilon t).$$

It is true that

$$M_t^\epsilon = 1 + \frac{\epsilon}{\lambda}\int_0^t M_t^\epsilon (dN_t - \lambda dt).$$

Now, the sought after derivative can be written

$$\begin{aligned}\frac{\partial}{\partial \epsilon} v^\epsilon\Big|_{\epsilon=0} = \lim_{\epsilon\to 0} E\left[\frac{\Phi(\tau^\epsilon) - \Phi(\tau)}{\epsilon}\right] = \lim_{\epsilon\to 0} E\left[\Phi(\tau)\left(\frac{M_\tau^\epsilon - 1}{\epsilon}\right)\right] \\ = E\left[\Phi(\tau)\left(\int_0^\tau \frac{1}{\lambda}(dN_t - \lambda dt)\right)\right].\end{aligned}$$

□

As will be seen in the next chapter, it is not surprising that the stochastic weight in Proposition 5.1.1 takes exactly the same form as the well known weight obtained via the likelihood ratio approach (c.f. Joshi and Kainth (2003)).

### 5.1.2 Cox processes

A natural extension of the above modelling of the default mechanism is to make the jump intensity stochastic. This provides us with a more realistic description of credit spreads, which in the real world are far from constant as implied by the

assumption of constant intensities. Jump processes of this type are known as Cox processes.

We let the jump intensity evolve according to

$$d\lambda_t = b(\lambda_t)dt + a(\lambda_t)dW_t, \tag{5.2}$$

and investigate the effect of perturbed intensity exactly as before. Hence, we define the perturbed jump process $N_t^\epsilon$ as the process with constantly perturbed intensity $\lambda_t + \epsilon$.

Proposition 5.1.2: For $v^\epsilon$ defined as

$$v^\epsilon = E\left[\Phi(\tau^\epsilon)\right],$$

we have

$$\frac{\partial}{\partial \epsilon} v^\epsilon \Big|_{\epsilon=0} = E\left[\Phi(\tau)\left(\frac{1}{\lambda_\tau} - \tau\right)\right].$$

*Proof.* Changing measure[1] as in the proof of Proposition 5.1.1 we note that

$$M_t^\epsilon = 1 + \int_0^t M_{t^-}^\epsilon \frac{\epsilon}{\lambda_t}(dN_t - \lambda_t dt),$$

and the result follows exactly as before. □

In practice an intensity model as (5.2) is calibrated to the market. Each name has several outstanding credit obligations with different maturities resulting in a default intensity curve—the credit curve. That is, for each maturity $T_i$, $i = 1, \ldots, n$, we can find a $\hat{\lambda}_{T_i}$ such that the risk neutral probability (implied by the market value of the debt) of survival of the company until maturity of the contract can be written as

$$P(T_i) = e^{-\hat{\lambda}_{T_i} T_i},\ i = 1, \ldots, n.$$

The intensity model (5.2) would then be chosen as to best replicate these probabilities. For a perfect fit we would get

$$E\left[e^{-\int_0^{T_i} \lambda_t dt}\right] = e^{-\hat{\lambda}_{T_i} T_i},\ i = 1, \ldots, n,$$

[1] For a very general Girsanov theorem see Schönbucher (2003) Theorem 4.8.

and thus,

$$E\left[e^{-\int_0^{T_i}(\lambda_t+\epsilon)dt}\right] = e^{-(\hat{\lambda}_{T_i}+\epsilon)T_i},\ i = 1,\ldots,n.$$

From the above we see that the sensitivity to the perturbed Cox process of Proposition 5.1.2 corresponds to the sensitivity to a parallel shift in the credit curve. In practical situations the calibration never yields a perfect fit as assumed above, but the argument can be carried through the minimisation procedure of the calibration.

### 5.1.3 Example

As an example of a Cox process based credit model consider the two factor Cox–Ingersoll–Ross (CIR) model (c.f. Schönbucher (2003) section 7.2). Both the risk–free interest rate and the jump intensity are taken to be combinations of the stochastic processes

$$dx_t^{(1)} = (\alpha_1 - \beta_1 x_t^{(1)})dt + \sigma_1\sqrt{x_t^{(1)}}dW_t^{(1)}$$
$$dx_t^{(2)} = (\alpha_2 - \beta_2 x_t^{(2)})dt + \sigma_2\sqrt{x_t^{(2)}}dW_t^{(2)},$$

where $dW^{(1)}$ and $dW^{(2)}$ are independent Brownian motions. The risk–free interest rate $r$ and the jump intensity $\lambda$ are expressed as

$$r_t = w\,x_t^{(1)} + (1-w)x_t^{(2)}$$
$$\lambda_t = \bar{w}\,x_t^{(1)} + (1-\bar{w})x_t^{(2)},$$

where $w$ and $\bar{w}$ are positive constants in order to ensure that the interest rate and the intensity stay positive.

Note how this model can only generate positive correlation between the interest rate and the jump intensity. This is typically not the case in the real world, but the CIR model still serve the purpose of fitting the credit curve satisfactory.

Under the CIR assumptions there exist closed–form solutions to the risky and the risk–free bond prices. If we assume a recovery rate for the risky bond we can use these closed form solutions to calibrate the parameters to a credit curve and a risk–free yield curve. We can then use Monte Carlo techniques to price and hedge exotic claims with the estimated parameters.

In Figure 5.1 the CIR fit to the Libor curve, the Bombardier credit spread curve and the parallel shifted Bombardier credit spread curve is presented. Assuming a recovery rate of 50% this calibration resulted in two sets of parameters: one set for the original credit spread curve and one set for the parallel shifted credit spread curve. These two sets of parameters can be used for a finite difference approximation serving as a benchmark for the sensitivity calculation suggested by Proposition 5.1.2. In Figure 5.2 the Monte Carlo convergence plot for the sensitivity to parallel shift in credit curve of a simple claim on Bombardier paying \$1 at default if this happens before maturity is shown.

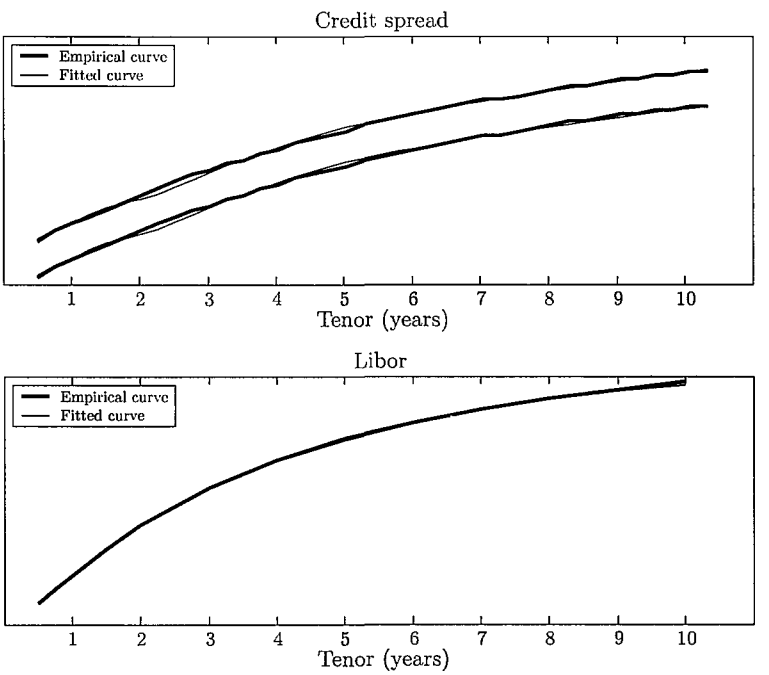


**Figure 5.1:** CIR fit to Libor curve, Bombardier credit spread curve and parallel shifted Bombardier credit spread curve as of 23 April 2004. Sources: Mark-It partners (BOMB SNRFOR USD MR) and Citigroup.

The advantage of the stochastic weight approach is threefold. Not only is the convergence faster, we also avoid a potentially costly second calibration. The most important benefit is however that the two calibration errors might affect the price in different directions. This introduces bias in the Monte Carlo simulation of the Greek, and bias is far worse than variance. In the Bombardier example above it

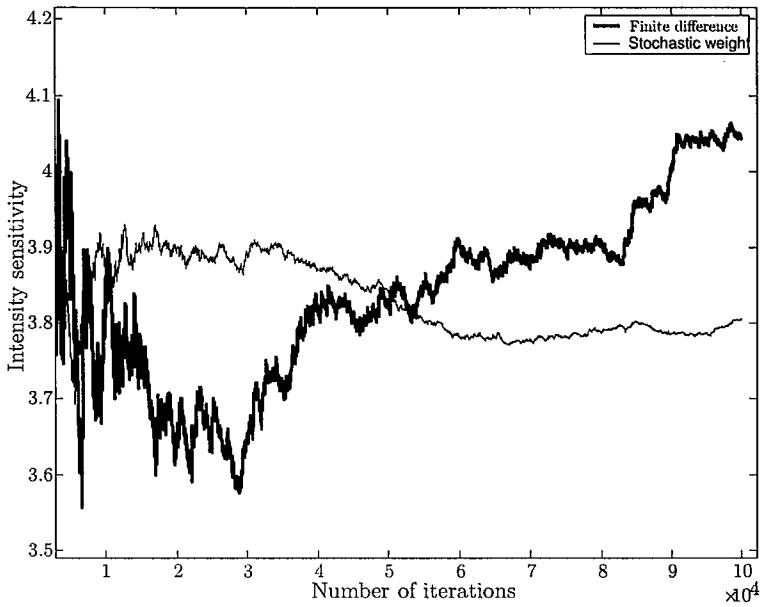


**Figure 5.2:** Monte Carlo simulation of the sensitivity to parallel shifts in credit spread curve for a 5 year digital credit default swap on Bombardier as of 23 April 2004.

was found that in order to obtain robustness in the two calibrations, the weights $w$ and $\bar{w}$ should be kept constant.

## 5.2 Correlation sensitivity for baskets of credits

### 5.2.1 Correlated default events

When pricing credit derivatives depending on more than one reference credit, default time correlation must be taken into account. Given the above presentation, the first method to achieve default correlation that springs to mind is perhaps to extend the Cox process framework with several correlated default intensities. If the intensities are strongly positively correlated, high jump intensities have a tendency of occurring simultaneously implying positive default time correlations. Unfortunately, this dependence is too weak and this approach gives far too low default correlations compared to observed ones.

A cure for unrealistic default correlations is to introduce jumps in the intensity process. If we let a default by one company trigger a jump in another, we obtain

a positive relationship between default times, which is strong enough for realistic default correlations. See for example the model presented in Davis and Lo (2001). As presented in Schönbucher and Schubert (2000) jumps in the intensity process turns out to be equivalent to the copula approach presented briefly below.

The copula approach models the individual default times as the first jump times of either Poisson processes as in Section 5.1.1 or Cox processes as in Section 5.1.2, and then uses a copula to impose a dependance structure. This was pioneered in Li (2000) where constant intensities were used. An algorithm for generating $n$ correlated default times in what is commonly known as the Li model can be outlined as

1. Simulate $\tilde{Z}_1, \ldots, \tilde{Z}_n$ as correlated standard normal random numbers.
2. Generate the default times $\tau_i$, $i = 1, \ldots, n$ through $N(\tilde{Z}_i) = 1 - e^{-\lambda_i \tau_i}$, where $N$ is the normal cumulative distribution function.

Note how the resulting default times are exponentially distributed and connected via a Gaussian copula.

### 5.2.2 Correlation sensitivity

The connection of random default times through copulae seems far from the process based approach in the preceding, but the algorithm in the previous section gives us a hint to how to construct stochastic weights for correlation sensitivity. The price of a claim on a basket of credits is written as

$$v = E\left[\Phi(\tau_1, \ldots, \tau_n)\right].$$

But we know that $\tau_i = -\frac{1}{\lambda}\log\left(1 - N(\tilde{Z}_i)\right)$, so we can rewrite the price of the claim as

$$v = E\left[\widetilde{\Phi}(\tilde{Z}_1, \ldots, \tilde{Z}_n)\right],$$

for some payoff function $\widetilde{\Phi}$. Further, the correlated normal random variables $\tilde{Z}_1, \ldots, \tilde{Z}_n$ can be seen as the values of an $n$–dimensional diffusion process $X_t$ at time 1, where $X_t$ has zero drift and a diffusion matrix replicating the relevant correlation. We write

$$dX_t = \sigma dW_t, \quad X_0 = 0_n, \tag{5.3}$$

where $0_n$ is an $n$–dimensional column vector of zeros. With the problem rewritten on this form we can use the diffusion matrix result of Proposition 4.1.4 to obtain stochastic weights for the correlation Greeks.

In a basket of $n$ credits there are $\frac{1}{2}n(n-1)$ default correlations. In practical applications however, the problem is usually simplified and correlations are assumed to be the same; either the same across all credits or the same within industry sectors. In the following we will illustrate the correlation sensitivity technique for a basket of two credits and then show how to obtain weights for the homogenous correlation case.

We start with the correlated system of equation (5.3) and let

$$\sigma = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix}.$$

The perturbed process in this case, $X_t^\epsilon = \sigma^\epsilon dW$, has a perturbed diffusion matrix according to

$$\sigma^\epsilon = \begin{pmatrix} 1 & 0 \\ \rho+\epsilon & \sqrt{1-(\rho+\epsilon)^2} \end{pmatrix},$$

where $\epsilon$ is small enough for $\rho+\epsilon$ to be less than 1. As in the diffusion matrix result we need the variation process to the parameter $\epsilon$:

$$Z_t^\epsilon = \begin{pmatrix} 0 & 0 \\ 1 & -\frac{\rho+\epsilon}{\sqrt{1-(\rho+\epsilon)^2}} \end{pmatrix} \begin{pmatrix} W_t^{(1)} \\ W_t^{(2)} \end{pmatrix}.$$

**Proposition 5.2.1:** Let the perturbed process $X_t^\epsilon$ translate to the default times $\tau_i^\epsilon$ for $i = 1, \ldots, n$. For $v^\epsilon(\rho)$ defined as

$$v^\epsilon(\rho) = E\left[\Phi(\tau_1^\epsilon, \ldots, \tau_n^\epsilon)\right],$$

we have

$$\begin{aligned} \frac{\partial v}{\partial \rho} = \frac{\partial}{\partial \epsilon} v^\epsilon \Big|_{\epsilon=0} &= E\left[\Phi(\tau_1, \ldots, \tau_n)\delta^{(0)}\left(\sigma^{-1}Z_1\right)\right] \\ &= E\left[\Phi(\tau_1, \ldots, \tau_n)\frac{1}{1-\rho^2}\left(\left(\sqrt{1-\rho^2}\, W_1^{(1)} + \rho\, W_1^{(2)}\right)W_1^{(2)} + \rho\right)\right]. \end{aligned}$$

*Proof.* The only difference from Proposition 4.1.4 is the way the diffusion matrix is perturbed. This does not pose any problems in carrying through the argument as before. □

*Remark.* In the general case with more than two credits the sensitivity to one pairwise correlation coefficient can always be obtained in exactly the same way as in Proposition 5.2.1 by labelling the names under consideration as number one and number two.

The common way of moving from a correlation matrix to a diffusion matrix is by Cholesky decomposition. As hinted in the proof of Proposition 5.2.1 this is what makes the correlation sensitivity different from the diffusion matrix sensitivity. For a multidimensional correlation matrix the $\epsilon$ perturbation is translated, via the Cholesky decomposition, to the diffusion matrix in a non–trivial way.

One way of dealing with the homogenous correlation matrix case is to introduce an additional process acting as a common factor. This will be outlined in the next section where it will also be shown that the common factor approach will result in higher variance than the approach presented in the following.

Assume that the basket contains $n$ credits and that the correlation coefficient to be used in the Li–model, $\rho$, is the same for all pairs. The strategy will be to preserve the structure of the correlation matrix when constructing the diffusion matrix so that all the diagonal elements are the same and all the off–diagonal elements are the same. Two conditions must be satisfied: the variance of each component of $X_t = \sigma W_t$ must be equal to 1 and the correlation between two components must be $\rho$, i.e.

$$E[dX_t^{(i)} dX_t^{(j)}] = \begin{cases} dt, \text{ for } i = j \\ \rho\, dt, \text{ for } i \neq j. \end{cases}$$

We end up with the diffusion matrix

$$\sigma = \begin{pmatrix} \sqrt{1-(n-1)\hat{\rho}^2} & \hat{\rho} & \dots & \hat{\rho} \\ \hat{\rho} & \sqrt{1-(n-1)\hat{\rho}^2} & \dots & \hat{\rho} \\ \vdots & & & \\ \hat{\rho} & \hat{\rho} & \dots & \sqrt{1-(n-1)\hat{\rho}^2} \end{pmatrix}, \quad (5.4)$$

where $\hat{\rho}$ is the solution to the equation

$$\rho = \hat{\rho}\left(2\sqrt{1-(n-1)\hat{\rho}^2} + (n-2)\hat{\rho}\right), \tag{5.5}$$

which can be solved numerically.

Proposition 5.2.2: By substituting $\hat{\rho}+\epsilon$ for $\hat{\rho}$ in (5.4) we obtain the perturbed default times $\tau_i^\epsilon$, $i = 1,\ldots,n$. For $v^\epsilon(\hat{\rho})$ defined as

$$v^\epsilon(\hat{\rho}) = E\left[\Phi(\tau_1^\epsilon,\ldots,\tau_n^\epsilon)\right],$$

we have

$$\begin{aligned}\frac{\partial v}{\partial \hat{\rho}} = \frac{\partial}{\partial \epsilon} v^\epsilon\Big|_{\epsilon=0} &= E\left[\Phi(\tau_1,\ldots,\tau_n)\delta^{(0)}\left(\sigma^{-1}\eta W_\cdot\right)\right] \\ &= E\left[\Phi(\tau_1,\ldots,\tau_n)\left(W_1^*(\sigma^{-1}\eta)^* W_1 - \mathrm{Tr}(\sigma^{-1}\eta)\right)\right],\end{aligned}$$

where

$$\eta = \begin{pmatrix} \frac{(n-1)\hat{\rho}}{\sqrt{1-\hat{\rho}^2}} & 1 & \ldots & 1 \\ \vdots & & & \\ 1 & 1 & \ldots & \frac{(n-1)\hat{\rho}}{\sqrt{1-\hat{\rho}^2}} \end{pmatrix}.$$

*Proof.* The variation process with respect to $\epsilon$, $Z_t^\epsilon$, has diffusion matrix $\eta$. The steps in the proof of Proposition 5.2.1 completes the argument. □

Corollary 5.2.3: By using the chain rule we get the sensitivity of the claim to the correlation coefficient:

$$\frac{\partial v}{\partial \rho} = \frac{\partial \hat{\rho}}{\partial \rho}\frac{\partial v}{\partial \hat{\rho}}.$$

### 5.2.3 Examples and extensions

As mentioned in the previous section there is an alternative way to Corollary 5.2.3 of dealing with the homogenous default correlation matrix of $n$ credits. Instead of constructing a diffusion matrix using the solution to (5.5) we can introduce an extra Brownian motion and an extra row in the diffusion matrix so that

$$\begin{pmatrix} X_t^{(1)} \\ \vdots \\ X^{(n)} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ \sqrt{\rho} & \sqrt{1-\rho} & 0 & \ldots & 0 \\ \sqrt{\rho} & 0 & \sqrt{1-\rho} & \ldots & 0 \\ \vdots & & & & \\ \sqrt{\rho} & 0 & 0 & \ldots & \sqrt{1-\rho} \end{pmatrix} \begin{pmatrix} W_t^{(0)} \\ \vdots \\ W_t^{(n)} \end{pmatrix}.$$

The stochastic weight for the correlation sensitivity can now be obtained in the same fashion as in Proposition 5.2.2. In Figure 5.3 we can see that this *extension of diffusion matrix* approach appears to yield a slightly noisier stochastic weight than does the *homogenous diffusion matrix* approach derived in Proposition 5.2.2. To investigate why we should focus on the quadratic form

$$W_1^*(\sigma^{-1}\eta)^*W_1.$$

Both approaches will yield a quadratic form as the one above and these are the sources of randomness for the stochastic weights. For all values of $\rho$ the elements of $\sigma^{-1}\eta$ tend to have larger magnitude in the extension of diffusion matrix approach than in the homogenous diffusion matrix approach leading to the convergence behaviour of Figure 5.3.

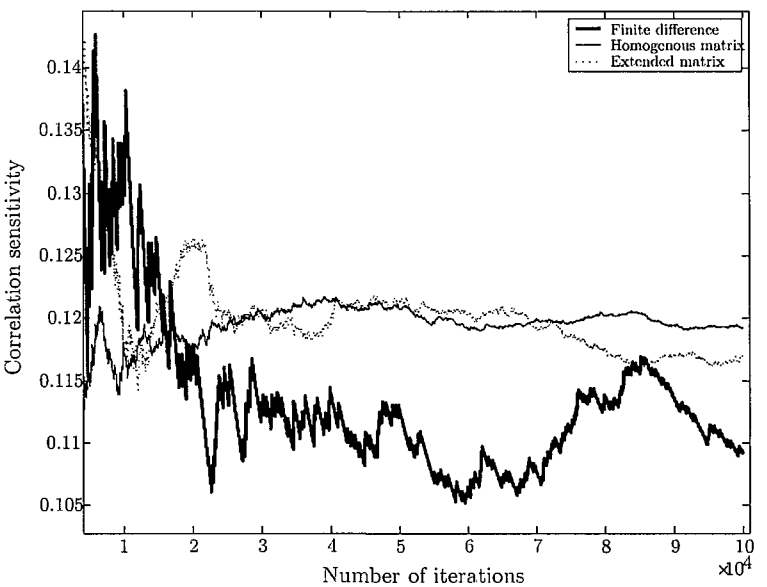


**Figure 5.3:** Monte Carlo simulation of the correlation sensitivity for a one year digital second to default credit swap on a basket of two credits. Both jump intensities were 0.3 and the correlation was 0.3.

Since the Gaussian copula does not have tail dependence academic research has investigated different copulae. Perhaps the most logical second step is to look at the $t$–copula, which is defined from the multivariate $t$–distribution the same way as the Gaussian copula is defined from the multivariate normal distribution.

Even though at first sight the Gaussian copula looks to be as far as we can go with the Gaussian Malliavin calculus approach, it turns out that we can make the extension to the $t$–copula.

The multivariate $t$–distribution with $k$ degrees of freedom is defined as the distribution of a normally distributed random vector (with some correlation matrix) divided by the square root of a scaled $\chi^2$–distributed random variable with $k$ degrees of freedom. The $\chi^2$–distributed random variable with $k$ degrees of freedom can in turn be expressed as a sum of $k$ squared normally distributed random variables. Everything fits perfectly in the above Gaussian framework, and the correlation sensitivity can be obtained with the stochastic weight of Proposition 5.2.1. A convergence plot is shown in Figure 5.4.

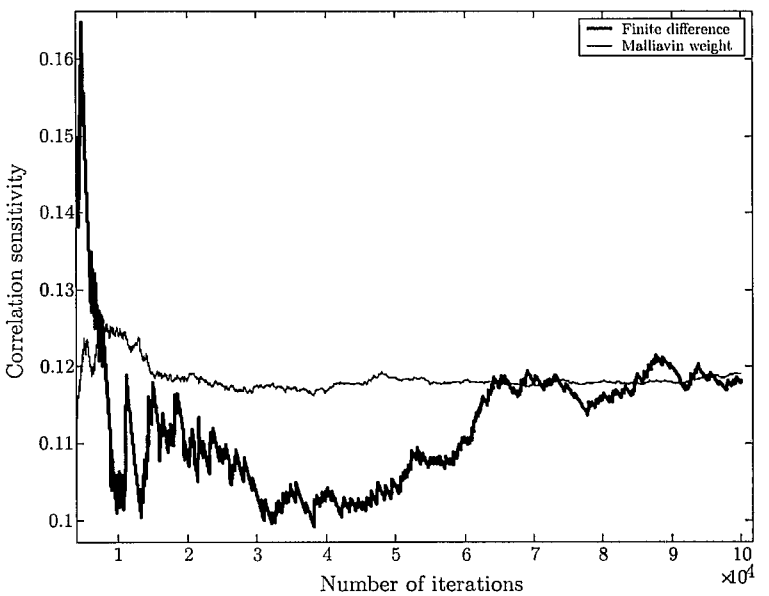


**Figure 5.4:** Monte Carlo simulation of the correlation sensitivity for a one year digital second to default credit swap on a basket of two credits. A $t$–copula with 2 degrees of freedom was used, both jump intensities were 0.3 and the correlation was 0.3.

## 5.3 Alternative methods

In the examples of Sections 5.1.3 and 5.2.3 the stochastic weight approaches were compared to naïve finite difference methods. It could be argued that this makes the

results look better than they actually are since there are well known methods to improve on the Monte Carlo convergence of credit derivatives. An implementation in practice would typically incorporate *stratified sampling* and *importance sampling*.

Stratified sampling is an important tool when the probabilities of default in a basket are significantly different in magnitude. In the examples above the probabilities were chosen constant across credits to avoid that type of effects, but in practice there is no such luxury as choosing.

A serious rival to the stochastic weight approach is to use finite difference approximations of Monte Carlo runs where importance sampling has been used.[2] By changing measure we can 'force' the appropriate number of defaults to happen in each iteration with probability one. A comparison between importance sampling and stochastic weight is shown in Figure 5.5 where the constant intensity single name model of Section 5.1.1 was simulated.

There is no point in combining importance sampling and stochastic weights. Stochastic weights are efficient only for discontinuous payoff functions as seen in Section 4.3 Chapter 4, and by using importance sampling the discontinuity is effectively removed since the indicator function $\mathbf{1}_{\{\tau<T\}}$ will always be equal to one.

The drawback with importance sampling is that it is payoff specific, and hence not as flexible as the stochastic weight approach. The importance sampling algorithm also slows up the Monte Carlo run, especially for large baskets. At some point it is more time efficient to use the faster but less accurate method and let it run more iterations.

[2] See Joshi and Kainth (2003).

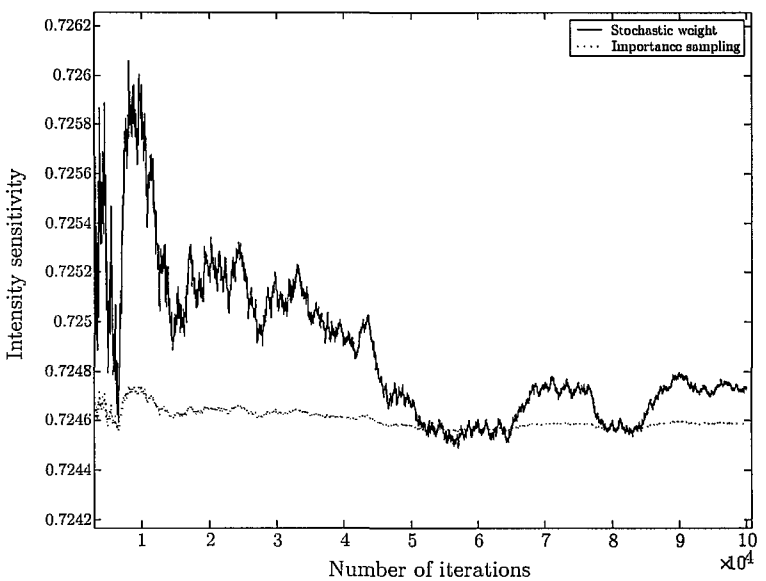


**Figure 5.5:** Monte Carlo simulation of the intensity sensitivity for a one year digital credit default swap. The constant jump intensity was 0.3.

# 6. THE LIKELIHOOD RATIO AND MALLIAVIN METHODS

## 6.1 Preliminaries

### 6.1.1 Likelihood ratio weights

In this chapter we consider general Greeks in that we do not consider sensitivities to a particular parameter of the price process. Following Broadie and Glasserman (1996) we assume that the claim payoff is a function of the random variable $H(\omega, \lambda)$ on a probability space $(\Omega, \mathcal{F}, P)$, where $\omega \in \Omega$ and $\lambda$ denotes the parameter to which sensitivity is of interest. The value of the claim is expressed as

$$v(\lambda) = E[\Phi(H(\omega, \lambda))].$$

If we also assume for a moment that the payoff admits a density, which is differentiable with respect to $\lambda$, we can derive a stochastic sensitivity weight as

$$\begin{aligned}\frac{\partial v(\lambda)}{\partial \lambda} &= \frac{\partial}{\partial \lambda} \int_{\mathbb{R}} \Phi(y) p(y, \lambda) dy = \int_{\mathbb{R}} \Phi(y) \frac{\frac{\partial p(y,\lambda)}{\partial \lambda}}{p(y, \lambda)} p(y, \lambda) dy \\ &= \int_{\mathbb{R}} \Phi(y) \frac{\partial}{\partial \lambda} \left(\log p(y, \lambda)\right) p(y, \lambda) dy = E\left[\Phi(H) \frac{\partial}{\partial \lambda} \left(\log p(H, \lambda)\right)\right].\end{aligned}$$

The stochastic weight for a particular sensitivity is not unique. Indeed, any weights with agreeing conditional expectations with respect to $\sigma(H)$ will do. In this context it should be noticed that the likelihood ratio method always yields a weight that is measurable with respect to $\sigma(H)$, and as shown in Fournié et al. (2001) this is the optimal weight in terms of variance reduction of the Monte–Carlo simulation.

There are striking similarities between Monte–Carlo sensitivity weights derived using Malliavin techniques and the *likelihood ratio* method. In fact, in many

circumstances when the likelihood ratio method is applicable the expressions for the weights turn out to be identical when derived using the two different methods.

The objective of this chapter is to state conditions under which a likelihood ratio weight is attainable, and to shed some light on the connection between the Malliavin weight and the likelihood ratio weight for a particular sensitivity.

### 6.1.2 General Malliavin weight expressions

At the end of the next section we will compare the formal expressions for the likelihood ratio weight to the Malliavin weight. To be able to do so we need to derive an expression for the Malliavin weight in the general setting of this chapter where no particular sensitivity is considered. The derivation presented here is an alternative to Proposition 2.1 in Fournié et al. (2001).

We work on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F})_t, P)$ where the filtration is generated by a separable simple Lévy process $X_t$ as defined in Chapter 3 Definition 3.3.1. The derivative price is

$$v(\lambda) = E[\Phi(X_T)], \tag{6.1}$$

where $\lambda$ is a parameter on which the price process depend, for example its initial value. We denote by $Z_t$ the variation process with respect to $\lambda$ as in Chapter 4.

In general, if $F \in \mathbb{D}^{(0)}$ is satisfying the conditions for the chain rule, Theorem 3.2.4, $G \in L^2(\Omega)$, $h_t \in L^2(\Omega \times \mathbb{R}^+)$ and $f(x)$ continuously differentiable with bounded derivative, then we have

$$\begin{aligned} E[f'(F)G] =& E\left[f'(F)\frac{\int_0^\infty (D_t^{(0)}F)h_t dt}{\int_0^\infty (D_t^{(0)}F)h_t dt}G\right] \\ =& E\left[\int_0^\infty f'(F)D_t^{(0)}F\frac{h_t G}{\int_0^\infty (D_s^{(0)}F)h_s ds}dt\right]. \end{aligned}$$

So by the chain rule and the definition of the Skorohod integral we get

$$E[f'(F)G] = E\left[f(F)\delta^{(0)}\left(\frac{Gh}{\int_0^\infty (D_s^{(0)}F)h_s ds}\right)\right], \tag{6.2}$$

assuming that $\frac{Gh_t}{\int_0^\infty (D_s^{(0)}F)h_s ds} \in \mathrm{Dom}\,\delta^{(0)}$. Using the arguments in the proof of Proposition 3.2 in Fournié et al. (1999), we can use equation (6.2) on general real functions

in $L^2([0,T])$ not necessarily continuously differentiable. With $f$ replaced by $\phi$, $F$ by $X_T$ and $G$ by $Z_T$ we get

$$\frac{\partial v(\lambda)}{\partial \lambda} = E\left[\phi(X_T)\delta^{(0)}\left(\frac{Zh}{\int_0^T (D_s^{(0)}X_T)h_s ds}\right)\right],$$

where $h_t$ is an arbitrary function such that $\frac{Zh_t}{\int_0^\infty (D_s^{(0)}X_T)h_s ds} \in \operatorname{Dom}\delta^{(0)}$. With, for example, $\lambda = X_0$ and $h_t = (D_t^{(0)}X_T)^{-1}$ we recover the well known delta result. We have proved the following proposition.

Proposition 6.1.1: Let $X_t$ be separable and depending on a parameter $\lambda$; the variation process denoted by $Z_t$. If there exist a process $h_t$ such that

$$\frac{Zh_t}{\int_0^\infty (D_s^{(0)}X_T)h_s ds} \in \operatorname{Dom}\delta^{(0)},$$

then

$$\frac{\partial v(\lambda)}{\partial \lambda} = E\left[\phi(X_T)\delta^{(0)}\left(\frac{Zh}{\int_0^T (D_s^{(0)}X_T)h_s ds}\right)\right].$$

## 6.2 Differentiability of a density and the Likelihood ratio method

We continue this chapter with a result stating sufficient conditions for when a random variable has a density, which is differentiable with respect to a general parameter. As a byproduct we will obtain explicit expressions for the density and its derivative, and we conclude with a discussion relating these expressions to the Malliavin sensitivity–weight.

Proposition 6.2.1: Let $F$ be a random variable in $\mathbb{D}^{(0)}$ satisfying the conditions in Theorem 3.2.4 and differentiable with respect to a parameter $\lambda$; the derivative is denoted by $Y$. Suppose that $h_t^{(1)}$, $h_t^{(2)}$ and $h^{(3)}$ are such that

$$\frac{h_t^{(1)}}{\int_0^\infty (D_s^{(0)}F)h_s^{(1)} ds} \in \operatorname{Dom}\delta^{(0)} \quad \text{and} \quad \frac{Gh_t^{(3)}}{\int_0^\infty (D_s^{(0)}F)h_t^{(3)} ds} \in \operatorname{Dom}\delta^{(0)}$$

where $G = \delta^{(0)}\left(\frac{Yh^{(2)}}{\int_0^\infty (D_s^{(0)}F)h_s^{(2)} ds}\right)$. Then the law of $F$ admits a bounded density differentiable with respect to $\lambda$ and with bounded derivative. Furthermore, we have

for the density $p(y,\lambda)$:

$$p(y,\lambda) = E\left[\mathbf{1}_{\{F>y\}}\delta^{(0)}\left(\frac{h^{(1)}}{\int_0^\infty (D_s^{(0)}F)h_s^{(1)}ds}\right)\right]$$

$$\frac{\partial p(y,\lambda)}{\partial\lambda} = E\left[\mathbf{1}_{\{F>y\}}\delta^{(0)}\left(\frac{Gh^{(3)}}{\int_0^\infty (D_s^{(0)}F)h_s^{(3)}ds}\right)\right].$$

*Proof.* By approximating the density by continuously differentiable functions with bounded derivatives, we can use equation (6.2) to get the result. To that end, define $q_\epsilon$ as

$$q_\epsilon(y,\lambda) = E\left[\frac{1}{\sqrt{2\pi\epsilon}}e^{-\frac{(F-y)^2}{2\epsilon}}\right].$$

It is clear that $q_\epsilon$ is a density, and for the corresponding measure with fixed $\lambda$ we get, by letting $\epsilon \to 0$,

$$Q_\epsilon(a \le F \le b) = \int_a^b E\left[\frac{1}{\sqrt{2\pi\epsilon}}e^{-\frac{(F-y)^2}{2\epsilon}}\right]dy = E\left[\int_a^b \frac{1}{\sqrt{2\pi\epsilon}}e^{-\frac{(F-y)^2}{2\epsilon}}dy\right]$$

$$= E\left[N(b,F,\epsilon) - N(a,F,\epsilon)\right] \to E\left[\mathbf{1}_{\{a\le F\le b\}}\right] = P(a \le F \le b),$$

where $N(x,F,\epsilon)$ is the cumulative normal distribution function with mean $F$ and variance $\epsilon$. So if $q_\epsilon$ converges, it converges to $p$. We have by (6.2),

$$\begin{aligned} q_\epsilon(y,\lambda) &= E\left[\frac{1}{\sqrt{2\pi\epsilon}}e^{-\frac{(F-y)^2}{2\epsilon}}\right] \\ &= E\left[\int_{-\infty}^F \frac{1}{\sqrt{2\pi\epsilon}}e^{-\frac{(z-y)^2}{2\epsilon}}dz\,\delta^{(0)}\left(\frac{h^{(1)}}{\int_0^\infty (D_s^{(0)}F)h_s^{(1)}ds}\right)\right] \\ &\to E\left[\mathbf{1}_{\{F>y\}}\delta^{(0)}\left(\frac{h^{(1)}}{\int_0^\infty (D_s^{(0)}F)h_s^{(1)}ds}\right)\right] \text{ as } \epsilon \to 0. \end{aligned}$$

So $q_\epsilon$ converges, at least pointwise, to $p$. To prove that $p(y,\lambda)$ is differentiable in $\lambda$ we need uniform convergence of $\frac{\partial q_\epsilon}{\partial\lambda}$ at least in a neighborhood of every $\lambda$. By

differentiating inside the expectation we get, by using (6.2) twice,

$$\begin{aligned}\frac{\partial q_\epsilon(y,\lambda)}{\partial\lambda} &= E\left[\frac{\partial}{\partial F}\left(\frac{1}{\sqrt{2\pi\epsilon}}e^{-\frac{(F-y)^2}{2\epsilon}}\right)Y\right]\\ &= E\left[\frac{1}{\sqrt{2\pi\epsilon}}e^{-\frac{(F-y)^2}{2\epsilon}}\delta^{(0)}\left(\frac{Yh^{(2)}}{\int_0^\infty (D_s^{(0)}F)h_s^{(2)}ds}\right)\right]\\ &= E\left[\int_{-\infty}^{F}\frac{1}{\sqrt{2\pi\epsilon}}e^{-\frac{(z-y)^2}{2\epsilon}}dz\,\delta^{(0)}\left(\frac{Gh^{(3)}}{\int_0^\infty (D_s^{(0)}F)h_s^{(3)}ds}\right)\right]\\ &\to E\left[\mathbf{1}_{\{F>y\}}\delta^{(0)}\left(\frac{Gh^{(3)}}{\int_0^\infty (D_s^{(0)}F)h_s^{(3)}ds}\right)\right] \text{ as } \epsilon\to 0.\end{aligned}$$

Now, consider $\lambda$ in $[\hat{\lambda}-\eta,\hat{\lambda}+\eta]$. The Schwarz inequality yields that

$$\begin{aligned}&\sup_{\lambda\in[\hat{\lambda}-\eta,\hat{\lambda}+\eta]}\left|E\left[\left(N(F,y,\epsilon)-\mathbf{1}_{\{F>y\}}\right)\delta^{(0)}\left(\frac{Gh^{(3)}}{\int_0^\infty (D_s^{(0)}F)h_s^{(3)}ds}\right)\right]\right|\\ &\qquad\le \sup_{\lambda\in[\hat{\lambda}-\eta,\hat{\lambda}+\eta]}\sqrt{E\left[\left|N(F,y,\epsilon)-\mathbf{1}_{\{F>y\}}\right|^2\right]}\\ &\qquad\qquad\sqrt{E\left[\left(\delta^{(0)}\left(\frac{Gh^{(3)}}{\int_0^\infty (D_s^{(0)}F)h_s^{(3)}ds}\right)\right)^2\right]}.\end{aligned}$$

Since the Skorohod integral is bounded in $L^2$ for any given $\lambda$, we can find a constant such that

$$E\left[\left(\delta^{(0)}\left(\frac{Gh^{(3)}}{\int_0^\infty (D_s^{(0)}F)h_s^{(3)}ds}\right)\right)^2\right] < C_1,$$

for $\lambda\in[\hat{\lambda}-\eta,\hat{\lambda}+\eta]$. Furthermore,

$$\begin{aligned}E\left[\left|N(z,y,\epsilon)-\mathbf{1}_{\{F>y\}}\right|^2\right] &= \int_{\mathbb{R}}\left|N(z,y,\epsilon)-\mathbf{1}_{\{z>y\}}\right|^2 p(z)dz\\ &\le \sqrt{\int_{\mathbb{R}}\left|N(z,y,\epsilon)-\mathbf{1}_{\{z>y\}}\right|^4 dz\,E[p(F,\lambda)]},\end{aligned}$$

and since $p$ is bounded given $\lambda$ we can find a constant such that $E[p(F,\lambda)] < C_2$ for $\lambda\in[\hat{\lambda}-\eta,\hat{\lambda}+\eta]$. Hence, we have uniform convergence (in $\lambda$) of $\frac{\partial q_\epsilon}{\partial\lambda}$ in a neighbourhood of every $\lambda$.

We have shown that

$$\lim_{\epsilon\to 0}\frac{\partial q_\epsilon(y,\lambda)}{\partial\lambda} = \frac{\partial p(y,\lambda)}{\partial\lambda} = E\left[\mathbf{1}_{\{F>y\}}\delta^{(0)}\left(\frac{Gh^{(3)}}{\int_0^\infty (D_s^{(0)}F)h_s^{(3)}ds}\right)\right]. \qquad \square$$

As mentioned in Section 6.1.1 the likelihood ratio method, when applicable, always yields the optimal weight in terms of variance reduction. To investigate this further we derive the formal expression for the likelihood ratio weight.

We assume that the conditions for Proposition 6.2.1 are satisfied, and as in the proof of the same proposition we will approximate the density of $X_T$ as in (6.1) by continuously differentiable functions with bounded derivatives. For the likelihood ratio weight we have

$$\pi_\epsilon(y) := \frac{\partial q_\epsilon(y,\lambda)}{\partial \lambda} \Big/ q_\epsilon(y,\lambda) = \frac{E\left[\frac{1}{\sqrt{2\pi\epsilon}} e^{-\frac{(X_T-y)^2}{2\epsilon}} \delta^{(0)}\left(\frac{Yh^{(2)}}{\int_0^\infty (D_s^{(0)} X_T) h_s^{(2)} ds}\right)\right]}{E\left[\frac{1}{\sqrt{2\pi\epsilon}} e^{-\frac{(X_T-y)^2}{2\epsilon}}\right]}.$$

If we were to evaluate $\pi_\epsilon$ at some random point, say $y = F$, we would have to rewrite the previous equation as

$$\pi_\epsilon(F) = \frac{E\left[\frac{1}{\sqrt{2\pi\epsilon}} e^{-\frac{(X_T-F)^2}{2\epsilon}} \delta^{(0)}\left(\frac{Yh^{(2)}}{\int_0^\infty (D_s^{(0)} X_T) h_s^{(2)} ds}\right) \Big| \sigma(F)\right]}{E\left[\frac{1}{\sqrt{2\pi\epsilon}} e^{-\frac{(X_T-F)^2}{2\epsilon}} \Big| \sigma(F)\right]}.$$

By setting $F = X_T$ and letting $\epsilon \to 0$ we get

$$\frac{\partial v(\lambda)}{\partial \lambda} = E\left[\phi(X_T) E\left[\delta^{(0)}\left(\frac{Yh^{(2)}}{\int_0^T (D_s^{(0)} X_T) h_s^{(2)} ds}\right) \Big| \sigma(X_T)\right]\right]. \tag{6.3}$$

Comparing (6.3) with the expression in Proposition 6.1.1 we find that the likelihood ratio weight is, indeed, the conditional expectation of the Malliavin weight with respect to the sigma algebra generated by $X_T$. Note that if the conditions for Proposition 6.2.1 are satisfied, then we can always find a $h$ as in Proposition 6.1.1 simply by choosing $h$ equal to $h^{(2)}$. The converse is however not true; the conditions for existence of a Malliavin weight are weaker than those for existence of a Likelihood ratio weight.

# REFERENCES


J. Bertoin. *Lévy Processes*. Cambridge University Press, 1996.

M. Broadie and P. Glasserman. Estimating security price derivatives using simulation. *Management Science*, 42(2):269–285, 1996.

M.H.A. Davis and V. Lo. Modelling default correlation in bond portfolios. In C. Alexander, editor, *Mastering Risk Volume 2: Applications*, pages 141–151. Financial Times Prentice Hall, 2001.

L. Debnath and P. Mikusiński. *Introduction to Hilbert Spaces with Applications*. Academic Press, 2nd edition, 1999.

Y. El-Khatib and N. Privault. Computations of greeks in a market with jumps via the malliavin calculus. *Finance and Stochastics*, 8(2):161–179, 2004.

K.D. Elworthy and X.-M. Li. Formulae for the derivatives of heat semigroups. *Journal of Functional Analysis*, 125:252–286, 1994.

E. Fournié, J.-M. Lasry, J. Lebuchoux, and P.-L. Lions. Applications of Malliavin calculus to Monte Carlo methods in finance. II. *Finance and Stochastics*, 5(2):201–236, 2001.

E. Fournié, J.-M. Lasry, J. Lebuchoux, P.-L. Lions, and N. Touzi. Applications of Malliavin calculus to Monte Carlo methods in finance. *Finance and Stochastics*, 3 (4):391–412, 1999.

P. Glasserman and D.D. Yao. Some guidelines and guarantees for common random numbers. *Management Science*, 38(6):884–908, 1992.

M. Joshi and D. Kainth. Rapid and accurate development of prices and greeks for nth to default credit swaps in the Li model. *Working paper*, 2003.

I. Karatzas and S.E. Shreve. *Methods of Mathematical Finance*. Springer–Verlag, 1998.

J. A. León, J. L. Solé, F. Utzet, and J. Vives. On Lévy processes, Malliavin calculus and market models with jumps. *Finance and Stochastics*, 6(2):197–225, 2002.

D.X. Li. On default correlation: a copula function approach. *Working paper 99-07, Risk Metrics Group*, 2000.

D. Nualart. *The Malliavin Calculus and Related Topics*. Springer–Verlag, 1995.

D. Nualart and W. Schoutens. Chaotic and predictable representations for Lévy processes. *Stochastic Processes and their Applications*, 90(1):109–122, 2000.

B. Øksendal. An introduction to Malliavin calculus with applications to economics. *NHH Preprint Series*, 1996.

L.C.G. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales*, volume 2. Cambridge University Press, 2000.

P.J. Schönbucher. *Credit derivatives pricing models. Models, pricing and implementation*. John Wiley & Sons Ltd., 2003.

P.J. Schönbucher and D. Schubert. Copula–dependent default risk in intensity models. *Working paper, Bonn University*, 2000.