Malliavin Monte Carlo Greeks for jump diffusions

Mark H.A. Davis
Department of Mathematics, Imperial College
London SW7 2AZ
mark.davis@imperial.ac.uk

Martin P. Johansson
Imperial College and Citigroup*
Citigroup Centre, 33 Canada Square, 6th Floor, Canary Wharf
London E14 5LB
martin.p.johansson@citigroup.com

July 3, 2005

Abstract
In recent years efficient methods have been developed for calculating derivative price sensitivities using Monte Carlo simulation. Malliavin calculus has been used to transform the simulation problem in the case where the underlying follows a Markov diffusion process. In this work, recent developments in the area of Malliavin calculus for Levy processes are applied and slightly extended. This allows for derivation of similar stochastic weights as in the continuous case for a certain class of jump-diffusion processes.

Keywords: jump process, Lévy process, Monte Carlo estimation, mathematical finance

1 Introduction

For many financial applications Monte Carlo simulation is the favoured pricing tool because of its flexibility. Pricing is, however, only the first step in managing a trade. As the contract start to run it needs to be protected against unfavourable price moves and its risk needs to be properly managed. The price sensitivities with respect to the model parameters—the Greeks—are vital inputs in this context. For calculating the Greeks Monte Carlo simulation leaves much to be desired. The slow convergence, especially for discontinuous payoff functions, is well known, and for this reason less flexible pricing tools are often used.

*The analysis and conclusions set forth are those of the authors. Citigroup is not responsible for any statement or conclusion herein, and opinions or theories presented herein do not necessarily reflect the position of the institution.
The Greeks are calculated as differentials of the derivative price, which can be expressed as an expectation (in risk–neutral measure) of the discounted payoff:

\[ v(x) = E \left[ \phi(X_T) \mid X_0 = x \right]. \]

When this expectation is estimated using Monte Carlo simulation the derivatives can be found by finite difference approximations. If the terms in the finite difference approximation are estimated using independent random number sequences the convergence rate is at best \( n^{-1/3} \), but by using the same random numbers for all terms the convergence rate can be improved to \( n^{-1/2} \), which is the best that can be achieved using Monte Carlo methods.\(^1\)

The theoretical convergence rates for finite difference approximations are not met for discontinuous payoff functions. Fournié et al. (1999) propose a method with faster convergence for the case when the underlying price process is 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

\[ E \left[ \phi(X_T) \pi \mid X_0 = x \right], \]

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 \( X_t = f(X_t^c, X_d^c) \), where \( X_t^c \) is a Markov diffusion for which drift and diffusion coefficients are assumed to have bounded continuous derivatives, whereas \( X_d^c \) is driven by the Poisson process and not dependent on the the initial value \( X_0 \).

This includes the stochastic volatility model

\[
\begin{align*}
    dX_t^{(1)} &= \mu_1 \left( X_t^{(2)}, \ldots, X_t^{(n)} \right) X_t^{(1)} dt + \sigma_1 \left( X_t^{(2)}, \ldots, X_t^{(n)} \right) X_t^{(1)} dW_t \\
    &+ \sum_{k=1}^{m} (\alpha_k - 1) X_t^{(k)} (dN_t^{(k)} - \lambda_k dt), \quad X_0^{(1)} = x_1 \\
    dX_t^{(i)} &= \mu_i \left( X_t^{(2)}, \ldots, X_t^{(n)} \right) dt + \sigma_i \left( X_t^{(2)}, \ldots, X_t^{(n)} \right) dW_t, \\
    X_0^{(i)} &= x_i, \quad i = 2, \ldots, n,
\end{align*}
\]

and the jump diffusion version of the Vasicek model for interest rates

\[ dr_t = a(b - r)dt + \sigma dW_t + \sum_{k=1}^{m} \alpha_k (dN_t^{(k)} - \lambda_k dt), \]

where \( \{W_t, t \geq 0\} \) is a Brownian motion in \( \mathbb{R}^n \), \( \{N^{(k)}, t \geq 0\} \), \( k = 1, \ldots, m \) are independent Poisson processes with intensities \( \lambda_k \) and \( \alpha_k \) are positive deterministic constants.

\(^{1}\)See Glasserman and Yao (1992).
Malliavin-style stochastic calculus for jump processes has been studied since the 1980s. Early works such as Bismut (1983) or Bichteler et al. (1987) exclude the case of fixed jump size, while Carlen and Pardoux (1990) study the pure Poisson case (no Brownian motion component). More recently, Nualart and Schoutens (2000) developed a theory of chaos expansions for functionals of Lévy processes when the Lévy measure satisfies an exponential moment condition. León et al. (2002) have used this as the basis for a Malliavin calculus. A drawback of this approach is that there is no general chain rule. We develop their approach to include a ‘Skorokhod integral’ and show that a restricted form of chain rule is enough to give formulas for Monte Carlo Greeks which are valid for important applications.

Since the first version of this paper was written, at least two other treatments of similar problems have appeared, both dealing with stochastic differential equations driven by Brownian motion and compound Poisson components. Bavouzet and Massaoud (2005) and Bally et al. (2005) reduce the problem to a setting in which only ‘finite-dimensional’ Malliavin calculus is required. El-Khatib and Privault (2004) consider 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. Forster et al. (2005) take a dynamical systems approach which treats the Poisson component as a quasi-deterministic parameter, permitting them to bypass jump process Malliavin calculus altogether. While each of these approaches has advantages for specific applications, no one of them uniformly dominates the others.

In Section 2 some theoretical results are stated and the existing theory extended for later use. In Section 3 the random weights for calculating the Greeks are derived and in Section 4 some numerical examples are presented. Some derivations are left to the Appendix.

2 Malliavin calculus for simple Lévy processes

In this section some of the theory of Malliavin calculus for simple Lévy processes as developed in León et al. (2002) will be revisited. The theory will also be extended with some results we will need in the following sections.

On a filtered probability space \((\Omega, \mathcal{F}, (\mathcal{F}_t), P)\), define a simple Lévy process as given by

\[
X_t = \sigma W_t + \alpha_1 N^{(1)}_t + \cdots + \alpha_m N^{(m)}_t, \quad t \geq 0,
\]

where \(\{W_t, t \geq 0\}\) is a standard Brownian motion and \(\{N^{(j)}_t, t \geq 0\}\), \(j = 1, \ldots, m\), are 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 \(X_t\), which is the same filtration as \(\sigma\{W_t, N^{(j)}_t \mid j = 1, \ldots, m\}\).

The idea in León et al. (2002) is to represent random variables on the probability space \((\Omega, \mathcal{F}, P)\) by iterated integrals, and from there, develop a Malliavin calculus as in the Gaussian case. See Nualart (1995) for a comprehensive treatment of the Malliavin calculus in the Gaussian setting.
We use the notation
\[ G_0(t) = W_t \]
\[ G_j(t) = N_j(t) - \lambda_j t, \quad j = 1, \ldots, m, \]
and define \( L_{n}^{(i_1, \ldots, i_n)}(f) \) to be the iterated integral of some deterministic square integrable function \( f \) with respect to \( G_{i_1}, \ldots, G_{i_n} \):
\[
L_{n}^{(i_1, \ldots, i_n)}(f) = \int_0^\infty \left( \int_0^{t_2} \cdots \left( \int_0^{t_n} f(t_1, \ldots, t_n) \, dG_{i_1}(t_1) \right) \cdots \right) dG_{i_n}(t_n).
\]

It will also be useful to introduce the notation
\[
\Sigma_n = \left\{ (t_1, \ldots, t_n) \in \mathbb{R}_+^n : 0 < t_1 < t_2 < \cdots < t_n \right\},
\]
and
\[
\Sigma_n^{(k)}(t) = \left\{ (t_1, \ldots, t_{k-1}, t_{k+1}, \ldots, t_n) \in \Sigma_{n-1} : \right.
\]
\[
0 < t_1 < \cdots < t_{k-1} < t < t_{k+1} < \cdots < t_n \}
\]

León et al. (2002) formulate the following chaotic representation property, which is a modification of the more general result of Nualart and Schoutens (2000).

**Theorem 2.1 (León, Solé, Utzet and Vives)** Let \( F \in L^2(\Omega, \mathcal{F}, P) \), then \( F \) has a representation of the form
\[
F = E[F] + \sum_{n=1}^{\infty} \sum_{0 \leq i_1, \ldots, i_n \leq m} L_{n}^{(i_1, \ldots, i_n)}(f_{i_1, \ldots, i_n}),
\]
where \( f_{i_1, \ldots, i_n} \in L^2(\Sigma_n) \).

The reason for introducing the chaotic representation is because we want to define a Malliavin derivative similar to the Gaussian analogue. In this discontinuous case the expressions in the definition of the iterated integral and the chaotic representation property are a bit more involved since we are integrating with respect to different martingales instead of only Brownian motion. Therefore, we cannot restrict our attention to symmetric functions.

For the definition of the Malliavin derivative we must first define the spaces of random variables which are differentiable in the \( l \)-th direction (\( l = 0, \ldots, m \)). We use the same notation as in León et al. (2002).

**Definition 2.2 (Differentiable in the \( l \)-th direction)** We say that \( F \) is differentiable in the \( l \)-th direction (\( l = 0, \ldots, m \)) if \( F \in \mathbb{D}^{(l)} \) where
\[
\mathbb{D}^{(l)} = \{ F \in L^2(\Omega), F = E[F] + \sum_{n=1}^{\infty} \sum_{0 \leq i_1, \ldots, i_n \leq m} L_{n}^{(i_1, \ldots, i_n)}(f_{i_1, \ldots, i_n}) : \right.
\]
\[
\sum_{n=1}^{\infty} \sum_{0 \leq i_1, \ldots, i_n \leq m} \sum_{k=1}^{n} 1_{i_k = l} \lambda_{i_1} \ldots \lambda_{i_{k-1}} \lambda_{i_{k+1}} \cdots \lambda_{i_n}
\]
\[
\int_0^{\infty} \| f_{i_1, \ldots, i_n}(\ldots, t, \ldots) 1_{\Sigma_n^{(k)}(t)} \|_{L^2([0, \infty)^{n-k})}^2 \, dt < \infty \},
\]

4
and $\lambda_0 = 1$.

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 $L^2(\Omega)$ and it can also be shown that if we remove the Poisson processes the space $\mathbb{D}^{(0)}$ collapses to the classical Gaussian space $\mathbb{D}^{1,2}$ discussed e.g. in Nualart (1995).

**Definition 2.3 (Malliavin derivative)** For $F \in \mathbb{D}^{(l)}$ such that

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

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

$$D_l^{(l)}F = \sum_{n=1}^{\infty} \sum_{0 \leq i_1, \ldots, i_n \leq m} \sum_{k=1}^{n} 1_{\{i_k = l\}} L_n^{(i_1, \ldots, i_k-1, i_{k+1}, \ldots, i_n)}(f_{i_1, \ldots, i_n}(\ldots, t, \ldots)) \mathbb{1}_{\{\sum_{k=1}^{n} (i_k - 2)_+(t)\}}.$$

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

**Theorem 2.4 (León, Solé, Utzet and Vives)** If $F \in \bigcap_{j=0}^{m} \mathbb{D}^{(j)}$ then

$$F = \mathbb{E}[F] + \int_0^{\infty} p(D^{(0)}_t F) dW_t + \sum_{j=1}^{m} \int_0^{\infty} p(D^{(j)}_t F) d(N_t^{(j)} - \lambda_j t) \quad (2.1)$$

**Theorem 2.5 (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 continuously differentiable function with bounded partial derivatives in the variable $x$, and that $Z \in \mathbb{D}^{(0)}$. Then $F \in \mathbb{D}^{(0)}$ and

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

where $D^{(0)}_t Z$ 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 section. It allows us to calculate the Malliavin derivative in $W$ direction using the classical rules. This, together with the integration by parts formula Theorem 2.6, which will be proved later in this section, 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

$$G_j(t) = W^{(j)}_t, \quad j = 1, \ldots, d$$

$$G_j(t) = (N^{(j-d)}_t - \lambda_{j-d} t), \quad j = d+1, \ldots, d+m.$$

Definition 2.2 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 as $G_j(t) = W^{(j)}_t, j = 1, \ldots, d$ and $G_j(t) = (N^{(j-d)}_t - \lambda_{j-d} t), j = d+1, \ldots, d+m$. The chain rule is given by

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

where $D^{(0)}_t Z$ is the usual Gaussian Malliavin derivative.
motions together in one $d$-dimensional column vector denoted simply by $W_t$. For a random variable $F \in \mathbb{D}^{(d)}$ 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 2.3.

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

$$D_t^{(l)} F = \sum_{n=1}^{\infty} \sum_{0 \leq i_1, \ldots, i_n \leq m} \sum_{k=1}^{n} 1_{\{i_k = l+d\}} L_{n-1, i_1, \ldots, i_{k-1}, i_{k+1}, \ldots, i_n} \left( f_{i_1, \ldots, i_n}(\ldots, t, \ldots) \right)$$

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 denoting the Malliavin derivative with respect to the $i$th Poisson process. Now the above theory still holds true with $d$-dimensional Brownian noise and applying the results we can prove the following theorem.

**Theorem 2.6 (Integration by parts formula)** Let $F \in \bigcap_{j=1}^{m+d} \mathbb{D}^{(j)}$, $u$ be a previsible process in $L^2(\Omega \times \mathbb{R}^+; \mathbb{R}^d)$ and $v$ be a previsible process in $L^2(\Omega \times \mathbb{R}^+)$. Then

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

$$E \left[ F \int_0^\infty v_t (dN_t^{(l)} - \lambda_l dt) \right] = E \left[ \int_0^\infty (D_t^{(l)} F) v_t \lambda_t dt \right],$$

where $l = 1, \ldots, m$ and $*$ denotes transpose.

**Proof.** This follows from the Clark-Ocone representation (2.1) for $F$. \qed

**Remark.** When $u_t$ is a matrix–process the Brownian part of Theorem 2.6 becomes

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

where the integration is done component–wise.

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

$$\left| E \left[ \int_0^\infty (D_t^{(0)} F) u_t \lambda_t dt \right] \right| < c\|F\|_{L^2(\Omega)},$$

for some constant $c$ depending on $u$ and any $F \in \mathbb{D}^{(0)}$. We say that $u$ is Skorohod integrable in the $l$–th direction or $u \in \text{Dom} \delta^{(l)}$. 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

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

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 both the continuous and the discontinuous case the integral is a linear operator due to the linearity of the Malliavin derivative. It is likely that the Skorohod integral in Definition 2.7 has other properties in common with its Gaussian analogue. Another such similarity that will be discussed here is the result presented in Proposition 2.10, which provides a good tool for calculating the Skorohod integral with respect to Brownian motion. For the proof we will need the following Lemmas, which are also interesting in their own right.

**Lemma 2.8** Let $g(t) \in L^2(\mathbb{R}^+)$ and $H \in L^2(\Omega)$ with finite chaos expansion

$$H = E[H] + \sum_{n=1}^N \sum_{0 \leq i_1, \ldots, i_n \leq m} F^{(i_1, \ldots, i_n)}(h_{i_1, \ldots, i_n}).$$

Then $Hg \in \text{Dom} \delta^{(l)}$ for $l = 1, \ldots, m$.

**Proof.** It is enough to consider the case when $H$ is a single iterated integral:

$$H = L^{(j_1, \ldots, j_q)}(h_{j_1, \ldots, j_q}).$$

The general result follows from linearity. For any $F \in \mathbb{D}^{(l)}$ we have

$$E \left[ \int_0^\infty (D_t^{(l)} F) g(t) \lambda_t dt \right]$$

$$= \lambda_l \int_0^\infty E \left[ \left( \sum_{n=1}^\infty \sum_{0 \leq i_1, \ldots, i_n \leq m} \sum_{k=1}^n 1_{\{i_k = l\}} L^{(i_1, \ldots, i_n)}(f_{i_1, \ldots, i_n}(\ldots, t, \ldots) 1_{\{i_k = l\}}) \right) \right] \times (L^{(j_1, \ldots, j_q)}(h_{j_1, \ldots, j_q})) g(t) dt.$$
so we get (note how the infinite sum for $F$ simplifies due to the above fact)

$$
E\left[\int_0^\infty (D^{(l)}_t F) H g(t) \lambda dt\right] = \\
\lambda_t \int_0^\infty \left( \sum_{0 \leq i_1, \ldots, i_{q+1} \leq m} \sum_{k=1}^{q+1} \mathbf{1}_{\{i_k=t\}} \mathbf{1}_{\{(i_1, \ldots, i_{k-1}, i_{k+1}, \ldots, i_{q+1})=(j_1, \ldots, j_q)\}} \right) \\
\lambda_{i_1} \cdots \lambda_{i_{k-1}} \lambda_{i_k+1} \cdots \lambda_{i_{q+1}} \int_0^{t_{k+1}} \cdots \int_0^{t_{k-1}} \cdots \int_0^{t_2} \cdots \int_0^{t_1} h_{j_1, \ldots, j_q}(t_1, \ldots, t_{k-1}, t_{k+1}, \ldots, t_{q+1}) g(t_1) dt_1 \cdots dt_{k-1} dt_{k+1} \cdots dt_{q+1} \right) dt.
$$

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

$$
E\left[\int_0^\infty (D^{(l)}_t F) H g(t) \lambda dt\right] = E \left[ \left( E[F] + \sum_{n=1}^\infty \sum_{0 \leq i_1, \ldots, i_n \leq m} \int_0^\infty \int_0^{t_n-} \cdots \int_0^{t_2-} f_{i_1, \ldots, i_n}(t_1, \ldots, t_n) dG_{i_1}(t_1) \cdots dG_{i_n-1}(t_{n-1}) dG_{i_n}(t_n) \right) \right. \\
\times \left. \left( \sum_{k=1}^{q+1} \int_0^{t_{k+2}} \cdots \int_0^{t_{k+1}} \cdots \int_0^{t_k-} h_{j_1, \ldots, j_q}(t_1, \ldots, t_{k-1}, t_{k+1}, \ldots, t_{q+1}) g(t_k) dG_{j_1}(t_1) \cdots dG_{j_{k-1}}(t_{k-1}) dG_{j_k}(t_k) \cdots dG_{j_q}(t_{q+1}) \right) \right].
$$

An application of Schwarz inequality concludes the proof. \hfill \square

**Corollary 2.9** The Skorohod integral is a densely defined operator in $L^2(\Omega \times \mathbb{R}^+)$.

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

**Remark.** The Skorohod integral of $H g$ as in the above Lemma is given by

$$
\delta^{(l)}(H g) = \int_0^\infty E[H] g(t_1) dG_{i_1}(t_1) + \sum_{n=1}^N \sum_{0 \leq i_1, \ldots, i_n \leq m} \sum_{k=1}^{n+1} \int_0^\infty \\
\cdots \int_0^{t_{k+2}} \cdots \int_0^{t_{k+1}} \cdots \int_0^{t_k-} h_{i_1, \ldots, i_n}(t_1, \ldots, t_{k-1}, t_{k+1}, \ldots, t_{n+1}) g(t_k) dG_{i_1}(t_1) \cdots dG_{i_{k-1}}(t_{k-1}) dG_{i_k}(t_k) \cdots dG_{i_n}(t_{n+1}).
$$

(2.2)

We see that Skorohod integration increases the order of the chaos expansion with a $dG$ integral mixed in at every possible position. This is in line with Skorohod integration in the Gaussian case.
Proposition 2.10 For $F u$ in $L^2(\Omega \times \mathbb{R}^+)$, where $F \in \mathcal{D}^{(0)}$ and $u_t$ previsible, we have
\[
\delta^{(0)}(F u) = F \int_0^\infty u_t dW_t - \int_0^\infty (D_{t}^{(0)} F) u_t dt.
\] (2.3)
in the sense that $F u \in \text{Dom} \delta^{(0)}$ if and only if the right hand side of (2.3) is in $L^2(\Omega)$.

Proof. It is not very surprising that when $u_t$ is stochastic but previsible, the expression for the Skorohod integral of $F u$ is very similar to the expression (2.2). In Appendix A we show that
\[
\delta^{(i)}(F u) = \int_0^\infty E[F] u_t dG^{(i)}(t_1) + \sum_{n=1}^\infty \sum_{0 \leq t_1, \ldots, t_n \leq m} \sum_{k=1}^{n+1} \int_0^\infty 
\cdots \int_0^{t_{k-1}-} \int_0^{t_k-} \cdots \int_0^{t_{k+1}-} \int_0^{t_k} \int_0^{t_{k+1}} \cdots \int_0^{t_{n+1}-} f_{i_1, \ldots, i_n}(t_1, \ldots, t_{k-1}, t_k+1, \ldots, t_{n+1}) dG_{i_1}(t_1) \\
\cdots dG_{i_{k-1}}(t_{k-1}) u_{t_k} dG_{i_k}(t_k) dG_{i_k}(t_{k+1}) \ldots dG_{i_n}(t_{n+1}).
\]
We will now focus on the expression $F \int_0^\infty u_t dG_0(t)$, 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
\[
\int_0^\infty \int_0^{t_1-} \cdots \int_0^{t_2-} f_{i_1, \ldots, i_n}(t_1, \ldots, t_n) dG_{i_1}(t_1) \cdots dG_{i_{n-1}}(t_{n-1}) dG_{i_n}(t_n) \int_0^\infty u_t dG_0(t),
\]
and the strategy is to prove that
\[
\int_0^\infty \int_0^{t_1-} \cdots \int_0^{t_2-} f_{i_1, \ldots, i_n}(t_1, \ldots, t_n) dG_{i_1}(t_1) \cdots dG_{i_{n-1}}(t_{n-1}) dG_{i_n}(t_n) \int_0^\infty u_t dG_0(t)
\]
\[= \sum_{k=1}^{n+1} \int_0^\infty \cdots \int_0^{t_{k+1}-} \int_0^{t_k-} \cdots \int_0^{t_{k-1}-} \int_0^{t_k} \int_0^{t_{k+1}} \cdots \int_0^{t_{n+1}-} f_{i_1, \ldots, i_n}(t_1, \ldots, t_{k-1}, t_k+1, \ldots, t_{n+1})
\]
\[\cdots dG_{i_1}(t_1) \cdots dG_{i_{k-1}}(t_{k-1}) u_{t_k} dG_0(t_k) dG_{i_k}(t_{k+1}) \cdots dG_{i_n}(t_{n+1}) + \sum_{k=1}^{n} \int_0^\infty \cdots \int_0^{t_k-} f_{i_1, \ldots, i_n}(t_1, \ldots, t_n) dG_{i_1}(t_1) \cdots u_{t_k} d[G_{i_k}, G_0]_{t_k} \ldots dG_{i_n}(t_n).
\] (2.4)

The first sum of the right hand side of (2.4) is the contribution from the chaos expansion of $\delta^{(i)}(F u)$ and the second sum is the contribution from the chaos expansion of $\int_0^\infty (D_{t}^{(0)} F) u_t dt$, establishing (2.3). Equation (2.4) can be shown by induction using Itô’s formula. \hfill \Box

Remark. In the case when $F$ is a $d$-dimensional random column–vector and $u_t$ is a $(d \times d)$ matrix–process Proposition 2.10 translates to
\[
\delta^{(0)}(u F) = F^* \int_0^\infty u_t dW_t - \int_0^\infty \text{Tr} \left( (D_{t}^{(0)} F) u_t \right) dt,
\]
with the convention that the Itô integral for a matrix process is a column–vector.
3 Monte Carlo Greeks

In this section we will apply the results from the previous section to calculate the Greeks for a certain class of price processes. We focus our attention on jump diffusion processes in $\mathbb{R}^d$ for which we make a particular assumption.

**Assumption 3.1 (Separability)** Assume that $b$, $\sigma$ and $\alpha$ are continuously differentiable functions with bounded derivatives and consider Markov jump diffusions, $X_t \in \bigcap_{j=1}^{d+m} \mathbb{P}^{(j)}$, 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

$$X_t = f(X^c_t, X^d_t), \quad X^c_0 = x.$$

Here $X^c_t$ satisfies a stochastic differential equation

$$dX^c_t = b_c(X^c_t)dt + \sigma_c(X^c_t)dW_t, \quad X^c_0 = x,$$

with smooth coefficients $b_c$ and $\sigma_c$ while $X^d_t$ is adapted to the natural filtration $\mathcal{F}_t^N$ of the Poisson processes $(N_{t}^{(1)}, \ldots, N_{t}^{(m)})$. In particular, $X^d_t$ does not depend on $x$. We say that the jump diffusion process is separable.

An important property of the class of jump diffusions that satisfies Assumption 3.1 is that they are differentiable with respect to Brownian motion and the Malliavin derivative can be calculated using the chain rule given in Theorem 2.5. This, together with the integration by parts formula (Theorem 2.6), 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).

Examples of separable pricing processes are the stochastic volatility model

$$dX^{(1)}_t = b_1 \left( X^{(2)}_t, \ldots, X^{(d)}_t \right) dt + \sigma_1 \left( X^{(2)}_t, \ldots, X^{(d)}_t \right) dW_t$$

$$+ \sum_{k=1}^{m} (\hat{\alpha}_k - 1) X^{(1)}_{t-} (dN^{(k)}_t - \lambda_k dt), \quad X^{(1)}_0 = x_1,$$

and the jump–diffusion version of the Vasicek model for interest rates

$$dr_t = a(b - r_{t-})dt + \sigma dr_t + \sum_{k=1}^{m} \hat{\alpha}_k (dN_t^{(k)} - \lambda_k dt),$$

where $\hat{\alpha}_k$, $k = 1, \ldots, m$ are deterministic constants. The claim that the jump–diffusion version of the Vasicek model is separable can be proved by
solving the short rate SDE analytically, and the claim regarding the stochastic volatility model will be proved in section 4.

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

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

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 $Y_t = \nabla_x X_t^c$, and it is also known that, since $b_c$ and $\sigma_c$ are continuously differentiable functions with bounded derivatives and since $X_t^c$ is continuous, the Malliavin derivative of $X_t^c$ can be written as

$$D_t^{(0)} X_t^c = Y_t Y_s^{-1} \sigma_c(X_s^c) 1_{s \leq t}. \tag{3.2}$$

We define the payoff function

$$\phi = \phi(X_{t_1}, \ldots, X_{t_n}), \tag{3.3}$$

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

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

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

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

We will also need the following technical lemma:

**Lemma 3.2** Let $g : \mathbb{R} \to \mathbb{R}$ be a continuously differentiable function with bounded derivative and let $G$ be a separable random variable in $\bigcap_{j=1}^{d+m} D^{(j)}$. Then

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

**Proof.** That $g(G) \in \bigcap_{j=1}^{d+m} D^{(j)}$ is given by the chain rule (Theorem 2.5). For $k \in [d+1, d+m]$ we can write the Malliavin derivative as

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

where $\omega + \theta_t^{(k)}$ means “an extra jump in the $k$th Poisson process at time $t$”. We have,

$$E \left[ \int_0^\infty \left( D_t^{(k)} g(G) \right)^2 dt \right] = E \left[ \int_0^\infty \left( g(G(\omega + \theta_t^{(k)}) - g(G(\omega) \right)^2 dt \right]$$

$$\leq E \left[ M^2 \int_0^\infty \left( G(\omega + \theta_t^{(k)}) - G(\omega) \right)^2 dt \right] < \infty,$$

where $M$ is a bound on the derivative of $g$. \hfill $\square$

---

11

---

2See Nualart (1995) section 2.3.1.

3See León et al. (2002) Proposition 2.4(b).
3.1 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

\[ dX^\epsilon_t = \left( b(X^\epsilon_t) + \epsilon \gamma(X^\epsilon_{t-}) \right) dt + \sigma(X^\epsilon_t) dW_t + \sum_{k=1}^{m} \alpha_k(X^\epsilon_{t-}) (dX^{(k)}_t - \lambda_k dt) \]

\[ X^\epsilon_0 = x, \]

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

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

we have

\[ \frac{\partial}{\partial \epsilon} v^\epsilon(x) \bigg|_{\epsilon=0} = E \left[ \phi(X_{t_1}, \ldots, X_{t_n}) \int_0^T (\sigma^{-1}(X_{t-}) \gamma(X_{t-}))^* dW_t \right] \]

**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 for more general path-dependent claims $\phi = \phi(X(\cdot))$ and also for processes which are not separable as in Assumption 3.1.

3.2 Variations in the initial condition

This is where we for the first time will make full use of the theory developed in Section 2. We rely heavily on the chain rule and the integration by parts formula, which allow us to repeat the steps in Fournié et al. (1999).

First define the set $\Gamma$, of square integrable functions 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 (3.3).

**Proposition 3.4** Assume that the diffusion matrix $\sigma_c$ is uniformly elliptic. Then for any $a(t) \in \Gamma$

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

**Proof.** First assume that $\phi$ is continuously differentiable with bounded gradient. In this case it is possible to “differentiate inside the expectation”.\footnote{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 $\frac{dX^\epsilon}{dt}$ and the fact that the map $x \mapsto X_t$ is a.s. continuous even in the presence of jumps.} 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 (Assumption 3.1) $X^d_t$ does not depend on $x$ so that

$$\nabla v(x) = E \left[ \sum_{i=1}^{n} \nabla_i \phi(X_{t_1}, \ldots, X_{t_n}) \frac{\partial X_{t_i}}{\partial x} \right].$$

We want to rewrite $Y_{t_i}$ in terms of the Malliavin derivative for $X_{t_i}$ using (3.2), but to do that we must also use the chain rule (Theorem 2.5). We have

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

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

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

Inserting in (3.4) 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 3.2 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 can thus use the chain rule again to get

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

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 (Theorem 2.6) to get

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

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 3.4 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.

### 3.3 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 2.10 to interpret the result.

As in Section 3.1 we need to define the perturbed process with respect to the property under investigation; in this case the diffusion coefficient:

\[
\begin{align*}
    dX_t^\epsilon &= b(X_{t-}^\epsilon)dt + (\sigma_{t-}^\epsilon + \epsilon\gamma(X_{t-}^\epsilon))dW_t + \sum_{k=1}^m \alpha_k(X_{t-}^\epsilon)(dN_t^{(k)} - \lambda_kdt) \\
    X_0^\epsilon &= x,
\end{align*}
\]

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 \( \epsilon \)

\[
\begin{align*}
    dZ_t^\epsilon &= b'(X_{t-}^\epsilon)Z_{t-}^\epsilon dt + \sum_{i=1}^d (\sigma_i'(X_{t-}^\epsilon) + \epsilon\gamma_i'(X_{t-}^\epsilon))Z_{t-}^i dW_t^i + \gamma(X_{t-}^\epsilon)dW_t \\
    Z_0^\epsilon &= 0_n,
\end{align*}
\]

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

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

In this setting Proposition 3.5 is the analog of Proposition 3.3.

**Proposition 3.5** Assume that the diffusion matrix \( \sigma_c \) is uniformly elliptic and that for \( \beta_i = \left( \frac{\partial X_t^\epsilon}{\partial X_t^\epsilon} \right)^{-1} \) \( Z_i \), \( i = 1, \ldots, n \) we have \( \sigma_c^{-1}(X^\epsilon)Y \beta \in \text{Dom} \delta^{(0)} \). 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 \( a \in \Gamma_n \)

\[
\frac{\partial}{\partial \epsilon} v^\epsilon \bigg|_{\epsilon=0} = E \left[ \phi(X_{t_1}^\epsilon, \ldots, X_{t_n}^\epsilon) \delta^{(0)}(\sigma_c^{-1}(X^\epsilon)Y \hat{\beta}) \right],
\]

where

\[
\hat{\beta}_i = \sum_{i=1}^n a(t) (\beta_{t_i} - \beta_{t_{i-1}}) 1_{\{t \in [t_{i-1}, t_i)\}},
\]

14
for $t_0 = 0$. Furthermore, if $\beta \in \mathbb{D}^{(0)}$ the Skorohod integral can be calculated according to

$$
\delta^{(0)}\left(\sigma^{-1}_c(X^c)Y\tilde{\beta}\right) = \sum_{i=1}^{n} \left\{ \beta^*_{t_i} \int_{t_{i-1}}^{t_i} a(t) \left( \sigma^{-1}_c(X^c_t)Y_t \right)^* dW_t \\
- \int_{t_{i-1}}^{t_i} a(t) \text{Tr} \left( (D^{(0)}_t \beta_t) \sigma^{-1}_c(X^c_t)Y_t \right) dt \\
- \int_{t_{i-1}}^{t_i} a(t) \left( \sigma^{-1}_c(X^c_t)Y_t \beta_{t_{i-1}} \right)^* dW_t \right\}
$$

Proof. As in the proof of Proposition 3.4 we may assume $\phi$ to be continuously differentiable with bounded gradient, and we can take the differential under the expectation to get

$$
\frac{\partial}{\partial \epsilon} \nu' (x) = E \left[ \sum_{i=1}^{n} \nabla_i \phi(X^c_{t_i}, \ldots, X^c_{t_n}) Z_t \right]. \quad (3.5)
$$

Next we define $\beta_t = \left( \frac{\partial X^c_t}{\partial X_t} Y_t \right)^{-1} Z_t = \left( \frac{\partial X^c_t}{\partial X_t} Y_t \right)^{-1} Z_{t=0}$, and proceed as in the proof of Proposition 3.4. Because of (3.2) and since $a \in \Gamma_n$ we can write

$$
\int_0^T \frac{\partial X^c_t}{\partial X^c_{t_i}} (D^{(0)}_t X^c_{t_i}) \sigma^{-1}_c(X^c_t)Y_t \tilde{\beta}_t dt = \int_0^{t_i} \frac{\partial X^c_t}{\partial X^c_{t_i}} Y_t \tilde{\beta}_t dt \\
= \frac{\partial X^c_t}{\partial X^c_{t_i}} Y_t \sum_{j=1}^{i} \int_{t_j}^{t_{j-1}} a(t)(\beta_{t_j} - \beta_{t_{j-1}}) dt \\
= \frac{\partial X^c_t}{\partial X^c_{t_i}} Y_t \beta_{t_i} = Z_{t_i}.
$$

Inserting into (3.5) we get, by using the chain rule and Lemma 3.2,

$$
\left. \frac{\partial}{\partial \epsilon} \nu'(x) \right|_{\epsilon=0} = E \left[ \int_0^T \sum_{i=1}^{n} \nabla_i \phi(X^c_{t_i}, \ldots, X^c_{t_n}) \frac{\partial X^c_t}{\partial X^c_{t_i}} (D^{(0)}_t X^c_{t_i}) \sigma^{-1}_c(X^c_t)Y_t \tilde{\beta}_t dt \right] \\
= E \left[ \int_0^T (D^{(0)}_t \phi(X^c_{t_i}, \ldots, X^c_{t_n})) \sigma^{-1}_c(X^c_t)Y_t \tilde{\beta}_t dt \right] \\
= E \left[ \phi(X^c_{t_i}, \ldots, X^c_{t_n}) \delta^{(0)}(\sigma^{-1}_c(X^c_t)Y\tilde{\beta}) \right],
$$

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

$$
\left. \frac{\partial}{\partial \epsilon} \nu'(x) \right|_{\epsilon=0} = E \left[ \phi(X^c_{t_i}, \ldots, X^c_{t_n}) \delta^{(0)}(\sigma^{-1}_c(X^c_t)Y\tilde{\beta}) \right],
$$

and Proposition 2.10 can be used to calculate the Skorohod integral. \hfill \Box

Remark. The assumption of $\sigma^{-1}_c(X^c)Y\tilde{\beta} \in \text{Dom} \delta^{(0)}$ 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.
3.4 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 3.6** Suppose that the diffusion matrix $\sigma$ is uniformly elliptic. For $j = 1, \ldots, m$ we have

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

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

$$dX_t^\epsilon = b(X_t^\epsilon) dt + \sigma(X_t^\epsilon) dW_t + \alpha_1(X_t^\epsilon) (dN_t^1 - (\lambda_1 + \epsilon) dt) + \sum_{k=2}^m \alpha_k(X_t^\epsilon) (dN_t^k - \lambda_k dt), \quad X_0^\epsilon = x,$$

where $\epsilon$ is a positive deterministic parameter and $N_t^1$ is a Poisson process with intensity $\lambda_1 + \epsilon$. As in the proof of Proposition 3.3 we note that by changing measure we can write

$$E\left[\phi(X_t^\epsilon, \ldots, X_n^\epsilon)\right] = E\left[M_T^\epsilon \phi(X_t, \ldots, X_n)\right],$$

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,

$$M_T^W = \exp\left(-\epsilon \int_0^T (\sigma^{-1}(X_{t-}) \alpha_1(X_{t-}))^* dW_t - \frac{\epsilon^2}{2} \int_0^T \|\sigma^{-1}(X_{t-}) \alpha_1(X_{t-})\|^2 ds\right)$$

$$M_T^N = \left(1 + \frac{\epsilon}{\lambda_1}\right)^{N_t^{(1)}} \exp(-\epsilon t).$$

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

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

□

**Remark.** As was the case for variations in the drift coefficient the above result holds true even for more general path-dependent claims and non-separable price dynamics.
3.5 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 3.5—the diffusion coefficient result. To that end, consider the perturbed process

\[
\begin{align*}
    dX_t^\epsilon &= b(X_t^\epsilon)dt + \sigma(X_t^\epsilon)dW_t + (\alpha_1(X_t^\epsilon) + \epsilon \gamma(X_t^\epsilon))(dN_t^{(1)} - \lambda_1 dt) \\
    &+ \sum_{k=2}^{m} \alpha_k(X_t^\epsilon)(dN_t^{(k)} - \lambda_k dt), \quad X_0^\epsilon = x,
\end{align*}
\]

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

\[
\begin{align*}
    dZ_t^\epsilon &= b'(X_t^\epsilon)Z_t^\epsilon dt + \sum_{i=1}^{d} \sigma_i'(X_t^\epsilon)Z_t^\epsilon dW_t^{(i)} \\
    &+ (\alpha_1'(X_t^\epsilon) + \epsilon \gamma'(X_t^\epsilon))Z_{t-}^\epsilon (dN_t^{(1)} - \lambda_1 dt) \\
    &+ \sum_{k=2}^{m} \alpha_k'(X_t^\epsilon)Z_{t-}^\epsilon (dN_t^{(k)} - \lambda_k dt) + \gamma(X_t^\epsilon)(dN_t^{(1)} - \lambda_1 dt), \quad Z_0^\epsilon = 0.
\end{align*}
\]

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 3.5.

**Proposition 3.7** Assume that the diffusion matrix \( \sigma_c \) is uniformly elliptic and that for \( \beta_t = \left( \frac{\partial X}{\partial Y_{t_i}} Y_{t_i} \right)^{-1} Z_{t_i}, i = 1, \ldots, n \) we have \( \sigma_c^{-1}(X^c)Y \beta \in \text{Dom} \delta^{(0)} \). For \( v^\epsilon(x) \) defined as

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

we have for any \( a \in \Gamma_n \)

\[
\frac{\partial}{\partial \epsilon} v^\epsilon \bigg|_{\epsilon=0} = \mathbb{E} \left[ \phi(X_{t_1}, \ldots, X_{t_n}) \delta^{(0)} \left( \sigma_c^{-1}(X^c)Y \beta \right) \right],
\]

where

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

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

\[
\delta^{(0)}(\sigma_c^{-1}(X^c)Y, \tilde{\beta}) = \sum_{i=1}^{n} \left\{ \beta_{t_i}^* \int_{t_{i-1}}^{t_i} a(t) \left( \sigma_c^{-1}(X^c)_t Y_t \right)^* dW_t \\
    - \int_{t_{i-1}}^{t_i} a(t) \text{Tr} \left( (D_{t_i}^{(0)} \beta_{t_i}) \sigma_c^{-1}(X^c)_t Y_t \right) dt \\
    - \int_{t_{i-1}}^{t_i} a(t) \left( \sigma_c^{-1}(X^c)_t Y_t \beta_{t_{i-1}} \right)^* dW_t \right\}
\]

17
Proof. Exactly the same proof as for Proposition 3.5 with redefined $\beta_t$ and $Z_t$.

Remark. As long as $\hat{\alpha}_k \neq 0$ for $k = 1, \ldots, m$ we have $\beta_{i_t} \in \bigcap_{j=1}^{d+m} \mathbb{D}^{(j)}$ for the stochastic volatility and interest rate models as presented earlier. See the Appendix for the explicit calculations in the case of the Heston stochastic volatility model.

4 Examples

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

We start by proving that the separability assumption (Assumption 3.1) holds for the stochastic volatility model

\[
\begin{align*}
\frac{dX_t}{X_t} &= b_1 \left( X_{t-}^{(1)}, \ldots, X_{t-}^{(d)} \right) dt + \sigma_1 \left( X_{t-}^{(1)}, \ldots, X_{t-}^{(d)} \right) \frac{dW_t}{X_t}, \\
&\quad + \sum_{k=1}^{m} \left( \alpha_k - 1 \right) \frac{dN_k}{X_t}, \\
\end{align*}
\]

(4.1)

where $b_1$ and $\sigma_1$ 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

\[
\begin{align*}
\frac{dX_t^c}{X_t^c} &= \left( b_1 \left( X_{t-}^{c(1)}, \ldots, X_{t-}^{c(d)} \right) + \sum_{k=1}^{m} \lambda_k \left( 1 - \alpha_k \right) \right) \frac{dW_t}{X_t^c}, \\
&\quad + \sigma_1 \left( X_{t-}^{c(1)}, \ldots, X_{t-}^{c(d)} \right) \frac{dW_t}{X_t^c}, \\
X_0^c &= x_0, \\
\end{align*}
\]

(4.2)

and the $d$-dimensional discontinuous process

\[
\begin{align*}
\hat{X}_t^{(1)} &= \alpha_1^{N_1(1)} \cdots \alpha_m^{N_m(1)} X_t^{c(1)}, \\
\hat{X}_t^{(i)} &= X_t^{c(i)}, \\
\end{align*}
\]

(4.3)

It is clear that $(X^{(1)}, \ldots, X^{(d)})$ and $(\hat{X}^{(2)}, \ldots, \hat{X}^{(d)})$ are indistinguishable due to pathwise uniqueness of the solutions to the SDEs, and by applying the Itô formula to $\hat{X}_t^{(1)}$ we find that $\hat{X}_t^{(1)} = X_t^{c(1)}$ satisfies (4.2).

We summarise the result in the following Lemma.

Lemma 4.1 Let $X_t$, $X_t^c$ and $\hat{X}_t$ be as defined in (4.1), (4.2) and (4.3) respectively. Then $\hat{X}_t = X_t$ a.s.

Next we consider a jump–diffusion version of the Black–Scholes model. Let the price process follow

\[
\begin{align*}
\frac{dX_t}{X_t} &= \mu X_t dt + \sigma X_t dW_t + (\alpha - 1) X_t (dN_t - \lambda dt), \\
X_0 &= x,
\end{align*}
\]

18
where $\mu$, $\sigma$ and $\alpha$ are constants and the intensity of the Poisson process is $\lambda$. The price process is clearly separable since it is a special case of the stochastic volatility model defined in (4.1). In particular we see that the continuous price process contribution $X^c_t$ is identical to the Black–Scholes price process with modified drift. By straightforward derivations we conclude that for a contingent claim
\[ v = E \left[ e^{-rT} \phi(X_T) \right], \] (4.4)
we have
\[
\begin{align*}
\text{Rho}_{BS} &= \frac{\partial v}{\partial \lambda} = E \left[ e^{-rT} \phi(X_T) \frac{W_T}{\sigma} \right] - E \left[ T e^{-rT} \phi(X_T) \right] \\
\text{Delta}_{BS} &= \frac{\partial v}{\partial x} = E \left[ e^{-rT} \phi(X_T) \left( \frac{W_T}{\sigma} - w_T - \frac{1}{\sigma} \right) \right] \\
\text{Gamma}_{BS} &= \frac{\partial^2 v}{\partial x^2} = E \left[ e^{-rT} \phi(X_T) \left( \frac{W_T^2}{\sigma^2} - \frac{1}{\sigma^2} \right) \right] \\
\text{Vega}_{BS} &= \frac{\partial v}{\partial \sigma} = E \left[ e^{-rT} \phi(X_T) \left( \frac{W_T^2}{\sigma} - W_T - \frac{1}{\sigma} \right) \right] \\
\text{Lambda}_{BS} &= \frac{\partial v}{\partial \lambda} = E \left[ e^{-rT} \phi(X_T) \left( \frac{N_T - \lambda T}{\lambda} - \frac{(\alpha - 1) W_T}{\sigma} \right) \right] \\
\text{Alpha}_{BS} &= \frac{\partial v}{\partial \alpha} = E \left[ e^{-rT} \phi(X_T) \left( \frac{N_T - \lambda T}{\alpha} \right) \frac{W_T}{\sigma} \right].
\end{align*}
\]
Rho might need some further explanation. Rho is defined as the sensitivity of the claim (4.4) with respect to interest rate, and here the interest rate appears both in the discount factor and in the drift. Analogous to section 3.1 we define the Rho as the sensitivity of the claim with respect to the factor $\epsilon$, as in the perturbed interest rate $r + \epsilon r$, in the point $\epsilon = 0$. Two examples involving Delta are shown in Figure 1 and Figure 2.

The jump–diffusion version of the Heston model
\[
\begin{align*}
dX^{(1)}_t &= r X^{(1)}_t dt + \sqrt{X^{(2)}_t} X^{(1)}_t dW^{(1)}_t + (\alpha - 1) X^{(1)}_t (dN_t - \lambda dt), \quad X^{(1)}_0 = x_1 \\
dX^{(2)}_t &= \kappa (\theta - X^{(2)}_t) dt + \sigma \sqrt{X^{(2)}_t} dW^{(2)}_t, \quad X^{(2)}_0 = x_2,
\end{align*}
\]
with $E[dW^{(1)}_t dW^{(2)}_t] = \rho dt$ is also a special case of (4.1) and the separability assumption therefore holds. $X^{(1)}_t$ is the process for the security on which the contingent claim is written, and $X^{(2)}_t$ is a process governing the volatility. The price of the claim in this setting is
\[ v = E \left[ e^{-rT} \phi(X^{(1)}_T) \right], \]
and the continuous process $X^c_t$ is identical to the price process in the original Heston model, but again with modified drift. The stochastic weights for calculating the Greeks for the jump–diffusion version are found to be:
Figure 1: Monte Carlo simulation of 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 $\lambda = 0.1$.

\[
\text{Rho}_H = E\left[ e^{-rT} \phi(X_T^{(1)}) \left( \int_0^T \frac{dW_t^{(1)}}{\sqrt{X_t^{(2)}}} - \frac{\rho}{\sqrt{1-\rho^2}} \int_0^T \frac{dW_t^{(2)}}{\sqrt{X_t^{(2)}}} \right) \right] \\
- E\left[ T e^{-rT} \phi(X_T^{(1)}) \right]
\]

\[
\text{Delta}_H = E\left[ e^{-rT} \phi(X_T^{(1)}) \frac{1}{\pi T} \left( \int_0^T \frac{dW_t^{(1)}}{\sqrt{X_t^{(2)}}} - \frac{\rho}{\sqrt{1-\rho^2}} \int_0^T \frac{dW_t^{(2)}}{\sqrt{X_t^{(2)}}} \right) \right]
\]

\[
\text{Gamma}_H = E\left[ e^{-rT} \phi(X_T^{(1)}) \frac{1}{\pi T^2} \left( \left( \int_0^T \frac{dW_t^{(1)}}{\sqrt{X_t^{(2)}}} - \frac{\rho}{\sqrt{1-\rho^2}} \int_0^T \frac{dW_t^{(2)}}{\sqrt{X_t^{(2)}}} \right)^2 \\
- \frac{1}{1-\rho^2} \int_0^T \frac{dt}{X_t^{(2)}} - T \left( \int_0^T \frac{dW_t^{(1)}}{\sqrt{X_t^{(2)}}} - \frac{\rho}{\sqrt{1-\rho^2}} \int_0^T \frac{dW_t^{(2)}}{\sqrt{X_t^{(2)}}} \right) \right) \right]
\]

\[
\text{Vega}_H = E\left[ e^{-rT} \phi(X_T^{(1)}) \frac{1}{\pi T} \left( \int_0^T \frac{Y_t^{(1,2)}}{X_t^{(1)} \sqrt{X_t^{(2)}}} dW_t^{(1)} \\
+ \int_0^T \frac{Y_t^{(2,2)} - \rho X_t^{(1)} Y_t^{(1,2)}}{\sqrt{1-\rho^2} X_t^{(1)} \sqrt{X_t^{(2)}}} dW_t^{(2)} \right) \right]
\]
Monte Carlo simulation of Delta_{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 $\lambda = 0.1$.

The derivations are placed in the Appendix and examples are shown in Figure 3 and Figure 4, where the stochastic Euler scheme has been used in the Monte Carlo 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

\begin{align*}
\text{Vega}_x &= E \left[ e^{-rT} \phi(X_T^{(1)}) \frac{1}{T} \left( \int_0^T W_r^{(1)} \sqrt{X_t^{(2)}} dt - \int_0^T \frac{dW_r^{(1)}}{\sqrt{X_t^{(2)}}} \right) \right] \\
\text{Lambda}_x &= E \left[ e^{-rT} \phi(X_T^{(1)}) \left( N_T - \frac{\alpha}{\sqrt{X_T^{(2)}}} \right) \left( \int_0^T \frac{dW_r^{(1)}}{\sqrt{X_t^{(2)}}} - \frac{\rho}{\sqrt{1-\rho^2}} \int_0^T \frac{dW_r^{(2)}}{\sqrt{X_t^{(2)}}} \right) \right] \\
\text{Alpha}_x &= E \left[ e^{-rT} \phi(X_T^{(1)}) \left( \frac{N_T}{\sqrt{X_T^{(2)}}} - \lambda \right) \left( \int_0^T \frac{dW_r^{(1)}}{\sqrt{X_t^{(2)}}} - \frac{\rho}{\sqrt{1-\rho^2}} \int_0^T \frac{dW_r^{(2)}}{\sqrt{X_t^{(2)}}} \right) \right]
\end{align*}

The two vegas are different in the sense that Vega_{x} is the sensitivity to changes in the initial value of the volatility process and Vega_{P} is the sensitivity to a perturbation as in section 3.3. Since the initial value of the volatility process is just another parameter coming from the calibration Vega_{x} is perhaps not as interesting as Vega_{P} which is the analogue of Vega_{BS}. 

Figure 2: Monte Carlo simulation of Delta_{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 3: Monte Carlo simulation of $\Delta \tau_H$ 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: Monte Carlo simulation of $\Delta \tau_H$ 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 $\lambda = 0.1$. 

22
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.

A Skorohod integral for \( Fu_t \) when \( u_t \) previsible

Claim: For \( Fu_t \) in \( L^2(\Omega \times \mathbb{R}^+) \) where

\[
F = E[F] + \sum_{n=1}^{\infty} \sum_{0 \leq t_1, \ldots, t_n \leq m} L_{n}^{t_1, \ldots, t_n}(f_{t_1, \ldots, t_n}),
\]

and \( u_t \) previsible, we have

\[
\delta^{(1)}(Fu) = \int_{0}^{\infty} E[F]u_t dG_t(t_1) + \sum_{n=1}^{\infty} \sum_{0 \leq t_1, \ldots, t_n \leq m} \sum_{k=1}^{n+1} \int_{0}^{\infty} \ldots \int_{0}^{t_{k+1}} \ldots \int_{0}^{t_{k+1}} \ldots \int_{0}^{t_{k+1}} \ldots f_{t_1, \ldots, t_n}(t_1, \ldots, t_{k-1}, t_{k+1}, \ldots, t_{n+1}) dG_t(t_1) \ldots dG_{t_{k-1}}(t_{k-1}) u_t \ldots dG_{t_{n}}(t_{n+1}),
\]

in the sense that \( Hu \in \text{Dom} \delta^{(1)} \) if the right hand side is in \( L^2(\Omega \times \mathbb{R}^+) \).

Proof: By linearity of the Skorohod integral it is enough to prove the claim for the case when \( F \) is a single iterated integral. Take any \( H \in \mathbb{D}^{(1)} \).

\[
E \left[ \int_{0}^{\infty} (D^{(1)}H) Fu_t dt \right] = \int_{0}^{\infty} E \left[ \sum_{n=1}^{\infty} \sum_{0 \leq t_1, \ldots, t_n \leq m} \sum_{k=1}^{n} 1_{\{h_k = t\}} 1_{\{\Sigma^{(h)}_k(t)\}} \int_{0}^{t_{k+2}} \ldots \int_{0}^{t_{k+1}} \ldots \int_{0}^{t_{k+1}} \ldots f_{t_1, \ldots, t_n}(t_1) \ldots dG_{t_{k+1}}(t_{k+1}) \ldots dG_{t_n}(t_n) \right] dt
\]

\[
= E \left[ \sum_{n=1}^{\infty} \sum_{0 \leq t_1, \ldots, t_n \leq m} \sum_{k=1}^{n} 1_{\{h_k = t\}} \left( \int_{0}^{t_{k+2}} \ldots \int_{0}^{t_{k+1}} \ldots \int_{0}^{t_{k+1}} \ldots f_{t_1, \ldots, t_n}(t_1) \ldots dG_{t_{k+1}}(t_{k+1}) \ldots dG_{t_n}(t_n) \right) \right] dt
\]

\[
= E \left[ \sum_{n=1}^{\infty} \sum_{0 \leq t_1, \ldots, t_n \leq m} \sum_{k=1}^{n} 1_{\{h_k = t\}} \left( \int_{0}^{t_{k+2}} \ldots \int_{0}^{t_{k+1}} \ldots \int_{0}^{t_{k+1}} \ldots dG_{t_{k-1}}(t_{k-1}) u_t \ldots dG_{t_{n}}(t_{n+1}) \right) \right].
\]
For the $D_{l}^{(1)}H$ contribution, the integrals outside the $dt$ integral can be matched up with the outer integrals in the $F$–contribution with an isometry argument:

$$E \left[ \int_{0}^{\infty} (D_{l}^{(1)}H) Fu_{i} dt \right]$$

$$= E \left[ \sum_{n=1}^{\infty} \sum_{i_{1} \leq i_{t}, \ldots, i_{n} \leq m} \sum_{k=1}^{n} 1_{\{i_{k}=l\}} 1_{\{(i_{k+1}, \ldots, i_{n})=(j_{n-k+1}, \ldots, j_{n})\}} \lambda_{i_{k+1}} \cdots \lambda_{i_{n}} \int_{0}^{t_{i_{k+1}}} \cdots \int_{0}^{t_{i_{k}}} \int_{0}^{t_{2}} \sum_{q=1}^{t_{i_{k}}-t_{2}} \cdots \int_{0}^{t_{1}} h_{i_{1}, \ldots, i_{n}} (\ldots) dG_{i_{1}}(t_{1}) \cdots dG_{i_{n}}(t_{n}) \right]$$

$$= E \left[ \sum_{n=1}^{\infty} \sum_{i_{1} \leq i_{t}, \ldots, i_{n} \leq m} \sum_{k=1}^{n} 1_{\{i_{k}=l\}} 1_{\{(i_{k+1}, \ldots, i_{n})=(j_{n-k+1}, \ldots, j_{n})\}} \lambda_{i_{k+1}} \cdots \lambda_{i_{n}} \int_{0}^{t_{i_{k+1}}} \cdots \int_{0}^{t_{i_{k}}} \int_{0}^{t_{2}} \sum_{q=1}^{t_{i_{k}}-t_{2}} \cdots \int_{0}^{t_{1}} h_{i_{1}, \ldots, i_{n}} (\ldots) dG_{i_{1}}(t_{1}) \cdots dG_{i_{n}}(t_{n}) \right]$$

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

$$E \left[ \int_{0}^{\infty} (D_{l}^{(1)}H) Fu_{i} dt \right]$$

$$= E \left[ \sum_{n=1}^{\infty} \sum_{i_{1} \leq i_{t}, \ldots, i_{n} \leq m} \sum_{k=1}^{n} 1_{\{i_{k}=l\}} 1_{\{(i_{k+1}, \ldots, i_{n})=(j_{n-k+1}, \ldots, j_{n})\}} \lambda_{i_{k+1}} \cdots \lambda_{i_{n}} \int_{0}^{t_{i_{k+1}}} \cdots \int_{0}^{t_{i_{k}}} \int_{0}^{t_{2}} \sum_{q=1}^{t_{i_{k}}-t_{2}} \cdots \int_{0}^{t_{1}} h_{i_{1}, \ldots, i_{n}} (\ldots) dG_{i_{1}}(t_{1}) \cdots dG_{i_{n}}(t_{n}) \right]$$

$$= E \left[ \sum_{n=1}^{\infty} \sum_{i_{1} \leq i_{t}, \ldots, i_{n} \leq m} \left( \int_{0}^{\infty} \int_{0}^{t_{i_{k+1}}} \cdots \int_{0}^{t_{i_{k}}} \int_{0}^{t_{2}} \sum_{q=1}^{t_{i_{k}}-t_{2}} \cdots \int_{0}^{t_{1}} h_{i_{1}, \ldots, i_{n}} (\ldots) dG_{i_{1}}(t_{1}) \cdots dG_{i_{n}}(t_{n}) \right) \sum_{k=1}^{n} 1_{\{i_{k}=l\}} 1_{\{(i_{k+1}, \ldots, i_{n})=(j_{n-k+1}, \ldots, j_{n})\}} \lambda_{i_{k+1}} \cdots \lambda_{i_{n}} \int_{0}^{t_{i_{k+1}}} \cdots \int_{0}^{t_{i_{k}}} \int_{0}^{t_{2}} \sum_{q=1}^{t_{i_{k}}-t_{2}} \cdots \int_{0}^{t_{1}} h_{i_{1}, \ldots, i_{n}} (\ldots) dG_{i_{1}}(t_{1}) \cdots dG_{i_{n}}(t_{n}) \right]$$

$$= E \left[ \sum_{n=1}^{\infty} \sum_{i_{1} \leq i_{t}, \ldots, i_{n} \leq m} \left( \int_{0}^{\infty} \int_{0}^{t_{i_{k+1}}} \cdots \int_{0}^{t_{i_{k}}} \int_{0}^{t_{2}} \sum_{q=1}^{t_{i_{k}}-t_{2}} \cdots \int_{0}^{t_{1}} h_{i_{1}, \ldots, i_{n}} (\ldots) dG_{i_{1}}(t_{1}) \cdots dG_{i_{n}}(t_{n}) \right) \sum_{k=1}^{n} 1_{\{i_{k}=l\}} 1_{\{(i_{k+1}, \ldots, i_{n})=(j_{n-k+1}, \ldots, j_{n})\}} \lambda_{i_{k+1}} \cdots \lambda_{i_{n}} \int_{0}^{t_{i_{k+1}}} \cdots \int_{0}^{t_{i_{k}}} \int_{0}^{t_{2}} \sum_{q=1}^{t_{i_{k}}-t_{2}} \cdots \int_{0}^{t_{1}} h_{i_{1}, \ldots, i_{n}} (\ldots) dG_{i_{1}}(t_{1}) \cdots dG_{i_{n}}(t_{n}) \right]$$

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

B **Stochastic weights in the Heston model**

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
\[ b(X_t) = \begin{pmatrix} rX_t^{(1)} \\ \kappa(\theta - X_t^{(2)}) \end{pmatrix}, \quad c(X_t) = \begin{pmatrix} (\alpha - 1)X_t^{(1)} \\ 0 \end{pmatrix} \] and
\[ a(X_t) = \begin{pmatrix} \sqrt{X_t^{(2)}}X_t^{(1)} \\ \sigma\sqrt{X_t^{(2)}} & \sigma\sqrt{X_t^{(2)}}(1-\rho^2) \end{pmatrix}, \]
with the notation used in section 4. The inverse of the diffusion matrix becomes
\[ a(X_t)^{-1} = \frac{1}{\sigma^2 X_t^{(1)}X_t^{(2)}} \begin{pmatrix} \sigma\sqrt{X_t^{(2)}}(1-\rho^2) & 0 \\ -\sigma\sqrt{X_t^{(2)}} & \sigma\sqrt{X_t^{(2)}}X_t^{(1)} \end{pmatrix}. \]

**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
\[ \left( \frac{1}{\sqrt{X_t^{(2)}}}, -\frac{\rho}{\sqrt{X_t^{(2)}}(1-\rho^2)} \right), \]
and we get the expression in section 4

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

The first variation process is in this case a matrix process
\[ dY_t = b'(X_t^\epsilon)Y_t dt + a_1'(X_t^\epsilon)Y_t dW_t^{(1)} + a_2'(X_t^\epsilon)Y_t dW_t^{(2)}, \quad Y_0 = I, \]
with
\[ b'(X_t^\epsilon) = \begin{pmatrix} r + \lambda(1-\alpha) \\ 0 \\ -\kappa \end{pmatrix}, \]
\[ a_1'(X_t^\epsilon) = \begin{pmatrix} \sqrt{X_t^{(2)}} & \frac{1}{2} \frac{X_t^{(1)}}{\sqrt{X_t^{(2)}}} \\ 0 & \frac{1}{2} \frac{\sigma\rho}{\sqrt{X_t^{(2)}}} \end{pmatrix}, \]
\[ a_2'(X_t^\epsilon) = \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{2} \frac{\sigma\sqrt{1-\rho^2}}{\sqrt{X_t^{(2)}}} \end{pmatrix}. \]
For the Delta we are interested in the first row of the matrix \((a^{-1}(X_t^c)Y_t)^*\) integrated over the 2-dimensional Brownian motion. From the above we can deduce that the row vector of interest is
\[
\left( \frac{1}{X_t^{c(1)}} Y_t^{(1,1)}, \frac{1}{\sqrt{1 - \rho^2} X_t^{c(1)}} \frac{1}{X_t^{c(2)}} Y_t^{(1,1)} \right).
\]
The first–row first–column element in the first variation process can be seen to be
\[
Y_t^{(1,1)} = \frac{1}{x} X_t^{c(1)},
\]
since
\[
dY_t^{(1,1)} = (r + \lambda(1 - \alpha))Y_t^{(1,1)} dt + \sqrt{X_t^{c(2)}} Y_t^{(1,1)} dW_t^{(1)} + \frac{1}{2} \frac{X_t^{c(1)}}{\sqrt{X_t^{c(2)}}} Y_t^{(2,1)} dW_t^{(1)}
\]
and
\[
Y_0^{(1,1)} = 1,
\]
and \(Y_t^{(2,1)} = 0\) due to the fact that \(X_t^{c(2)}\) does not depend explicitly on \(X_t^{c(1)}\) and in particular not its initial value. This leaves us with the expression for DeltaH presented in section 4.

**Stochastic weight for Vega**\(^{x2}\) **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
\[
\left( \frac{1}{X_t^{c(1)}} Y_t^{(1,2)}, \frac{1}{\sqrt{1 - \rho^2} X_t^{c(1)}} \frac{1}{X_t^{c(2)}} \left[ \frac{1}{\sigma} Y_t^{(2,2)} - \rho X_t^{c(1)} Y_t^{(1,2)} \right] \right).
\]

**Stochastic weight for Vega**\(^{P}\) **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 \(X_t^{c(2)}\) does not depend explicitly on \(X_t^{c(1)}\) we can deduce that the variation process with respect to \(\epsilon, Z_t^\epsilon\), has a vanishing second
component. We write $Z^{(2)}_t = 0$. Furthermore,

$$
\beta_t = Y_t^{-1} \frac{\partial X^c_t}{\partial X_t} Z_t
$$

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

$$
= \left( Z_t^{(1)} \frac{\partial X_t^{(1)}}{\partial X_t^{(1)^c}} / Y_t^{(1,1)} \right).
$$

Note that by the Itô formula

$$
Z_t^{(1)} = X_t^{(1)} \left( W_t^{(1)} - \int_0^t \sqrt{X_s^{(2)}} \, ds \right),
$$

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 is the Malliavin derivative of $\beta_t$ in the direction of the Brownian motion. Using the chain rule (Theorem 2.5) on a sequence of continuously differentiable functions with bounded derivatives approximating $\sqrt{X_t^{(2)}}$, together with (3.2) we get

$$
D_t^{(0)} \beta_t^{(1)} = x_1 \left( (1, 0) - \frac{1}{2} \int_0^T \frac{1}{\sqrt{X_t^{(2)}}} D_t^{(0)} X_s^{(2)} \, ds \right)
$$

$$
= x_1 \left( (1, 0) - \frac{1}{2} \int_t^T \frac{\sqrt{X_t^{(2)}} \, Y_s^{(2,2)} (\rho, \sqrt{1 - \rho^2})}{\sqrt{X_t^{(2)}}} \, ds \right).
$$

To arrive at the expression presented in section 4 we note that

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

**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

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

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

27
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^P X_t^{(2)} does not depend explicitly on X_t^{(1)} so we get
\[ \beta_t = Y_t^{-1} \frac{\partial X_t^{(1)}}{\partial X_t} Z_t \]
\[ = \frac{1}{Y_t^{(1,1)} Y_t^{(2,2)}} \begin{pmatrix} Y_t^{(2,2)} & -Y_t^{(1,2)} \\ 0 & Y_t^{(1,1)} \end{pmatrix} \begin{pmatrix} \frac{\partial X_t^{(1)}}{\partial X_t^{(1)}} & \frac{\partial X_t^{(1)}}{\partial X_t^{(2)}} \\ \frac{\partial X_t^{(2)}}{\partial X_t^{(1)}} & \frac{\partial X_t^{(2)}}{\partial X_t^{(2)}} \end{pmatrix} \begin{pmatrix} Z_t^{(1)} \\ 0 \end{pmatrix} \]
\[ = \begin{pmatrix} Z_t^{(1)} \frac{\partial X_t^{(1)}}{\partial X_t^{(1)}} / Y_t^{(1,1)} \\ 0 \end{pmatrix}. \]

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 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 Section 4.

References


