KPMG Center of Excellence
in Risk Management

Titel: Representations for conditional expectations and applications to pricing and hedging of financial products in Lévy and jump-diffusion setting

Catherine Daveloose
Department of Applied Mathematics,
Computer Science and Statistics,
Ghent University,
Krijgslaan 281 S9, 9000 Gent, Belgium,
email: Catherine.Daveloose@UGent.be

Asma Khedher
Chair of Mathematical Finance,
Technische Universität München,
Parkring 11, 85748 Garching-Hochbrück, Germany,
email: asma.khedher@tum.de

Michèle Vanmaele
Department of Applied Mathematics,
Computer Science and Statistics,
Ghent University,
Krijgslaan 281 S9, 9000 Gent, Belgium,
email: michele.vanmaele@ugent.be
http://users.ugent.be/ mvmaele/

Date: December 13^{th}, 2014.

Eine Einrichtung der TUM – gefördert von der
KPMG AG Wirtschaftsprüfungsgesellschaft

Abstract

In this paper, we derive expressions for conditional expectations in terms of regular expectations without conditioning but involving some weights. For this purpose we apply two approaches: the conditional density method and the Malliavin method. We use these expressions for the numerical estimation of the price of American options and their deltas in a Lévy and jump-diffusion setting. Several examples of applications to financial and energy markets are given including numerical examples.

Keywords: Conditional expectation, Monte Carlo methods, Conditional density method, Malliavin calculus, Pricing, Lévy processes, American option, Reduction of variance

1. Introduction

In this paper we consider the problem of computing conditional expectations of functionals of Lévy processes and jump-diffusions. We apply the developed theory to the numerical estimation of American option prices and their deltas.

In general it is not possible to obtain analytical expressions for conditional expectations and thus numerical methods are called for. Several approaches appeared in this field. [16] built up a tree in order to obtain a discretisation of the underlying diffusion on a grid (see also [7]). [27] use a regression method on a truncated basis of L^2 and then choose a basis of polynomials for the numerical estimation of conditional expectations. [20] derive expressions for the conditional expectations in terms of regular expectations for diffusion models.

Considering a random variable F, a scalar random variable G, and a function f on $\mathbb{R}$, [20] provide the following representation for the conditional expectation

$$\mathbb{E}[f(F)|G=0] = \frac{\mathbb{E}[f(F)H(G)\pi]}{\mathbb{E}[H(G)\pi]},$$

where π is a random variable called weight and H is the Heaviside step function increased with some constant, $H(x) = \mathbf{1}_{\{x\geq 0\}} + c$, $c \in \mathbb{R}$. The authors use two approaches: the density method and the Malliavin method. The density method requires that the couple (F, G) has a density $p(x, y)$, $(x \in \mathbb{R}, y \in \mathbb{R})$ such that its log is C^1 in the first argument. In the Malliavin approach, they use a Malliavin derivative of the Wiener process and provide expressions for conditional expectations, where F and G are modelled by continuous diffusions. One of the goals in the present paper is to relax the conditions imposed on the random variables F and G and in particular to allow for random variables which do not necessarily have a known density and which might originate from processes with jumps.

We recall that the density method introduced in [20] requires the knowledge of the density of (F, G). However when F and G are random variables generated from jump processes, the density of the couple (F, G) is in general not known or very hard to compute. This shortcoming can be overcome by using the *conditional density method* introduced by [11]. For example, in the case of a geometric Lévy process, we only need the knowledge of the joint density of the continuous parts, which we do know. Thus to apply the conditional density method a separability assumption on the random variables F and G will be required. F and G should consist of a part with known density and another part with unknown density.

For the *Malliavin method*, we work with the Malliavin derivative for jump processes developed by [30]. The idea is to use the Malliavin derivative in the direction of the Wiener term in the jump-diffusion process. Using this approach there is no separability assumption imposed, since the

Malliavin calculus as presented by [30] does not require any, as opposed to the Malliavin calculus used in [17] or in [12].

Furthermore, we provide expressions for the derivative of conditional expectations using both approaches and we illustrate our results with several examples of models which are commonly used in financial and energy markets. Notice that we present our study in the one-dimensional case for the ease of notation, although all results can be extended to a setting in higher dimensions.

The representations that we develop are interesting from a probabilistic point of view. Indeed we derive expressions for the conditional expectations of functionals of random variables involving only unconditional expectations of these functionals. Moreover, these representations are interesting from a numerical point of view. In the present paper, we apply them to the numerical estimation of American option prices and their deltas, the delta being the sensitivity of the option price with respect to the state of the underlying asset. In complete markets the delta is known to be the number of assets to hold in a self-financing portfolio replicating the option. It is also used in incomplete markets as an imperfect hedge.

To perform the numerical experiments, American options are approximated, through a time discretisation, by Bermudan options (see [3]). We make use of a localisation technique and a control variable to minimise the variance. To reduce the memory capacity of the algorithm for the estimation of the American option price and the delta, we suggest to simulate the underlying stock price process backwards in time. This backward simulation technique turns out to be a specific application of Lévy bridges, see [2].

To check the accuracy of the proposed algorithms, we first compute European option prices and their deltas at time $t > 0$ where we assume a Merton model for the price process. We compare the values obtained by our algorithm to the analytical solutions proposed by [29]. Then considering the same model we estimate the prices and the deltas of American options, which we in turn compare to estimates found in the literature.

The fundamental difference between the [27] approach and the (conditional) density or Malliavin approach is the way the conditional expectations are approximated. Furthermore, the [27] method is unable to provide an ad hoc method for the computation of the delta. It has to be combined with other methods such as the likelihood ratio method or pathwise sensitivities based approaches in order to obtain an approximation of the delta. The approaches presented in this paper lead to representation formulas for the derivative of conditional expectations and consequently provide an estimation of the delta using its own specific method. For more about advantages and drawbacks from considering the [27] algorithm or the Malliavin approach algorithm we refer to [13]. In that paper the authors performed a numerical comparison and discussed the efficiency and the level of complexity of both algorithms for continuous processes.

The paper is organised as follows. In Section 2 and 3 we develop a representation for conditional expectations via the conditional density method and the Malliavin method, respectively. In Section 4, we present variance reduction techniques to obtain acceptable convergence results in numerical applications. In Section 5, we present numerical examples to illustrate our results. Section 6 concludes the paper.

2. Conditional expectation via the conditional density method

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space equipped with a filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$ for time horizon $T > 0$, satisfying the *usual conditions* (see [31]). We introduce the generic notation $L = \{L_t\}_{0\leq t\leq T}$, for a Lévy process on the given probability space. We set $L_0 = 0$ by convention and work with the right-continuous with left limits version of the Lévy process. Let $\Delta L_t := L_t - L_{t-}$ indicate the jump of the Lévy process L at time t. Denote the Lévy measure of L by $\ell(dz)$. Recall that $\ell(dz)$ is a σ-finite Borel measure on $\mathbb{R}_0 := \mathbb{R}\backslash\{0\}$.

In this paper, we express the realisation of a conditional expectation $\mathbb{E}[f(S_t)|S_s = \alpha]$ in terms of regular expectations. Here f is a Borel measurable function (think for instance of the payoff function of a call option), S is an $\mathcal{F}$-adapted price process which may have jumps, and α is a real number. We also rewrite its differential w.r.t. α, i.e. the delta, by only using unconditional expectations.

First, we state a general result for the conditional expectation $\mathbb{E}[f(F)|G=\alpha]$, where F and G are two random variables satisfying the following separability assumptions.

Assumptions 2.1 (Separability). *Let F and G be two random variables such that*

$$F = g_1(X,Y) \quad \text{and} \quad G = g_2(U,V).$$

Herein the random variables X, Y, U, and V are $\mathcal{F}_T$-measurable and have bounded moments. The couple (X,U) is independent of (Y,V). Moreover

1. *(X,U) has a density $p_{(X,U)}$ with respect to the Lebesgue measure,*
2. *$\log p_{(X,U)}(x,\cdot) \in C^1$, for all $x \in \mathbb{R}$, and*
3. *$\frac{\partial}{\partial u} \log p_{(X,U)}(x,\cdot)$ has at most polynomial growth at infinity, for all $x \in \mathbb{R}$.*

The functions g_1 and g_2 are Borel measurable and there exist a Borel measurable function g^* and a strictly increasing differentiable function h such that

$$g_2(u,v) = h^{-1}(u + g^*(v)), \tag{2.1}$$

om for all $(u,v) \in \operatorname{Dom} g_2 \cap (\mathbb{R} \times \operatorname{Dom} g^*)$.

In this section, we require the following assumption for the function f.

Assumption 2.2. *Let f be a Borel measurable function with at most polynomial growth at infinity.*

We apply the *conditional density method* as it is developed in [11]. This method does not require the density of the couple (F,G) but only the density of a part, which we denote (X,U). The density $p_{(X,U)}$ of (X,U) plays the most important role in this method. The results follow from straightforward computations based on properties of (conditional) expectations.

We denote the *Heaviside step function* increased by an arbitrary number $c \in \mathbb{R}$ by $H(x) := \mathbf{1}_{\{x \geq 0\}} + c$. The distributional derivative of this function equals the *Dirac delta function* δ_0. For a function $g \in C^1$ with a single root, the composition rule (see [32]) states that

$$\delta_0(g(x)) = \frac{\delta_0(x - x_1)}{|g'(x_1)|}, \tag{2.2}$$

where x_1 is such that $g(x_1) = 0$ and $g'(x_1) \neq 0$.

Theorem 2.3. *Let F and G be as described in Assumptions 2.1 and let the function f satisfy Assumption 2.2. Then it holds for any α in $\mathbb{R}$ that*

$$\mathbb{E}[f(F)|G=\alpha] = \frac{\mathbb{E}[f(F)H(G-\alpha)\pi_{(X,U)}]}{\mathbb{E}[H(G-\alpha)\pi_{(X,U)}]}, \tag{2.3}$$

where

$$\pi_{(X,U)} = -\frac{\partial}{\partial u} \log p_{(X,U)}(X,U). \tag{2.4}$$

Proof. Using the definition of the conditional expectation, we know that

$$\mathbb{E}[f(F)|G=\alpha] = \frac{\mathbb{E}[f(F)\delta_0(G-\alpha)]}{\mathbb{E}[\delta_0(G-\alpha)]}. \tag{2.5}$$

Moreover we have that

$$\begin{aligned} \mathbb{E}[f(F)\delta_0(G-\alpha)] &= \mathbb{E}[f(g_1(X,Y))\delta_0(g_2(U,V)-\alpha)] \\ &= \mathbb{E}\Big[\mathbb{E}\big[f(g_1(X,Y))\delta_0(g_2(U,V)-\alpha)|\sigma(Y,V)\big]\Big], \end{aligned} \tag{2.6}$$

KPMG Center of Excellence in Risk Management

where $\sigma(Y, V)$ is the filtration generated by Y and V. From Assumptions 2.1(1) we derive

$$\mathbb{E}[f(F)\delta_0(G-\alpha)] = \mathbb{E}\Big[\int_{\mathbb{R}^2} f(g_1(x,Y))\delta_0(g_2(u,V)-\alpha)p_{(X,U)}(x,u)dxdu\Big].$$

By using the composition rule (2.2) for the Dirac delta function and relation (2.1), we obtain that

$$\begin{aligned}
&\int_{\mathbb{R}} \delta_0(g_2(u,V)-\alpha)p_{(X,U)}(x,u)du \\
&= \int_{\mathbb{R}} \frac{\delta_0(u+g^*(V)-h(\alpha))}{\frac{\partial}{\partial u}g_2(h(\alpha)-g^*(V),V)} p_{(X,U)}(x,u)du \\
&= h'(\alpha)\int_{\mathbb{R}} \delta_0(u+g^*(V)-h(\alpha))p_{(X,U)}(x,u)du.
\end{aligned}$$

The Dirac delta function is the distributional derivative of the Heaviside step function. Hence by integration by parts we find that

$$\begin{aligned}
&\int_{\mathbb{R}} \delta_0(g_2(u,V)-\alpha)p_{(X,U)}(x,u)du \\
&= -h'(\alpha)\int_{\mathbb{R}} H(u+g^*(V)-h(\alpha))\frac{\partial}{\partial u}p_{(X,U)}(x,u)du \\
&= -h'(\alpha)\int_{\mathbb{R}} H(g_2(u,V)-\alpha)\Big(\frac{\partial}{\partial u}\log p_{(X,U)}(x,u)\Big)p_{(X,U)}(x,u)du.
\end{aligned}$$

Finally we conclude that

$$\begin{aligned}
\mathbb{E}[f(F)\delta_0(G-\alpha)] &= \mathbb{E}\Big[\mathbb{E}\big[f(F)H(G-\alpha)\Big\{-\frac{\partial}{\partial u}\log p_{(X,U)}(X,U)\Big\}h'(\alpha)\big|\sigma(Y,V)\big]\Big] \\
&= \mathbb{E}\Big[f(F)H(G-\alpha)\Big\{-\frac{\partial}{\partial u}\log p_{(X,U)}(X,U)\Big\}\Big]h'(\alpha).
\end{aligned}$$

Applying the latter result with $f \equiv 1$ for the denominator of (2.5) we prove the statement.

Notice that in expression (2.5), the conditional expectation is written in terms of regular expectations. However, when evaluating the Dirac delta function δ_0, it will turn out to be zero for many simulated values, which might result in a denominator equal to zero. This is not convenient for numerical experiments such as Monte Carlo estimations. Therefore we proceeded in the proof above by moving to the Heaviside step function. As a result we obtain representation (2.3).

It is clear that the weight $\pi_{(X,U)}$ in expression (2.3) can be replaced by any weight π in the set $\mathcal{W}$ defined as

$$\mathcal{W} := \{\pi : \mathbb{E}[\pi|\sigma(F,G)] = \pi_{(X,U)}\}. \tag{2.7}$$

However the weight $\pi_{(X,U)}$ remains the optimal weight in this set in the sense of minimal variance as we state in the following proposition. For the proof we refer to [20].

Proposition 2.4. *Let F and G be as described in Assumptions 2.1, let the function f satisfy Assumption 2.2, and take α in $\mathbb{R}$. For all the weights π in the set $\mathcal{W}$, defined by (2.7), we know that $\mathbb{E}[f(F)H(G-\alpha)\pi] = \mathbb{E}[f(F)H(G-\alpha)\pi_{(X,U)}]$ where $\pi_{(X,U)}$ is given by (2.4). The variance*

$$\mathcal{V}(\pi) := Var(f(F)H(G-\alpha)\pi)$$

is minimised over this set $\mathcal{W}$ at $\pi_{(X,U)}$.

The random variable Y does not appear in the denominator of expression (2.5). Thus it is more natural to condition on the filtration $\sigma(V)$ than on $\sigma(Y,V)$ in (2.6). Moreover the weight should not depend on X. These points lead to the following theorem.

Theorem 2.5. *Let the function f satisfy Assumption 2.2. Let F and G be as described in Assumptions 2.1, moreover assume that*

1. *U has a density p_U with respect to the Lebesgue measure,*
2. *$\log p_U \in C^1$, and*
3. *$\frac{\partial}{\partial u} \log p_U$ has at most polynomial growth at infinity.*

Then it holds for any α in $\mathbb{R}$ that

$$\mathbb{E}[f(F)|G = \alpha] = \frac{\mathbb{E}[f(F)H(G - \alpha)\pi_{(X,U)}]}{\mathbb{E}[H(G - \alpha)\pi_U]},$$

where $\pi_{(X,U)}$ is given by (2.4) and

$$\pi_U = -\frac{\partial}{\partial u} \log p_U(U). \tag{2.8}$$

Proof. This result can be obtained through similar computations as in the proof of Theorem 2.3.

Proposition 2.6. *Let G be as described in Assumptions 2.1 such that conditions (1)-(3) from Theorem 2.5 hold, and take α in $\mathbb{R}$. For all the weights π in the set*

$$\mathcal{W}' := \{\pi : \mathbb{E}[\pi|\sigma(G)] = \pi_U\}$$

we know that $\mathbb{E}[H(G - \alpha)\pi] = \mathbb{E}[H(G - \alpha)\pi_U]$ where π_U is given by (2.8). The variance

$$\mathcal{V}'(\pi) := Var(H(G - \alpha)\pi)$$

is minimised over this set $\mathcal{W}'$ at π_U.

In case it holds that $\mathbb{E}[-\frac{\partial}{\partial u} \log p_{(X,U)}(X,U)|\sigma(G)] = -\frac{\partial}{\partial u} \log p_U(U)$, then $\pi_{(X,U)} \in \mathcal{W}'$ and π_U turns out to be the optimal weight compared to $\pi_{(X,U)}$ for the denominator of the representation. However, [13] noticed from numerical tests that the use of the same weight in the numerator and denominator leads to numerical compensations which seem to stabilise the algorithm. Therefore we decide to use the same weight in numerator and denominator in our numerical experiments in Section 5.

In many applications in mathematical finance, one can make grateful use of Theorem 2.3. In fact, we are able to express a realisation of the conditional expectation of the form $\mathbb{E}[f(S_t)|S_s = \alpha]$ in terms of regular expectations. Here f is a function e.g. a payoff function, $(S_t)_{t\in[0,T]}$ represents a Markovian stock price process, $0 < s < t < T$, and α is a real number. The expressions with unconditional expectations allow us to use Monte Carlo simulations to evaluate such conditional expectations.

The next proposition considers such a representation of conditional expectations in terms of unconditional expectations, where the price process S is an exponential jump-diffusion process. This type of stock price model is common in finance and conditional expectations appear when determining European option prices at times $t > 0$ or American option prices.

Proposition 2.7. *Observe a price process S defined by $S_t = \mathrm{e}^{L_t}$, $\forall t \in [0,T]$, where L is a Lévy process with decomposition $L_t = \mu t + \beta W_t + \widetilde{N}_t$. Here W is a standard Brownian motion, $\widetilde{N}$ is a compound Poisson process independent of W, and μ and β are constant parameters. Then, for any function f fulfilling Assumption 2.2, any positive number α, and $0 < s < t < T$, it holds that*

$$\mathbb{E}[f(S_t)|S_s = \alpha] = \frac{\mathbb{E}[f(S_t)H(S_s - \alpha)\pi]}{\mathbb{E}[H(S_s - \alpha)\pi]},$$

where

$$\pi = \frac{tW_s - sW_t}{\beta s(t - s)}. \tag{2.9}$$

PROOF. Following the notation of Theorem 2.3, we set $F = S_t$ and $G = S_s$. The random variables $X = \beta W_t$, $Y = \mu t + \widetilde{N}_t$, $U = \beta W_s$ and $V = \mu s + \widetilde{N}_s$ and the functions $g_i(x, y) = \mathrm{e}^{x+y}$, $i \in \{1, 2\}$, $g^*(v) = v$ and $h(\alpha) = \log \alpha$ satisfy Assumptions 2.1. The scaled Brownian motions X and U have a joint normal distribution with density function

$$p_{(X,U)}(x, u) = \frac{1}{2\pi\beta^2\sqrt{(t-s)s}} \exp\left(-\frac{x^2 s - 2xus + u^2 t}{2\beta^2(t-s)s}\right). \tag{2.10}$$

To determine the weight in (2.3) we calculate

$$-\frac{\partial}{\partial u} \log p_{(X,U)}(x, u) = \frac{\partial}{\partial u} \frac{x^2 s - 2xus + u^2 t}{2\beta^2(t-s)s} = \frac{ut - xs}{\beta^2(t-s)s},$$

such that we obtain

$$\pi = \frac{Ut - Xs}{\beta^2 s(t-s)} = \frac{tW_s - sW_t}{\beta s(t-s)}.$$

According to Theorem 2.5, we can replace the weight in the denominator by $\pi = W_s/(s\beta)$. Since it holds that

$$\mathbb{E}\Big[\frac{tW_s - sW_t}{\beta(t-s)s}\big|\sigma(S_s)\Big] = \frac{W_s}{s\beta},$$

we conclude, by Proposition 2.6, that the weight $W_s/(s\beta)$ is preferred above π (2.9) for the denominator in minimal variance sense.

In the sequel we observe a model which is often used to price energy products (see for example [10]). The price process is given by an additive model

$$S_t = X_t + Y_t,\ \forall t \in [0, T] \quad \text{with } S_0 > 0. \tag{2.11}$$

The process Y is adapted to the filtration $\mathcal{F}$ and does not have to be specified here. The process X is a so called $\Gamma(a, b)$-Ornstein-Uhlenbeck process, see Section 17 in [33]. Namely, it is a process following the dynamics

$$dX_t = -\lambda X_t dt + dL_t, \quad X_0 = S_0, \tag{2.12}$$

where $\lambda > 0$ and L is a subordinator, admitting a stationary distribution for the process X which is here $\Gamma(a, b)$. Hence this means that X_t has a $\Gamma(a, b)$-distribution for all $t > 0$. The solution of the stochastic differential equation (2.12) equals

$$X_t = \mathrm{e}^{-\lambda t} S_0 + \int_0^t \mathrm{e}^{\lambda(r-t)} dL_r.$$

An interesting property of OU-processes is the fact that the autocorrelation is independent of the stationary distribution, it equals

$$\mathrm{Corr}(X_t, X_s) = \mathrm{e}^{\lambda(s-t)}, \qquad \forall\ 0 < s < t. \tag{2.13}$$

Proposition 2.8. *Let us observe the additive model described by* (2.11) *and* (2.12). *Then it holds for any function f satisfying Assumption 2.2, $0 < s < t < T$, and $\alpha \in \mathbb{R}$, that*

$$\mathbb{E}[f(S_t)|S_s = \alpha] = \frac{\mathbb{E}[f(S_t)H(S_s - \alpha)\pi]}{\mathbb{E}[H(S_s - \alpha)\pi]},$$

where

$$\pi = \frac{1-a}{X_s} + \frac{b}{1-\rho} - \frac{\mathcal{I}_a(v(X_t, X_s))\frac{\partial v}{\partial u}(X_t, X_s)}{\mathcal{I}_{a-1}(v(X_t, X_s))}.$$

Herein, $\mathcal{I}_a$ is the modified Bessel function of the first kind with index a,

$$\rho = \mathrm{e}^{\lambda(s-t)} \quad \textit{and} \quad v(x, u) = \frac{2\sqrt{\rho b^2 xu}}{1-\rho}.$$

Proof. As in Theorem 2.3, we put $F = S_t = X_t + Y_t$, $G = S_s = X_s + Y_s$, $(X, U) = (X_t, X_s)$, and $h(\alpha) = \alpha$ to satisfy Assumptions 2.1. To obtain the weight we need the density function of the vector (X_t, X_s). Since X is a $\Gamma(a,b)$-OU process we know X_t and X_s are both $\Gamma(a,b)$ distributed and by (2.13) we know $\mathrm{Corr}(X_t, X_s) = \mathrm{e}^{\lambda(s-t)} =: \rho$. According to [15], the density function of this bivariate gamma distribution with non-zero correlation equals

$$p_{(X_t,X_s)}(x,u) = \frac{(b^2xu)^{(a-1)/2}\exp\big(-(bx+bu)/(1-\rho)\big)}{\rho^{(a-1)/2}(1-\rho)\Gamma(a)}\mathcal{I}_{a-1}\Big(\frac{2\sqrt{\rho b^2xu}}{1-\rho}\Big),$$

where $\mathcal{I}_a$ is the modified Bessel function of the first kind with index a. We compute

$$\frac{\partial}{\partial u}\log p_{(X_t,X_s)}(x,u) = \frac{a-1}{2u} - \frac{b}{1-\rho} + \frac{\partial}{\partial u}\log\mathcal{I}_{a-1}\Big(\frac{2\sqrt{\rho b^2xu}}{1-\rho}\Big).$$

For the function $v(x,u) = 2\sqrt{\rho b^2xu}/(1-\rho)$, it holds that $\frac{\partial v}{\partial u}(x,u) = \sqrt{(\rho b^2x/u)}/(1-\rho)$ and $\frac{\partial v}{\partial u}(x,u)/v(x,u) = 1/(2u)$. Using the recurrence formulas for modified Bessel functions (see [14]), we get

$$\begin{aligned}
&\frac{\partial}{\partial u}\log(\mathcal{I}_{a-1}(v(x,u)))\\
&= \frac{1}{\mathcal{I}_{a-1}(v(x,u))}\mathcal{I}'_{a-1}(v(x,u))\frac{\partial v}{\partial u}(x,u)\\
&= \frac{1}{\mathcal{I}_{a-1}(v(x,u))}\frac{1}{2}\big(\mathcal{I}_{a-2}(v(x,u)) + \mathcal{I}_a(v(x,u))\big)\frac{\partial v}{\partial u}(x,u)\\
&= \frac{1}{\mathcal{I}_{a-1}(v(x,u))}\frac{1}{2}\big(\{\mathcal{I}_{a-2}(v(x,u)) - \mathcal{I}_a(v(x,u))\} + 2\mathcal{I}_a(v(x,u))\big)\frac{\partial v}{\partial u}(x,u)\\
&= \frac{1}{\mathcal{I}_{a-1}(v(x,u))}\frac{1}{2}\Big(\frac{2(a-1)}{v(x,u)}\mathcal{I}_{a-1}(v(x,u)) + 2\mathcal{I}_a(v(x,u))\Big)\frac{\partial v}{\partial u}(x,u)\\
&= \Big(\frac{(a-1)}{v(x,u)} + \frac{\mathcal{I}_a(v(x,u))}{\mathcal{I}_{a-1}(v(x,u))}\Big)\frac{\partial v}{\partial u}(x,u)\\
&= \frac{a-1}{2u} + \frac{\mathcal{I}_a(v(x,u))\frac{\partial v}{\partial u}(x,u)}{\mathcal{I}_{a-1}(v(x,u))}.
\end{aligned}$$

According to (2.3) we conclude the statement.

For a differentiable (payoff) function f, we deduce the following representation as a consequence of Theorem 2.3. From now on we shorten the notation for the density of the couple (X, U) to the function p.

Theorem 2.9. *Let F and G be as described in Assumptions 2.1 and consider a differentiable function f satisfying Assumption 2.2. Assume the existence of two functions q and r, with at most polynomial growth at infinity, such that*

$$q(x,u) + \frac{1}{p(x,u)}\frac{\partial}{\partial x}\big(r(x,u)p(x,u)\big) = -\frac{\partial}{\partial u}\log p(x,u). \tag{2.14}$$

Then it holds for any $\alpha \in \mathbb{R}$ that

$$\mathbb{E}[f(F)|G=\alpha] = \frac{\mathbb{E}[f(F)H(G-\alpha)\pi_1 - f'(F)H(G-\alpha)\pi_2]}{\mathbb{E}[H(G-\alpha)\pi_1]}, \tag{2.15}$$

where $\pi_1 = q(X,U)$ and $\pi_2 = r(X,U)\frac{\partial}{\partial x}g_1(X,Y)$.

Proof. The following calculations are justified by equation (2.14), integration by parts and properties of conditional expectations:

$$\begin{aligned}
&\mathbb{E}[f(F)H(G-\alpha)\pi_{(X,U)}] = \mathbb{E}[\mathbb{E}[f(F)H(G-\alpha)\pi_{(X,U)}|\sigma(Y,V)]] \\
&= \mathbb{E}\Big[\int_{\mathbb{R}^2} f(g_1(x,Y))H(g_2(u,V)-\alpha)\Big\{-\frac{\partial}{\partial u}\log p(x,u)\Big\}p(x,u)dxdu\Big] \\
&= \mathbb{E}\Big[\int_{\mathbb{R}^2} f(g_1(x,Y))H(g_2(u,V)-\alpha)\Big\{q(x,u)+\frac{1}{p(x,u)}\frac{\partial}{\partial x}\big(r(x,u)p(x,u)\big)\Big\}p(x,u)dxdu\Big] \\
&= \mathbb{E}\Big[\int_{\mathbb{R}^2} \Big\{f(g_1(x,Y))H(g_2(u,V)-\alpha)q(x,u) \\
&\qquad - f'(g_1(x,Y))\frac{\partial}{\partial x}g_1(x,Y)H(g_2(u,V)-\alpha)r(x,u)\Big\}p(x,u)dxdu\Big] \\
&= \mathbb{E}[\mathbb{E}[f(g_1(X,Y))H(g_2(U,V)-\alpha)\pi_1 - f'(g_1(X,Y))H(g_2(U,V)-\alpha)\pi_2|\sigma(Y,V)]] \\
&= \mathbb{E}[f(F)H(G-\alpha)\pi_1 - f'(F)H(G-\alpha)\pi_2].
\end{aligned}$$

Replacing the numerator of (2.3) by this result and the denominator by putting $f \equiv 1$, we obtain the statement.

The relation (2.14) and representation (2.15) are inspired by the result (4.25) - (4.26) from [20]. An important difference however should be noticed. The functions $p, q,$ and r are now functions in the part (X, U), not in the couple (F, G). It is assumed that the density function p of the couple (X, U) is known. If one chooses a function q or r, then one can obtain the other function by (2.14), thus there are infinitely many possibilities for the weights π_1 and π_2.

In the next proposition we consider the exponential jump-diffusion model presented in Proposition 2.7 and we apply the previous result to compute conditional expectations in this setting.

Proposition 2.10. *Assume the exponential Lévy model of Proposition 2.7. For any function f fulfilling Assumption 2.2, any positive number α, and $0 < s < t < T$, we have*

$$\mathbb{E}[f(S_t)|S_s=\alpha] = \frac{\mathbb{E}\Big[f(S_t)H(S_s-\alpha)\dfrac{W_s}{s\beta} - f'(S_t)H(S_s-\alpha)S_t\Big]}{\mathbb{E}\Big[H(S_s-\alpha)\dfrac{W_s}{s\beta}\Big]},$$

and

$$\mathbb{E}[f(S_t)|S_s=\alpha] = \frac{\mathbb{E}\Big[f(S_t)H(S_s-\alpha)\dfrac{W_t}{s\beta} - f'(S_t)H(S_s-\alpha)\dfrac{tS_t}{s}\Big]}{\mathbb{E}\Big[H(S_s-\alpha)\dfrac{W_t}{s\beta}\Big]},$$

among many other possible representations for $\mathbb{E}[f(S_t)|S_s=\alpha]$.

Proof. Based on Theorem 2.9 and the density function (2.10) we consider the following possibilities for the weights. First, choose $r(X,U) = 1$, then we find that

$$q(x,u) = -\frac{\partial}{\partial u}\log p(x,u) - \frac{\partial}{\partial x}\log p(x,u) = \frac{ut-xs}{\beta^2(t-s)s} + \frac{x-u}{\beta^2(t-s)} = \frac{u}{\beta^2 s}.$$

This implies that $\pi_1 = W_s/(s\beta)$. Since $\frac{\partial}{\partial x}g_1(X,Y) = \mathrm{e}^{X+Y} = S_t$, we obtain that $\pi_2 = S_t$. Secondly, if we take $r(X,U) = t/s$, then it turns out that $q(x,u) = x/(s\beta^2)$. Consequently $\pi_1 = W_t/(s\beta)$ and $\pi_2 = tS_t/s$.

The previous results and examples concern computations that appear in the *pricing* of European options at times $t > 0$ and of American options. In the next theorem we deduce a representation for the *delta*.

Theorem 2.11. *Let F and G be as described in Assumptions 2.1 and let the function f satisfy Assumption 2.2. Then it holds for any $\alpha \in \mathbb{R}$ that*

$$\frac{\partial}{\partial\alpha}\mathbb{E}[f(F)|G=\alpha] = \frac{\mathbb{B}_{F,G}[f](\alpha)\mathbb{A}_{F,G}[1](\alpha) - \mathbb{A}_{F,G}[f](\alpha)\mathbb{B}_{F,G}[1](\alpha)}{\mathbb{A}_{F,G}[1](\alpha)^2} h'(\alpha),$$

where

$$\begin{aligned}
\mathbb{A}_{F,G}[\cdot](\alpha) &= \mathbb{E}[\cdot(F)H(G-\alpha)\pi_{(X,U)}], \\
\mathbb{B}_{F,G}[\cdot](\alpha) &= \mathbb{E}[\cdot(F)H(G-\alpha)(-\pi^2_{(X,U)} + \pi^*_{(X,U)})], \\
\pi_{(X,U)} &= -\frac{\partial}{\partial u}\log p_{(X,U)}(X,U), \text{ and} \\
\pi^*_{(X,U)} &= -\frac{\partial^2}{\partial u^2}\log p_{(X,U)}(X,U).
\end{aligned}$$

Proof. From Theorem 2.3 it follows immediately that

$$\frac{\partial}{\partial\alpha}\mathbb{E}[f(F)|G=\alpha] = \frac{\frac{\partial}{\partial\alpha}\{\mathbb{A}_{F,G}[f](\alpha)\}\mathbb{A}_{F,G}[1](\alpha) - \mathbb{A}_{F,G}[f](\alpha)\frac{\partial}{\partial\alpha}\{\mathbb{A}_{F,G}[1](\alpha)\}}{\mathbb{A}_{F,G}[1](\alpha)^2}.$$

For the derivatives in the right hand side, it holds that

$$\frac{\partial}{\partial\alpha}\mathbb{E}[f(F)H(G-\alpha)\pi_{(X,U)}] = -\mathbb{E}[f(F)\delta_0(G-\alpha)\pi_{(X,U)}].$$

Along the lines of the proof of Theorem 2.3, we derive

$$\begin{aligned}
&\int_{\mathbb{R}} \delta_0(g_2(u,V)-\alpha)\Big(\frac{\partial}{\partial u}\log p(x,u)\Big)p(x,u)du \\
&= \int_{\mathbb{R}} \delta_0(u+g^*(V)-h(\alpha))h'(\alpha)\frac{\partial}{\partial u}p(x,u)du \\
&= -\int_{\mathbb{R}} H(u+g^*(V)-h(\alpha))h'(\alpha)\frac{\partial^2}{\partial u^2}p(x,u)du \\
&= -\int_{\mathbb{R}} H(g_2(u,V)-\alpha)h'(\alpha)\Big\{\Big(\frac{\partial}{\partial u}\log p(x,u)\Big)^2 + \frac{\partial^2}{\partial u^2}\log p(x,u)\Big\}p(x,u)du,
\end{aligned}$$

which concludes the proof.

3. Conditional expectation via Malliavin method

[20] used the Malliavin method to obtain representations for the conditional expectations. Their approach can be applied to continuous diffusions. This section extends this approach to allow for the computation of conditional expectations in Lévy and jump-diffusion framework. For this purpose we use a Malliavin derivative of the combination of Gaussian and pure jump Lévy noises, see e.g. [18] and [34]. In our setting, we use the Malliavin derivative developed by [30].

Let $(\Omega_W, \mathcal{F}_W, \mathbb{P}_W)$ and $(\Omega_J, \mathcal{F}_J, \mathbb{P}_J)$ be the canonical spaces for the Brownian motion and pure jump Lévy process, respectively. We can interpret

$$\Omega = \Omega_W \times \Omega_J, \quad \mathcal{F} = \mathcal{F}_W \otimes \mathcal{F}_J, \quad \mathbb{P} = \mathbb{P}_W \otimes \mathbb{P}_J,$$

such that $(\Omega, \mathcal{F}, \mathbb{P})$ is again a complete probability space in which Lévy processes are well-defined. We make use of the Malliavin calculus developed in [30]. The *Malliavin derivative in the Brownian direction* is defined in a subspace of $L^2(\Omega)$ and is essentially a derivative with respect to the Brownian part of a Lévy process L. We denote it by $D^{(0)}$. Its dual, the *Skorohod integral* is also defined in [30] and denoted by $\delta^{(0)}$. In this section we make use of some computational rules and properties which are summarised in Appendix A. We still denote the *Heaviside step function* increased by an arbitrary number $c \in \mathbb{R}$ by $H(x) := \mathbf{1}_{\{x\geq 0\}} + c$. In the following theorem we consider a representation for the conditional expectation $\mathbb{E}[f(F)|G=\alpha]$, for a function f satisfying the following assumption.

Assumption 3.1. *Let f be a Borel measurable function which is continuously differentiable ($f \in C^1$) and has a bounded derivative.*

Theorem 3.2. *Observe a function f satisfying Assumption 3.1 and let F and G be in $\mathbb{D}^{(0)}$. Let u be in $\operatorname{Dom}\delta^{(0)}$ such that $f(F)u$ is in $L^2(\Omega \times [0,T])$ and*

$$\mathbb{E}\Big[\int_0^T u_t D_t^{(0)} G dt | \sigma(F,G)\Big] = 1. \tag{3.1}$$

Then it holds for any $\alpha \in \mathbb{R}$ that

$$\mathbb{E}[f(F)|G=\alpha] = \frac{\mathbb{E}[f(F)H(G-\alpha)\delta^{(0)}(u) - f'(F)H(G-\alpha)\int_0^T u_t D_t^{(0)} F dt]}{\mathbb{E}[H(G-\alpha)\delta^{(0)}(u)]}. \tag{3.2}$$

Proof. Recall expression (2.5) for the conditional expectation. By relation (3.1), the chain rule, the duality formula, and integration by parts, we successively find

$$\begin{aligned}
\mathbb{E}[f(F)\delta_0(G-\alpha)] &= \mathbb{E}[f(F)\delta_0(G-\alpha)\mathbb{E}[\int_0^T u_r D_r^{(0)} G dr | \sigma(F,G)]] \\
&= \mathbb{E}[\mathbb{E}[\int_0^T f(F)\delta_0(G-\alpha) u_r D_r^{(0)} G dr | \sigma(F,G)]] \\
&= \mathbb{E}[\int_0^T f(F) u_r D_r^{(0)} H(G-\alpha) dr] = \mathbb{E}[H(G-\alpha)\delta^{(0)}(f(F)u)] \\
&= \mathbb{E}[H(G-\alpha)\{f(F)\delta^{(0)}(u) - \int_0^T u_r D_r^{(0)} f(F) dr\}] \\
&= \mathbb{E}[H(G-\alpha)\{f(F)\delta^{(0)}(u) - f'(F)\int_0^T u_r D_r^{(0)} F dr\}].
\end{aligned}$$

Thus we obtain the numerator in (3.2). Then applying the latter result to $f \equiv 1$ proves the statement.

The latter theorem provides us with a representation formula for the conditional expectation $\mathbb{E}[f(F)|G=\alpha]$ for f being a continuously differentiable function. However in many applications in finance, we often have to consider non smooth functions. Hence we consider the following assumption.

Assumption 3.3. *For a given random variable F with density p_F, let f be a Borel measurable function in $L^2(\mathbb{R}, p_F)$.*

In order to deal with the potential non-smoothness of f, we include an additional assumption on the process u introduced in Theorem 3.2 leading to the following theorem.

Theorem 3.4. *Let F and G be in $\mathbb{D}^{(0)}$ and f a function as in Assumption 3.3. Consider a process u in $\operatorname{Dom}\delta^{(0)}$, guaranteeing $f(F)u$ is in $L^2(\Omega\times[0,T])$, satisfying (3.1) and, in addition,*

$$\mathbb{E}\Big[\int_0^T u_t D_t^{(0)} F dt | \sigma(F,G)\Big] = 0. \tag{3.3}$$

Then the following representation holds for $\alpha \in \mathbb{R}$

$$\mathbb{E}[f(F)|G=\alpha] = \frac{\mathbb{E}[f(F)H(G-\alpha)\delta^{(0)}(u)]}{\mathbb{E}[H(G-\alpha)\delta^{(0)}(u)]}. \tag{3.4}$$

Proof. *i)* First, let us consider a Borel measurable function f which is continuously differentiable with bounded derivative, such that we can apply Theorem 3.2. Because of the properties of conditional expectations and relation (3.3), we have in representation (3.2) that

$$\begin{aligned}\mathbb{E}\Big[f'(F)H(G-\alpha)\int_0^T u_t D_t^{(0)} F dt\Big] &= \mathbb{E}\Big[\mathbb{E}\Big[f'(F)H(G-\alpha)\int_0^T u_t D_t^{(0)} F dt|\sigma(F,G)\Big]\Big] \\ &= \mathbb{E}\Big[f'(F)H(G-\alpha)\mathbb{E}\big[\int_0^T u_t D_t^{(0)} F dt|\sigma(F,G)\big]\Big] = 0.\end{aligned}$$

Thus we obtain representation (3.4).
ii) Now we observe a Borel measurable function f in $L^2(\mathbb{R}, p_F)$. Since $C_K^\infty(\mathbb{R})$ is dense in $L^2(\mathbb{R}, p_F)$ there exists a row of functions f_n in $C_K^\infty(\mathbb{R})$ converging to f in L^2. In part *i)* we concluded that for any function f_n in this row representation (3.4) holds. By convergence arguments, we conclude that expression (3.4) also holds for the limit function f. See Proposition 3.2 in [19] for a rigorous proof in a similar setting concerning the computation of the Greeks.

Analogous to Proposition 2.4 in case of the conditional density method, we can also replace the weight obtained by the Malliavin method with a more general weight π satisfying

$$\mathbb{E}[\pi|\sigma(F,G)] = \delta^{(0)}(u).$$

Over the set of all weights for which this condition holds, the variance of $f(F)H(G-\alpha)\pi$ will be minimised at $\pi = \delta^{(0)}(u)$.

Now we assume S is modelled by the following *stochastic differential equation*

$$\begin{cases} dS_t = \mu(t,S_{t-})dt + \beta(t,S_{t-})dW_t + \int_{\mathbb{R}_0} \gamma(t,S_{t-},z)\widetilde{N}(dt,dz), \\ S_0 = s_0 > 0, \end{cases} \tag{3.5}$$

where W is a Wiener process and $\widetilde{N}$ is a compensated Poisson random measure with Lévy measure ℓ. We assume that $\beta(t,x) > 0$ for all $(t,x) \in [0,T] \times \mathbb{R}$. The coefficient functions $\mu(t,x)$, $\beta(t,x)$, and $\gamma(t,x,z)$ are continuously differentiable with bounded derivatives and Lipschitz continuous in the second argument, for all $(t,z) \in [0,T] \times \mathbb{R}_0$. The coefficients also satisfy the following linear growth condition

$$\mu^2(t,x) + \beta^2(t,x) + \int_{\mathbb{R}_0} \gamma^2(t,x,z)\ell(dz) \leq C(1+x^2),$$

for all $t \in [0,T]$, where C is a positive constant. The existence and uniqueness of the solution S is ensured by Theorem 9.1. Chap IV collected from [23].

The first variation process V related to S equals $\dfrac{\partial S}{\partial s_0}$ and satisfies

$$\begin{cases} dV_t = \mu_x(t,S_{t-})V_{t-}dt + \beta_x(t,S_{t-})V_{t-}dW_t + \int_{\mathbb{R}_0} \gamma_x(t,S_{t-},z)V_{t-}\widetilde{N}(dt,dz), \\ V_0 = 1. \end{cases}$$

The stock price S_t is in $\mathbb{D}^{(0)}$ for all $t \in [0,T]$, and its Malliavin derivative can be expressed in terms of the first variation process (see Theorem 3 and Proposition 7 in [30])

$$D_s^{(0)} S_t = V_t (V_{s-})^{-1} \beta(s, S_{s-}) \mathbf{1}_{\{s \leq t\}}. \tag{3.6}$$

The aim is to find a representation formula for the conditional expectation $\mathbb{E}[f(S_t)|S_s = \alpha]$, $0 < s < t < T$ and $\alpha \in \mathbb{R}$, containing only regular expectations. First we mention the following lemma. We do not present the proof since it is an adaptation of the proof of Lemma 4.1 in [20] to our setting.

KPMG Center of Excellence in Risk Management

Lemma 3.5. *It holds that*

$$D_s^{(0)} V_t = \Big\{ \beta_x(s, S_{s-}) V_t - \frac{\beta(s, S_{s-})\zeta_{s-} V_t}{V_{s-}^2} + \frac{\beta(s, S_{s-})\zeta_t}{V_{s-}} \Big\} \mathbf{1}_{\{s \leq t\}},$$

where $\zeta_t := \dfrac{\partial^2 S_t}{\partial s_0^2}$. *In other words* ζ *is the solution of the SDE*

$$\begin{cases} d\zeta_t = \Big[\mu_{xx}(t, S_{t-})V_{t-}^2 + \mu_x(t, S_{t-})\zeta_t\Big]dt + \Big[\beta_{xx}(t, S_{t-})V_{t-}^2 + \beta_x(t, S_{t-})\zeta_t\Big]dW_t \\ \qquad\qquad + \int_{\mathbb{R}_0} [\gamma_{xx}(t, S_{t-}, z)V_{t-}^2 + \gamma_x(t, S_{t-}, z)\zeta_t]\widetilde{N}(dt, dz), \\ \zeta_0 = 0. \end{cases}$$

Now we have all the ingredients to obtain a useful expression for the conditional expectation $\mathbb{E}[f(S_t)|S_s = \alpha]$, for $\alpha \in \mathbb{R}$. First we make use of Theorem 3.2 and later on we apply Theorem 3.4.

Proposition 3.6. *Let* f *satisfy Assumption 3.1,* $0 < s < t < T$*, and* $\alpha \in \mathbb{R}$*. In the setting described by the stochastic differential equation* (3.5) *we assume that*

$$\mathbb{E}\left[\int_0^T \left(\frac{V_{r-}}{\beta(r, S_{r-})}\right)^2 dr\right] < \infty \quad \text{and} \quad \mathbb{E}\left[\int_0^T \left(\frac{1}{sV_s}\frac{V_{r-}}{\beta(r, S_{r-})}\right)^2 dr\right] < \infty. \tag{3.7}$$

Then the following representation holds for the conditional expectation

$$\mathbb{E}[f(S_t)|S_s = \alpha] = \frac{\mathbb{E}[f(S_t)H(S_s - \alpha)\pi_1 - f'(S_t)H(S_s - \alpha)\pi_2]}{\mathbb{E}[H(S_s - \alpha)\pi_1]}, \tag{3.8}$$

where the Malliavin weights equal

$$\pi_1 = \frac{1}{sV_s}\Big(\int_0^s \frac{V_{r-}}{\beta(r, S_{r-})}dW_r + s\frac{\zeta_s}{V_s} + \int_0^s \Big[\frac{\beta_x(r, S_{r-})}{\beta(r, S_{r-})}V_{r-} - \frac{\zeta_{r-}}{V_{r-}}\Big]dr\Big) \quad \text{and} \quad \pi_2 = \frac{V_t}{V_s}. \tag{3.9}$$

Proof. To fulfill condition (3.1) we define

$$\tilde{u}_r = \frac{V_{r-}}{V_s \beta(r, S_{r-})}\frac{1}{s}\mathbf{1}_{\{r \leq s\}}.$$

Note that the process $V_-/\beta(\,\cdot\,, S_-)$ is predictable. By the first condition in (3.7) it turns out that this process is in $\operatorname{Dom}\delta^{(0)}$. Moreover by Lemma 3.5 and the chain rule it holds that $1/V_s$ is in $\mathbb{D}^{(0)}$. The second part of condition (3.7) allows us to conclude that $\tilde{u}$ is in $\operatorname{Dom}\delta^{(0)}$.

The first weight we calculate is the Skorohod integral of $\tilde{u}$. Thereto we perform integration by part,

$$\delta^{(0)}(\tilde{u}) = \frac{1}{V_s}\int_0^T \frac{V_{r-}}{\beta(r, S_{r-})s}\mathbf{1}_{\{r \leq s\}}dW_r - \int_0^T \frac{V_{r-}}{\beta(r, S_{r-})s}\mathbf{1}_{\{r \leq s\}}D_r^{(0)}\frac{1}{V_s}dr.$$

Because of the chain rule we can rewrite this as

$$\delta^{(0)}(\tilde{u}) = \frac{1}{sV_s}\int_0^s \frac{V_{r-}}{\beta(r, S_{r-})}dW_r + \frac{1}{s}\int_0^s \frac{V_{r-}}{\beta(r, S_{r-})}\frac{D_r^{(0)}V_s}{V_s^2}dr.$$

Now we make use of Lemma 3.5 and obtain that the latter equals

$$\frac{1}{sV_s}\int_0^s \frac{V_{r-}}{\beta(r, S_{r-})}dW_r + \frac{1}{s}\int_0^s \frac{V_{r-}}{\beta(r, S_{r-})}\frac{1}{V_s^2}\Big[\beta_x(r, S_{r-})V_s - \frac{\beta(r, S_{r-})\zeta_{r-}V_s}{V_{r-}^2} + \frac{\beta(r, S_{r-})\zeta_s}{V_{r-}}\Big]dr$$
$$= \frac{1}{sV_s}\Big(\int_0^s \frac{V_{r-}}{\beta(r, S_{r-})}dW_r + \int_0^s \Big[\frac{\beta_x(r, S_{r-})}{\beta(r, S_{r-})}V_{r-} - \frac{\zeta_{r-}}{V_{r-}} + \frac{\zeta_s}{V_s}\Big]dr\Big)$$

$$= \frac{1}{sV_s}\Big(\int_0^s \frac{V_{r-}}{\beta(r,S_{r-})}dW_r + s\frac{\zeta_s}{V_s} + \int_0^s \Big[\frac{\beta_x(r,S_{r-})}{\beta(r,S_{r-})}V_{r-} - \frac{\zeta_{r-}}{V_{r-}}\Big]dr\Big),$$

which is the mentioned expression for π_1.
The second weight in (3.2) is

$$\int_0^T \tilde{u}_r D_r^{(0)} S_t dr = \int_0^T \frac{1}{V_s}\frac{V_{r-}}{\beta(r,S_{r-})s}\mathbf{1}_{\{r\le s\}}V_t(V_{r-})^{-1}\beta(r,S_{r-})\mathbf{1}_{\{r\le t\}}dr = \int_0^s \frac{V_t}{sV_s}dr = \frac{V_t}{V_s}.$$

Theorem 3.4 can also be applied in this setting, which is interesting in case of non-differentiable functions f.

Proposition 3.7. *Consider again the setting defined by the stochastic differential equation* (3.5)*. For any function f as in Assumption 3.3 with $F = S_t$, $0 < s < t < T$, and $\alpha \in \mathbb{R}$ it holds, under conditions* (3.7)*, that*

$$\mathbb{E}[f(S_t)|S_s = \alpha] = \frac{\mathbb{E}[f(S_t)H(S_s-\alpha)\pi]}{\mathbb{E}[H(S_s-\alpha)\pi]},$$

where the Malliavin weight π differs from π_1 in (3.9) *as follows*

$$\pi = \pi_1 - \frac{1}{t-s}\frac{1}{V_s}\int_s^t \frac{V_{r-}}{\beta(r,S_{r-})}dW_r. \tag{3.10}$$

Proof. For the application of Theorem 3.4, we need the process

$$\hat{u}_r = \frac{V_{r-}}{V_s\beta(r,S_{r-})}\Big\{\frac{1}{s}\mathbf{1}_{\{r\le s\}} - \frac{1}{t-s}\mathbf{1}_{\{s\le r\le t\}}\Big\} = \tilde{u}_r - \frac{V_{r-}}{V_s\beta(r,S_{r-})}\frac{1}{t-s}\mathbf{1}_{\{s\le r\le t\}}. \tag{3.11}$$

By comparing this with the intermediate process used in the proof of Proposition 3.6, we conclude that $\hat{u}$ is in $\operatorname{Dom}\delta^{(0)}$. Moreover by the integration by parts formula and the fact that $V_-/\beta(\,\cdot\,,S_-)$ is predictable, we obtain

$$\begin{aligned}\delta^{(0)}(\hat{u}) &= \pi_1 + \delta^{(0)}\Big(-\frac{V_{r-}}{V_s\beta(r,S_{r-})}\frac{1}{t-s}\mathbf{1}_{\{s\le r\le t\}}\Big)\\ &= \pi_1 - \frac{1}{V_s}\frac{1}{t-s}\int_0^T \frac{V_{r-}}{\beta(r,S_{r-})}\mathbf{1}_{\{s\le r\le t\}}dW_r + \frac{1}{t-s}\int_0^T \frac{V_{r-}}{\beta(r,S_{r-})}\mathbf{1}_{\{s\le r\le t\}}D_r^{(0)}\frac{1}{V_s}dr,\end{aligned}$$

where π_1 is defined in (3.9). The last term equals zero, since by Lemma 3.5 the Malliavin derivative $D_r^{(0)}(1/V_s)$ introduces a factor $\mathbf{1}_{\{r\le s\}}$. This concludes the proof.

In the sequel we present two examples to illustrate our results from Propositions 3.6 and 3.7. The first example considers a linear SDE and the second concerns stochastic volatility models.

Example 3.8. Linear SDE. *We consider the following linear SDE*

$$\begin{cases} dS_t = \mu S_{t-}dt + \beta S_{t-}dW_t + \int_{\mathbb{R}_0}(\mathrm{e}^z-1)S_{t-}\widetilde{N}(dt,dz),\\ S_0 = s_0 > 0,\end{cases}$$

where μ and $\beta > 0$ are constants. We assume that $\int_{\mathbb{R}_0}(\mathrm{e}^z-1)^2\ell(dz) < \infty$, so that all assumptions imposed on model (3.5) *are satisfied. In this particular example, the first variation process V equals $V = S/s_0$ and $\zeta \equiv 0$ and conditions* (3.7) *are fulfilled. From Proposition 3.6 we find that the expression* (3.8) *holds with*

$$\pi_1 = \frac{s_0}{sS_s}\Big(\int_0^s \frac{1}{s_0\beta}dW_r + \int_0^s \frac{1}{s_0}dr\Big) = \frac{s_0}{sS_s}\Big(\frac{W_s}{s_0\beta} + \frac{s}{s_0}\Big) = \frac{1}{S_s}\Big(\frac{W_s}{s\beta} + 1\Big), \tag{3.12}$$

and

$$\pi_2 = \frac{S_t/s_0}{S_s/s_0} = \frac{S_t}{S_s}.$$

Substitution of the expressions for π_1 and π_2 into (3.8) leads to

$$\mathbb{E}[f(S_t)|S_s=\alpha] = \frac{\mathbb{E}\Big[f(S_t)H(S_s-\alpha)\frac{1}{S_s}\Big(\frac{W_s}{s\beta}+1\Big) - f'(S_t)H(S_s-\alpha)\frac{S_t}{S_s}\Big]}{\mathbb{E}\Big[H(S_s-\alpha)\frac{1}{S_s}\Big(\frac{W_s}{s\beta}+1\Big)\Big]},$$

where f satisfies Assumption 3.1, $0<s<t<T$, and $\alpha\in\mathbb{R}$.

On the other hand we can apply Proposition 3.7 for the linear SDE we are observing now. The weight π differs from the weight π_1 in Proposition 3.6, when the intermediate process is of the form (3.11), only by the second term in (3.10). In the present setting this term equals

$$\frac{s_0}{S_s}\Big(-\frac{1}{t-s}\Big)\int_s^t \frac{1}{s_0\beta}dW_r = -\frac{1}{\beta S_s}\frac{W_t-W_s}{t-s}.$$

Hence combining this with (3.12) gives

$$\pi = \frac{1}{S_s}\Big(\frac{W_s}{s\beta} - \frac{1}{\beta}\frac{W_t-W_s}{t-s}+1\Big) = \frac{1}{S_s}\Big(\frac{tW_s-sW_t}{s(t-s)\beta}+1\Big).$$

For any function f as in Assumption 3.3 with $F=S_t$, $0<s<t<T$, and $\alpha\in\mathbb{R}$ the conditional expectation can be rewritten as

$$\mathbb{E}[f(S_t)|S_s=\alpha] = \frac{\mathbb{E}\Big[f(S_t)H(S_s-\alpha)\frac{1}{S_s}\Big(\frac{tW_s-sW_t}{s(t-s)\beta}+1\Big)\Big]}{\mathbb{E}\Big[H(S_s-\alpha)\frac{1}{S_s}\Big(\frac{tW_s-sW_t}{s(t-s)\beta}+1\Big)\Big]}.$$

Example 3.9. Stochastic volatility models. *Let us consider the following model*

$$\begin{cases} dS_t = \mu S_{t-}dt + v(Y_{t-})S_{t-}dW_t^{(1)} + \int_{\mathbb{R}_0}(\mathrm{e}^z-1)S_{t-}\widetilde{N}(dt,dz),\\ dY_t = a(t,Y_{t-})dt + b(t,Y_{t-})dW_t^{(2)} + \int_{\mathbb{R}_0}\psi(z)\widetilde{N}(dt,dz),\end{cases} \tag{3.13}$$

with $S_0=s_0>0$ and $Y_0>0$. Herein $\widetilde{N}$ is the jump measure of a compound Poisson process with Lévy measure ℓ, and $W^{(1)}$ and $W^{(2)}$ are two correlated standard Brownian motions with

$$dW_t^{(1)}dW_t^{(2)} = \rho dt, \quad \rho\in(-1,1). \tag{3.14}$$

Moreover $\mu\in\mathbb{R}$, the functions a and b on $[0,T]\times\mathbb{R}$ are Lipschitz continuous and differentiable in the second argument for all t, v is a positive function which is Lipschitz continuous and differentiable on $\mathbb{R}$, $\int_{\mathbb{R}_0}(\mathrm{e}^z-1)^2\ell(dz)<\infty$, and ψ is a function on $\mathbb{R}$ such that $\int_{\mathbb{R}_0}\psi^2(z)\ell(dz)<\infty$. The process S may then perform the role of the stock price process, while $v(Y)$ can be interpreted as the stochastic volatility process. In many stochastic volatility models, the volatility $v(Y)$ equals $\sqrt{Y}$ and some conditions should be included to guarantee the non-negativity of the process Y. Some interesting examples are the Bates model (see [8]) and the Ornstein-Uhlenbeck stochastic volatility model (see [5, 6]).

From (3.14) we know there exists a Brownian motion $\widetilde{W}$, independent of $W^{(2)}$, such that we can express $W^{(1)}$ in terms of $\widetilde{W}$ and $W^{(2)}$ by

$$W_t^{(1)} = \rho W_t^{(2)} + \sqrt{1-\rho^2}\,\widetilde{W}_t.$$

Using the notations of Propositions 3.6 and 3.7, where we consider the Malliavin derivative in the direction of the Brownian motion $\widetilde{W}$, we have

$$V_t = \frac{S_t}{s_0}, \quad \beta(t,S_{t-}) = v(Y_{t-})S_{t-}\sqrt{1-\rho^2}, \quad \text{and} \quad \zeta_t = 0. \tag{3.15}$$

Applying Proposition 3.6, we find for the weights in representation (3.8)

$$\begin{aligned}\pi_1 &= \frac{s_0}{sS_s}\Big(\int_0^s \frac{S_{r-}/s_0}{v(Y_{r-})S_{r-}\sqrt{1-\rho^2}}d\widetilde{W}_r + \int_0^s \frac{v(Y_{r-})\sqrt{1-\rho^2}}{v(Y_{r-})S_{r-}\sqrt{1-\rho^2}}\frac{S_{r-}}{s_0}dr\Big)\\ &= \frac{s_0}{sS_s}\Big(\frac{1}{s_0\sqrt{1-\rho^2}}\int_0^s \frac{d\widetilde{W}_r}{v(Y_{r-})} + \frac{s}{s_0}\Big) = \frac{1}{S_s}\Big(\frac{1}{s\sqrt{1-\rho^2}}\int_0^s \frac{d\widetilde{W}_r}{v(Y_{r-})} + 1\Big)\\ &= \frac{1}{S_s}\Big(\frac{1}{s(1-\rho^2)}\Big\{\int_0^s \frac{dW_r^{(1)}}{v(Y_{r-})} - \rho\int_0^s \frac{dW_r^{(2)}}{v(Y_{r-})}\Big\} + 1\Big)\end{aligned}$$

and $\pi_2 = \dfrac{S_t/s_0}{S_s/s_0} = \dfrac{S_t}{S_s}$. *When we prefer not to use the derivative of the function* f*, we can apply Proposition 3.7. The weight is then given by*

$$\begin{aligned}\pi &= \pi_1 - \frac{1}{S_s}\frac{1}{(t-s)\sqrt{1-\rho^2}}\int_s^t \frac{d\widetilde{W}_r}{v(Y_{r-})}\\ &= \pi_1 - \frac{1}{S_s}\frac{1}{(t-s)(1-\rho^2)}\Big\{\int_s^t \frac{dW_r^{(1)}}{v(Y_{r-})} - \rho\int_s^t \frac{dW_r^{(2)}}{v(Y_{r-})}\Big\}.\end{aligned}$$

Considering the model (3.13), we derived a representation for $\mathbb{E}[f(S_t)|S_s=\alpha]$. In the sequel we observe the conditional expectation

$$\mathbb{E}[w(Y_T)|S_T=\alpha], \tag{3.16}$$

for a certain function $w:\mathbb{R}\mapsto\mathbb{R}_0$. Our motivation to consider the latter expression comes from a paper by [28], where the authors are interested in the computation of conditional moments of Y. Thus they consider (3.16) for $w(x)=x$ and $w(x)=x^2$. Moreover, in [35] the authors consider (3.16) for $w(x)=v^2(x)$, which is interesting for the study of stochastic local volatility.

We consider model (3.13) and a function w. It is clear that $D_r^{(0)}Y_t=0$ since Y only depends on $W^{(2)}$, which is independent of $\widetilde{W}$. Thus condition (3.3) is satisfied for any process u in $\operatorname{Dom}\delta^{(0)}$. Thus when condition (3.1) is fulfilled, the conditional expectation can be written in the form (3.4). From expression (3.6) and previous derivations (3.15) we deduce that

$$D_r^{(0)}(S_T-\alpha) = S_T v(Y_{r-})\sqrt{1-\rho^2},\ \text{for } r\le T.$$

Therefore the process satisfying condition (3.1) is given by

$$u_r = \big(TS_T v(Y_{r-})\sqrt{1-\rho^2}\big)^{-1}.$$

The Skorohod integral of this process is computed similarly as in the proof of Proposition 3.6 and it equals

$$\delta^{(0)}(u) = \frac{1}{TS_T\sqrt{1-\rho^2}}\int_0^T \frac{d\widetilde{W}_r}{v(Y_{r-})} - \int_0^T \frac{1}{Tv(Y_{r-})\sqrt{1-\rho^2}}D_r^{(0)}\Big(\frac{1}{S_T}\Big)dr.$$

By the chain rule, the second term in the last equation equals

$$\int_0^T \frac{1}{Tv(Y_{r-})\sqrt{1-\rho^2}}\frac{D_r^{(0)}S_T}{S_T^2}dr = \int_0^T \frac{1}{TS_T}dr = \frac{1}{S_T}.$$

Finally we conclude that

$$\mathbb{E}[w(Y_T)|S_T=\alpha] = \frac{\mathbb{E}\Big[w(Y_T)H(S_T-\alpha)\frac{1}{S_T}\Big(\frac{1}{T\sqrt{1-\rho^2}}\int_0^T \frac{d\widetilde{W}_r}{v(Y_{r-})}+1\Big)\Big]}{\mathbb{E}\Big[H(S_T-\alpha)\frac{1}{S_T}\Big(\frac{1}{T\sqrt{1-\rho^2}}\int_0^T \frac{d\widetilde{W}_r}{v(Y_{r-})}+1\Big)\Big]}.$$

Via the Malliavin method we can also deduce a representation for the delta in terms of unconditional expectations.

Theorem 3.10. *Consider the same setting as in Theorem 3.4 and assume that the Skorohod integral $\delta^{(0)}(u)$ is $\sigma(F,G)$-measurable. Then the delta is given by*

$$\frac{\partial}{\partial\alpha}\mathbb{E}[f(F)|G=\alpha] = \frac{\mathbb{B}_{F,G}[f](\alpha)\mathbb{A}_{F,G}[1](\alpha) - \mathbb{A}_{F,G}[f](\alpha)\mathbb{B}_{F,G}[1](\alpha)}{\mathbb{A}_{F,G}[1](\alpha)^2}, \tag{3.17}$$

where

$$\mathbb{A}_{F,G}[\cdot](\alpha) = \mathbb{E}[\cdot(F)H(G-\alpha)\delta^{(0)}(u)],$$
$$\mathbb{B}_{F,G}[\cdot](\alpha) = \mathbb{E}[\cdot(F)H(G-\alpha)\{-\delta^{(0)}(u)^2 + \int_0^T u_r D_r^{(0)}\delta^{(0)}(u)dr\}].$$

Proof. The structure of formula (3.17) follows clearly from the derivation of representation (3.4). Now we focus on the derivative

$$\mathbb{B}_{F,G}[f](\alpha) = \frac{\partial}{\partial\alpha}\mathbb{E}[f(F)H(G-\alpha)\delta^{(0)}(u)] = -\mathbb{E}[f(F)\delta_0(G-\alpha)\delta^{(0)}(u)].$$

By relation (3.1), the chain rule, the duality formula, and the integration by parts formula, we obtain

$$\begin{aligned}
&\mathbb{E}[f(F)\delta_0(G-\alpha)\delta^{(0)}(u)]\\
&= \mathbb{E}\Big[f(F)\delta_0(G-\alpha)\delta^{(0)}(u)\mathbb{E}\big[\int_0^T u_r D_r^{(0)}G dr|\sigma(F,G)\big]\Big]\\
&= \mathbb{E}\Big[\mathbb{E}\big[f(F)\delta_0(G-\alpha)\delta^{(0)}(u)\int_0^T u_r D_r^{(0)}G dr|\sigma(F,G)\big]\Big]\\
&= \mathbb{E}\Big[f(F)\delta_0(G-\alpha)\delta^{(0)}(u)\int_0^T u_r D_r^{(0)}G dr\Big]\\
&= \mathbb{E}\Big[\int_0^T f(F)\delta^{(0)}(u)u_r D_r^{(0)}H(G-\alpha)dr\Big]\\
&= \mathbb{E}\Big[H(G-\alpha)\delta^{(0)}\big(f(F)\delta^{(0)}(u)u\big)\Big]\\
&= \mathbb{E}\Big[H(G-\alpha)\Big\{f(F)\delta^{(0)}\big(\delta^{(0)}(u)u\big) - \int_0^T \delta^{(0)}(u)u_r D_r^{(0)}\big(f(F)\big)dr\Big\}\Big]\\
&= \mathbb{E}\Big[H(G-\alpha)\Big\{f(F)\{\delta^{(0)}(u)\delta^{(0)}(u) - \int_0^T u_r D_r^{(0)}\delta^{(0)}(u)dr\} - \delta^{(0)}(u)f'(F)\int_0^T u_r D_r^{(0)}F dr\Big\}\Big]\\
&= \mathbb{E}\Big[f(F)H(G-\alpha)\Big\{\delta^{(0)}(u)^2 - \int_0^T u_r D_r^{(0)}\delta^{(0)}(u)dr\Big\}\Big]\\
&\qquad - \mathbb{E}\Big[f'(F)H(G-\alpha)\delta^{(0)}(u)\int_0^T u_r D_r^{(0)}F dr\Big].
\end{aligned}$$

By expression (3.3), the latter expectation equals

$$\begin{aligned}
&\mathbb{E}\Big[f'(F)H(G-\alpha)\delta^{(0)}(u)\int_0^T u_r D_r^{(0)}F dr\Big]\\
&= \mathbb{E}\Big[\mathbb{E}\big[f'(F)H(G-\alpha)\delta^{(0)}(u)\int_0^T u_r D_r^{(0)}F dr|\sigma(F,G)\big]\Big]\\
&= \mathbb{E}\Big[f'(F)H(G-\alpha)\delta^{(0)}(u)\mathbb{E}\big[\int_0^T u_r D_r^{(0)}F dr|\sigma(F,G)\big]\Big] = 0.
\end{aligned}$$

Hence we conclude that

$$\mathbb{E}[f(F)\delta_0(G-\alpha)\delta^{(0)}(u)] = \mathbb{E}\Big[f(F)H(G-\alpha)\Big\{\delta^{(0)}(u)^2 - \int_0^T u_r D_r^{(0)}\delta^{(0)}(u)dr\Big\}\Big].$$

Similar to the results in the conditional density method case, we obtained several useful representations for conditional expectations and their derivatives in terms of unconditional expectations via the Malliavin method.

4. Variance reduction

In the representations considered in the previous sections the random variables whose expectation should be estimated can have a large variance. To obtain a smaller variance and satisfactory convergence results in the context of Monte Carlo simulations, one might include variance reduction techniques. In subsection 4.1 we study the *localisation* technique. This technique was used in [4]. We adapt it here to our conditional expectation representations. Moreover *control variables* may be included to reduce the variance. We handle this approach in subsection 4.2.

4.1. Localisation

We adapt the localisation technique of [4] for both methods; the conditional density method and the Malliavin method.

Proposition 4.1. *Assume the setting of Theorem 2.3. Then for any function* $\psi : \mathbb{R} \mapsto [0,\infty)$ *satisfying* $\int_{\mathbb{R}} \psi(t)dt = 1$ *and for all* $\alpha \in \mathbb{R}$*, we have*

$$\mathbb{E}[f(F)|G=\alpha] = \frac{\mathcal{J}^{\psi}_{F,G}[f](\alpha)}{\mathcal{J}^{\psi}_{F,G}[1](\alpha)},$$

where $\mathcal{J}^{\psi}_{F,G}[\cdot](\alpha)$ *is given by*

$$\mathcal{J}^{\psi}_{F,G}[\cdot](\alpha) = \mathbb{E}\Big[\cdot(F)\Big(\psi(G-\alpha)\frac{\partial}{\partial u}g_2(U,V) + \pi_{(X,U)}\big[H(G-\alpha)-\Psi(G-\alpha)\big]\Big)\Big]$$

where $\Psi(x) = \int_{-\infty}^{x}\psi(t)dt$.

Proof. For the numerator of representation (2.3) it holds that

$$\begin{aligned}\mathbb{E}[f(F)H(G-\alpha)\pi_{(X,U)}] &= \mathbb{E}[f(F)\psi(G-\alpha)\frac{\partial}{\partial u}g_2(U,V)] + \mathbb{E}[f(F)H(G-\alpha)\pi_{(X,U)}]\\ &\quad - \mathbb{E}[f(F)\Psi'(G-\alpha)\frac{\partial}{\partial u}g_2(U,V)].\end{aligned}$$

The last term equals

$$\mathbb{E}[f(F)\Psi'(G-\alpha)\frac{\partial}{\partial u}g_2(U,V)] = \mathbb{E}\Big[\int_{\mathbb{R}^2} f(g_1(x,Y))\Psi'(g_2(u,V)-\alpha)\frac{\partial}{\partial u}g_2(u,V)p(x,u)dxdu\Big].$$

Using the integration by parts formula, we get

$$\begin{aligned}&\mathbb{E}[f(F)\Psi'(G-\alpha)\frac{\partial}{\partial u}g_2(U,V)]\\ &= \mathbb{E}\Big[-\int_{\mathbb{R}^2} f(g_1(x,Y))\Psi(g_2(u,V)-\alpha)\frac{\partial}{\partial u}p(x,u)dxdu\Big]\\ &= \mathbb{E}\Big[\int_{\mathbb{R}^2} f(g_1(x,Y))\Psi(g_2(u,V)-\alpha)\Big(-\frac{\partial}{\partial u}\log p(x,u)\Big)p(x,u)dxdu\Big]\\ &= \mathbb{E}[\mathbb{E}[f(F)\Psi(G-\alpha)\pi_{(X,U)}|\sigma(Y,V)]] = \mathbb{E}[f(F)\Psi(G-\alpha)\pi_{(X,U)}],\end{aligned}$$

and the result follows.

KPMG Center of Excellence in Risk Management

Proposition 4.2. *Assume the setting of Theorem 3.4, then for any function* $\psi : \mathbb{R} \mapsto [0, \infty)$ *satisfying* $\int_{\mathbb{R}} \psi(t)dt = 1$ *and for all* $\alpha \in \mathbb{R}$, *we have*

$$\mathbb{E}[f(F)|G = \alpha] = \frac{\mathcal{J}^{\psi}_{F,G}[f](\alpha)}{\mathcal{J}^{\psi}_{F,G}[1](\alpha)},$$

where $\mathcal{J}^{\psi}_{F,G}[\cdot](\alpha)$ *is given by*

$$\mathcal{J}^{\psi}_{F,G}[\cdot](\alpha) = \mathbb{E}\Big[\cdot(F)\Big(\psi(G-\alpha) + \delta^{(0)}(u)\big[H(G-\alpha) - \Psi(G-\alpha)\big]\Big)\Big]$$

where $\Psi(x) = \int_{-\infty}^{x} \psi(t)dt$.

Proof. For the numerator of representation (3.4) it holds that

$$\begin{aligned}\mathbb{E}[f(F)H(G-\alpha)\delta^{(0)}(u)] = \mathbb{E}[f(F)\psi(G-\alpha)] + \mathbb{E}[f(F)H(G-\alpha)\delta^{(0)}(u)] \\ - \mathbb{E}[f(F)\Psi'(G-\alpha)].\end{aligned}$$

Applying the same arguments as in Theorem 3.4, we can show that the last term equals

$$\mathbb{E}[f(F)\Psi'(G-\alpha)] = \mathbb{E}[f(F)\Psi(G-\alpha)\delta^{(0)}(u)]$$

and the result follows.

Once we have introduced the localised versions of the representation formulas for the conditional expectation, one natural question arises, namely what is the optimal choice of the localising function ψ. To find this optimal function, we assume that the additional constant c in the function H is zero, i.e. $H(x) = \mathbf{1}_{\{x \geq 0\}}$. Let Z represent either the factor $\frac{\partial}{\partial u} g_2(U, V)$ in case of the conditional density method or the factor 1 when the Malliavin method is considered. Then, practically speaking, an expectation of the form

$$\mathcal{J}^{\psi}_{F,G}[\cdot](\alpha) = \mathbb{E}\Big[\cdot(F)\Big(\psi(G-\alpha)Z + \pi\big[H(G-\alpha) - \Psi(G-\alpha)\big]\Big)\Big]$$

is estimated via Monte Carlo simulation. More precisely if we denote by N the number of simulated values of F and G, we have the following estimation

$$\mathcal{J}^{\psi}_{F,G}[\cdot](\alpha) \approx \frac{1}{N}\sum_{q=1}^{N} \cdot(F^q)\Big(\psi(G^q-\alpha)Z^q + \pi^q\big[H(G^q-\alpha) - \Psi(G^q-\alpha)\big]\Big).$$

In order to reduce the variance, the idea is to minimise the integrated mean squared error with respect to the localising function ψ. Thus we have to solve the following optimisation problem (this criterion has been introduced by [25])

$$\inf_{\psi \in \mathcal{L}^1} I(\psi), \tag{4.1}$$

where $\mathcal{L}^1 = \{\psi : \mathbb{R} \mapsto [0, \infty) : \psi \in C^1(\mathbb{R}), \psi(+\infty) = 0, \int_{\mathbb{R}} \psi(t)dt = 1\}$ and I equals the integrated variance up to a constant (in terms of ψ)

$$I(\psi) = \int_{\mathbb{R}} \mathbb{E}[\cdot^2(F)\Big(\psi(G-\alpha)Z + \pi\big[H(G-\alpha) - \Psi(G-\alpha)\big]\Big)^2]d\alpha. \tag{4.2}$$

The choice of the optimal localising function ψ is given in the following proposition. It is obvious that the optimal localisation function will be different for the numerator and denominator since the optimisation problem is different. (The proof in [4] can easily be extended to the current setting.)

Proposition 4.3. *The infimum of the optimisation problem* (4.1) *with* $I(\psi)$ *given by* (4.2) *and* $H(x) = \mathbf{1}_{\{x\geq 0\}}$*, is reached at* $\psi^\star$*, where* $\psi^\star$ *is the probability density of the Laplace distribution with parameter* $\lambda^\star$*, i.e. for all* $t \in \mathbb{R}$*,* $\psi^\star(t) = \frac{\lambda^\star}{2}\mathrm{e}^{-\lambda^\star|t|}$*, where*

$$\lambda^\star = \Big(\frac{\mathbb{E}[\cdot^2(F)\pi^2]}{\mathbb{E}[\cdot^2(F)Z^2]}\Big)^{\frac{1}{2}}. \tag{4.3}$$

The localising function defined in the previous proposition is optimal in the sense of minimal variance, however it is not optimal in numerical experiments when it comes to the computational effort. Therefore [13] considered the exponential localising function

$$\psi(x) = \lambda^* \mathrm{e}^{-\lambda^* x}\mathbf{1}_{\{x\geq 0\}}, \tag{4.4}$$

where λ^* is given by (4.3). In paragraph 5.2.5 we show how the use of this function reduces the computational effort. We perform numerical experiments for both localising functions in Section 5.

The representations for the derivatives in Theorems 2.11 and 3.10 have a localised version too. We state the localised versions as well as the choice of the optimal localising function ψ in the following propositions. We do not present the proofs since they follow along similar lines as Propositions 4.1, 4.2, and 4.3.

Proposition 4.4. *Assume the setting of Theorem 2.11, then for any function* $\psi : \mathbb{R} \mapsto [0,\infty)$ *satisfying* $\int_{\mathbb{R}} \psi(t)dt = 1$ *and for all* $\alpha \in \mathbb{R}$*, we have*

$$\mathbb{B}_{F,G}[\cdot](\alpha) = \mathbb{E}\Big[\cdot(F)\Big(\psi(G-\alpha)(-\pi)Z + (-\pi^2+\pi^*)\big[H(G-\alpha) - \Psi(G-\alpha)\big]\Big)\Big]$$

where $\Psi(x) = \int_{-\infty}^{x}\psi(t)dt$*,*

$$Z = \frac{\partial}{\partial u} g_2(U,V), \quad \pi = \pi_{(X,U)}, \quad \text{and} \quad \pi^* = \pi^*_{(X,U)}.$$

Proposition 4.5. *Assume the setting of Theorem 3.10, then for any function* $\psi : \mathbb{R} \mapsto [0,\infty)$ *satisfying* $\int_{\mathbb{R}} \psi(t)dt = 1$ *and for all* $\alpha \in \mathbb{R}$*, we have*

$$\mathbb{B}_{F,G}[\cdot](\alpha) = \mathbb{E}\Big[\cdot(F)\Big(\psi(G-\alpha)(-\pi)Z + (-\pi^2+\pi^*)\big[H(G-\alpha) - \Psi(G-\alpha)\big]\Big)\Big]$$

where $\Psi(x) = \int_{-\infty}^{x}\psi(t)dt$*,*

$$Z = 1, \quad \pi = \delta^{(0)}(u), \quad \text{and} \quad \pi^* = \int_0^T u_r D_r^{(0)}\delta^{(0)}(u)dr.$$

The optimal localising functions minimise the integrated variance

$$\widetilde{I}(\psi) = \int_{\mathbb{R}} \mathbb{E}[\cdot^2(F)\Big(\psi(G-\alpha)(-\pi)Z + (-\pi^2+\pi^*)\big[H(G-\alpha) - \Psi(G-\alpha)\big]\Big)^2]d\alpha. \tag{4.5}$$

Proposition 4.6. *The infimum of the optimisation problem* $\inf_{\psi\in\mathcal{L}^1}\widetilde{I}(\psi)$*, with* $\widetilde{I}(\psi)$ *given by* (4.5)*, where* $H(x) = \mathbf{1}_{\{x\geq 0\}}$*, is reached at* $\widetilde{\psi}$*, where* $\widetilde{\psi}$ *is the probability density of the Laplace distribution with parameter* $\widetilde{\lambda}$*, i.e. for all* $t\in\mathbb{R}$*,* $\widetilde{\psi}(t) = \frac{\widetilde{\lambda}}{2}\mathrm{e}^{-\widetilde{\lambda}|t|}$*, where*

$$\widetilde{\lambda} = \Big(\frac{\mathbb{E}[\cdot^2(F)(-\pi^2+\pi^*)^2]}{\mathbb{E}[\cdot^2(F)\pi^2 Z^2]}\Big)^{\frac{1}{2}}.$$

4.2. Control variable

Another approach to obtain variance reduction (besides localisation) is to include a *control variable*, see e.g. Section 4.1 in [21]. The advantage of adding a control variable is to use the observed error in estimating a known quantity to adjust an estimator for an unknown quantity. In case of American option pricing, the control variable can be the corresponding European option price. The price of the American and respectively the European option with maturity T and payoff function Φ, on an asset with value α at time t is denoted by $P(t,\alpha)$, respectively $P^{\mathrm{Eu}}(t,\alpha)$. Let us define the function $P_\gamma(t,\alpha) := P(t,\alpha) - \gamma P^{\mathrm{Eu}}(t,\alpha)$, for a real number γ close to 1. Then it holds that

$$P_\gamma(t,\alpha) = \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}_{t,\alpha}\big[\mathrm{e}^{-\int_t^\tau r_u du}\big\{\Phi(S_\tau) - \gamma P^{\mathrm{Eu}}(\tau, S_\tau)\big\}\big],$$

where $\mathcal{T}_{t,T}$ denotes the set of all stopping times in $[t,T]$. The price of the American option at time 0 is given by $P(0,s_0) = P_\gamma(0,s_0) + \gamma P^{\mathrm{Eu}}(0,s_0)$ and its delta equals $\Delta(0,s_0) = \Delta_\gamma(0,s_0) + \gamma \Delta^{\mathrm{Eu}}(0,s_0)$. We can rewrite this formula for the American option price as

$$P(0,s_0) = \sup_{\tau \in \mathcal{T}_{0,T}} \mathbb{E}\big[\mathrm{e}^{-\int_0^\tau r_u du}\Phi(S_\tau) - \gamma\big\{\mathrm{e}^{-\int_0^\tau r_u du}P^{\mathrm{Eu}}(\tau,S_\tau) - P^{\mathrm{Eu}}(0,s_0)\big\}\big].$$

From this expression, the advantage of adding a control variable is clear. Indeed, the error between $P^{\mathrm{Eu}}(0,s_0)$ and an estimation of $\mathbb{E}[\mathrm{e}^{-\int_0^\tau r_u du}P^{\mathrm{Eu}}(\tau,S_\tau)]$ for each $\tau \in \mathcal{T}_{0,T}$ is used to adjust the estimation of the American option price $P(0,s_0) = \sup_{\tau\in\mathcal{T}_{0,T}} \mathbb{E}[\mathrm{e}^{-\int_0^\tau r_u du}\Phi(S_\tau)]$.

5. Numerical experiments

In this section we apply our results to estimate the price and delta of European options at times $t > 0$ and of American options at time zero. Then we illustrate our methods with numerical results in a specified jump-diffusion model. European options are considered to evaluate the accuracy of our representations since there are analytic formulas at hand in the Merton model, whereas there are non for American options.

5.1. General approach to determine European and American option prices and deltas

5.1.1. Prices and deltas of European options

European options may only be executed at time of maturity T. However they can be traded at any moment between time 0 and T. Consider the risk-free interest rate r and the underlying stock price process S, then the price at time $t > 0$ of a European option with payoff function Φ equals

$$P^{\mathrm{Eu}}(t,S_t) = \mathrm{e}^{-r(T-t)}\mathbb{E}[\Phi(S_T)|S_t]. \tag{5.1}$$

The delta at time t equals

$$\Delta^{\mathrm{Eu}}(t,S_t) = \mathrm{e}^{-r(T-t)}\frac{\partial}{\partial S}\mathbb{E}[\Phi(S_T)|S_t = S]\Big|_{S=S_t}. \tag{5.2}$$

5.1.2. Algorithm to estimate prices and deltas of American options

American options can be executed at any time prior to maturity. Since it is practically impossible to observe the possibility to execute the option at infinitely many times, an American option is often approximated by a Bermudan option with the same maturity and payoff function. To obtain this approximation, the time interval $[0,T]$ is discretised into n time periods with step size $\varepsilon = T/n$. The Bermudan option can then be executed at the n discrete times $iT/n, i = 1,\dots,n$. When the number of time periods increases, the Bermudan option converges to the American option (see [3]). Bermudan options can be priced through a Bellman dynamic programming principle, see [9] and [4]. Let Φ denote the payoff function and S the underlying stock price process with initial value s_0. Then the price of the Bermudan option $P(0,s_0)$ follows from the recursive computations

$$P(n\varepsilon, S_{n\varepsilon}) = \Phi(S_{n\varepsilon}) = \Phi(S_T),$$

$$P(k\varepsilon, S_{k\varepsilon}) = \max\Big\{\Phi(S_{k\varepsilon}), \mathrm{e}^{-r\varepsilon}\mathbb{E}\big[P((k+1)\varepsilon, S_{(k+1)\varepsilon})\big|S_{k\varepsilon}\big]\Big\}, \quad k = n-1,\ldots,1,0. \tag{5.3}$$

The sensitivity of the option price with respect to the initial value of the underlying asset, i.e. the delta of the option $\Delta(0, s_0) := \partial_{s_0}P(0, s_0)$, can be derived as follows

$$\begin{aligned} \Delta(\varepsilon, S_\varepsilon) &= \begin{cases} \mathrm{e}^{-r\varepsilon}\partial_\alpha\mathbb{E}[P(2\varepsilon, S_{2\varepsilon})|S_\varepsilon = \alpha]\Big|_{\alpha=S_\varepsilon} & \text{if } P(\varepsilon, S_\varepsilon) > \Phi(S_\varepsilon), \\ \partial_\alpha\Phi(\alpha)\Big|_{\alpha=S_\varepsilon} & \text{if } P(\varepsilon, S_\varepsilon) = \Phi(S_\varepsilon), \end{cases} \\ \Delta(0, s_0) &= \mathrm{e}^{-r\varepsilon}\mathbb{E}_{s_0}[\Delta(\varepsilon, S_\varepsilon)]. \end{aligned} \tag{5.4}$$

Hence to obtain a numerical estimation of the price and the delta at time zero, we proceed by estimating the prices and the deltas recursively and backwards in time. For estimations based on simulated values for the underlying stock price one can simulate the number of required paths N at the discrete time points and store them all before performing the recursive computations. On the other hand, since the pricing program and computations of the deltas go backwards in time, it is more convenient to simulate the stock price process simultaneously. Simulating the stock price process backwards in time too leads to more efficiency concerning memory capacity.

5.2. Implementations for put options in a Merton model

5.2.1. The setting

We consider European and American put options on a stock price process S defined by a Merton model. The put payoff function equals $\Phi(x) = (K-x)^+$. The stock price process S is modelled as

$$S_t = s_0 \exp\left(\left(r - \frac{\beta^2}{2}\right)t + \beta W_t + \sum_{i=1}^{N_t} Y_i\right), \tag{5.5}$$

where $r > 0$ is the risk-free interest rate, β is a positive constant, and W is a Wiener process. The jump part is determined by a Poisson process N with jump intensity μ and the random variables Y_i are i.i.d. with distribution $N(-\delta^2/2, \delta^2)$. Since we want to compare our results to the analysis in [1], we use the parameter setting as in his paper. That explains our choice of this specific connection between the jump mean and jump variance. This simplifies the Merton formula (5.8).

5.2.2. Representations

The conditional expectations and their derivatives in (5.1) - (5.4) can be estimated based on the representations we developed in the previous sections. In the present Merton setting, the representations obtained through Theorems 2.3, 3.4, 2.11, and 3.10 are as follows, for $0 < s < t < T$ and $\alpha \in \mathbb{R}$,

$$\begin{aligned} \mathbb{E}[f(S_t)|S_s = \alpha] &= \frac{\mathbb{E}[f(S_t)H(S_s-\alpha)\pi_{s,t}]}{\mathbb{E}[H(S_s-\alpha)\pi_{s,t}]} = \frac{\mathbb{A}_{t,s}[f](\alpha)}{\mathbb{A}_{t,s}[1](\alpha)}, \text{ and} \\ \frac{\partial}{\partial\alpha}\mathbb{E}[f(S_t)|S_s = \alpha] &= \frac{\mathbb{B}_{t,s}[f](\alpha)\mathbb{A}_{t,s}[1](\alpha) - \mathbb{A}_{t,s}[f](\alpha)\mathbb{B}_{t,s}[1](\alpha)}{\mathbb{A}_{t,s}[1](\alpha)^2}k(\alpha), \end{aligned} \tag{5.6}$$

where

$$\mathbb{B}_{t,s}[\cdot](\alpha) = \mathbb{E}[\cdot(S_t)H(S_s-\alpha)\{-\pi_{s,t}^2 + \pi_{s,t}^*\}].$$

Throughout this section we consider $H(x) = \mathbf{1}_{\{x\geq 0\}}$. Applying the conditional density method (CDM), it holds that $k(\alpha) = 1/\alpha$,

$$\pi_{s,t} = \frac{tW_s - sW_t}{\beta s(t-s)}, \quad \text{and} \quad \pi_{s,t}^* = \frac{t}{\beta^2 s(t-s)}.$$

For the Malliavin method (MM), we have $k(\alpha) = 1$,

$$\pi_{s,t} = \frac{1}{S_s}\left(\frac{tW_s - sW_t}{\beta s(t-s)} + 1\right), \quad \text{and} \quad \pi_{s,t}^* = -\frac{1}{S_s^2}\left(\frac{\beta(tW_s - sW_t) - t}{\beta^2 s(t-s)} + 1\right).$$

KPMG Center of Excellence in Risk Management

The regular expectations appearing in the representations (5.6) can easily be estimated by a Monte Carlo simulation. For example consider the estimation of the numerator of representation (5.6). We require N simulated values of S_t, S_s, and $\pi_{s,t}$, belonging to the same path. If we denote the j-th simulated values by S_t^j, S_s^j, and $\pi_{s,t}^j$, then we approximate

$$\mathbb{E}[\cdot(S_t)H(S_s-\alpha)\pi_{s,t}] \approx \frac{1}{N}\sum_{j=1}^{N} \cdot(S_t^j)H(S_s^j-\alpha)\pi_{s,t}^j. \tag{5.7}$$

5.2.3. Variance reduction techniques

As discussed in paragraph 4.1 we can include the localising technique. The estimation (5.7) is then replaced by

$$\mathbb{E}[\cdot(S_t)H(S_s-\alpha)\pi_{s,t}] \approx \frac{1}{N}\sum_{j=1}^{N} \cdot(S_t^j)\Big(\psi(S_s^j-\alpha)Z_s^j + \pi_{s,t}^j\big[H(S_s^j-\alpha)-\Psi(S_s^j-\alpha)\big]\Big),$$

where Z_s equals S_s in case of the CDM and 1 in case of the MM. The functions ψ and Ψ are defined by Proposition 4.3.

On the other hand we can include a control variable, see paragraph 4.2. For the estimation of the American option price $P(0,s_0)$ and delta $\Delta(0,s_0)$, we include the European option as a control variable. In the current setting, the European option price and delta can be obtained through Merton's approach. Consider the algorithm for the price of Bermudan options (5.3). To introduce the control variable we proceed in two steps. First we replace the put payoff function at each time $k\varepsilon$, $k=1,\dots,n$, by

$$\Phi_\gamma(k\varepsilon, S_{k\varepsilon}) = \Phi(S_{k\varepsilon}) - \gamma P^{\mathrm{Me}}(k\varepsilon, S_{k\varepsilon}),$$

where γ is a real number close to one and $P^{\mathrm{Me}}(k\varepsilon, S_{k\varepsilon})$ denotes the European option price, obtained through Merton's approach, at time $k\varepsilon$. Secondly, in the last step ($k=0$) we add $\gamma P^{\mathrm{Me}}(0,s_0)$ (respectively $\gamma\Delta^{\mathrm{Me}}(0,s_0)$) to obtain the American option price $P(0,s_0)$ (respectively the American option delta $\Delta(0,s_0)$).

The European option price in a Merton model is derived in [29] and is in our setting (5.5) for a put option given by the series

$$P^{\mathrm{Me}}(t,S_t) = \sum_{n=0}^{\infty} \frac{\mathrm{e}^{-\mu(T-t)}(\mu(T-t))^n}{n!} P_n^{\mathrm{BS}}(t,S_t). \tag{5.8}$$

Herein $P_n^{\mathrm{BS}}(t,S_t)$ is the Black-Scholes price of the European put option with the same maturity, strike, and interest rate r, and where the underlying stock price process has variance $v_n^2 = \beta^2 + n\delta^2/2$. The first 20 terms in the series are sufficient for a good approximation for the put option price.

5.2.4. Backward simulation

As remarked at the end of paragraph 5.1.2 the algorithm for the pricing of a Bermudan option goes backwards in time and we can simulate the different stochastic variables backwards in time too. For the Brownian motion we base the backward simulation on a Brownian bridge (see [4]). To simulate the compound Poisson process backwards in time, we base our method on results of [24] and [2]. We split the simulation of a compound Poisson process in the simulation of a Poisson process and in the simulation of the sum of the jump sizes. First we mention the following proposition implying a backward simulation algorithm for a Poisson process. This is covered by Lemma 3.1 in [2].

Proposition 5.1. *Let N be a Poisson process with intensity μ. For any time $t>0$ it holds that N_t has a* Poisson(μt) *distribution. Moreover for any $0<s<t$ it holds that N_s, conditioned on $N_t=z$, follows a* Binomial$(z, s/t)$ *distribution.*

Secondly we present the following proposition considering the (conditional) distribution of sums of independent and identically normal distributed variables. This result is a consequence of Brownian bridges, [24].

Proposition 5.2. *Consider the following sum*

$$C^{(k)} = \sum_{i=1}^{k} Y_i,$$

where Y_i are i.i.d. $N(\eta,\nu)$. For any $k>0$ it holds that $C^{(k)}$ has a $N(k\eta,k\nu)$ distribution. Moreover for any $0<j<k$ it holds that $C^{(j)}$, given that $C^{(k)}=z$, has a $N\big((j/k)z,(j/k)(k-j)\nu\big)$ distribution.

The backward simulation technique is interesting in numerical applications and following [2], this technique can also be derived for the Kou model, see [26].

5.2.5. Reduction of computational effort

[13] observed that the computational effort to estimate the American option prices by a Malliavin method can be reduced by sorting the estimated stock prices. Consider the Bermudan dynamic programming algorithm (5.3). For a fixed k in $\{n-1,\dots,1\}$ we estimate the conditional expectations for $q=1,\dots,N$ by our representations, including localisation, as follows

$$\mathbb{E}\big[P((k+1)\varepsilon,S_{(k+1)\varepsilon})\big|S_{k\varepsilon}=S^{(q)}_{k\varepsilon}\big] \approx \frac{\mathbb{J}[P_{(k+1)\varepsilon}](S^{(q)}_{k\varepsilon})}{\mathbb{J}[1](S^{(q)}_{k\varepsilon})},$$

where

$$\mathbb{J}[\cdot](\alpha) = \frac{1}{N}\sum_{j=1}^{N} \cdot^{(j)}\big(\psi(S^{(j)}_{k\varepsilon}-S^{(q)}_{k\varepsilon})Z^{(j)}_k + \pi^{(j)}_k\big(H(S^{(j)}_{k\varepsilon}-S^{(q)}_{k\varepsilon}) - \Psi(S^{(j)}_{k\varepsilon}-S^{(q)}_{k\varepsilon})\big)\big). \tag{5.9}$$

If we consider the exponential localising function (4.4), then it holds that

$$\mathbb{J}[\cdot](\alpha) = \frac{1}{N}\sum_{j=1}^{N} \cdot^{(j)} H(S^{(j)}_{k\varepsilon}-S^{(q)}_{k\varepsilon})\mathrm{e}^{\lambda^* S^{(q)}_{k\varepsilon}}\mathrm{e}^{-\lambda^* S^{(j)}_{k\varepsilon}}(\lambda^* Z^{(j)}_k + \pi^{(j)}_k). \tag{5.10}$$

Now let us sort the simulated paths such that the values $S^{(q)}_{k\varepsilon}$ increase for q going from 1 to N and let us indicate this by the superscript s, say $S^{s,(q)}_{k\varepsilon}$. Then we write for each q

$$\begin{aligned}&\mathbb{E}\big[P((k+1)\varepsilon,S_{(k+1)\varepsilon})\big|S_{k\varepsilon}=S^{s,(q)}_{k\varepsilon}\big]\\ &\quad= \frac{\mathrm{e}^{\lambda^*_P S^{s,(q)}_{k\varepsilon}}\sum_{j=q}^{N} P((k+1)\varepsilon,S^{s,(j)}_{(k+1)\varepsilon})\mathrm{e}^{-\lambda^*_P S^{s,(j)}_{k\varepsilon}}(\lambda^*_P Z^{s,(j)}_k + \pi^{s,(j)}_k)}{\mathrm{e}^{\lambda^*_1 S^{s,(q)}_{k\varepsilon}}\sum_{j=q}^{N}\mathrm{e}^{-\lambda^*_1 S^{s,(j)}_{k\varepsilon}}(\lambda^*_1 Z^{s,(j)}_k + \pi^{s,(j)}_k)}.\end{aligned}$$

Thus for q going from N to 1, the sums in the numerator and denominator get only one additional term. Hence to estimate $\mathbb{E}\big[P((k+1)\varepsilon,S_{(k+1)\varepsilon})\big|S_{k\varepsilon}=S^{s,(q)}_{k\varepsilon}\big]$ for each q, we can make use of the previously performed computations for $q+1$.

5.3. Numerical results for the Merton model

In this subsection we present the numerical results obtained via our representations in the context of European and American options. We compare our results to those reported by [1]. For this purpose we use the following parameter set for a put option on the underlying stock price process S,

$$\begin{array}{ll} S \text{ modelled by (5.5):} & s_0=40, r=0.08, \beta^2=0.05, \mu=5, \delta^2=0.05,\\ \text{put option:} & T=1, K\in\{30,35,40,45,50\}.\end{array} \tag{5.11}$$

Computations are performed in Matlab by a 2.80 GHz processor with 8GB.

5.3.1. Results for European option prices and deltas

From the point of risk management one might study the value of European options at times $t > 0$. For example one might have the intention to sell the option when the underlying stock price value has increased from the initial value $s_0 = 40$ to 42. Then the owner wants to know at which price the option will be sold. Since it is not known at which time this will happen, one has to price the option at several future times under the assumption that the underlying has changed to 42 at those times.

As an example we estimate the prices

$$P^{\text{Eu}}(s,\alpha) = \mathrm{e}^{-r(T-s)}\mathbb{E}[\Phi(S_T)|S_s = \alpha], \tag{5.12}$$

and deltas

$$\Delta^{\text{Eu}}(s,\alpha) = \mathrm{e}^{-r(T-s)}\frac{\partial}{\partial\alpha}\mathbb{E}[\Phi(S_T)|S_s = \alpha], \tag{5.13}$$

of a European put option with maturity $T = 1$ and strike $K = 45$ on the underlying S described in (5.11), at times $s \in \{0.1, 0.2, \dots, 0.9\}$ and for $\alpha \in \{35, 36, \dots, 45\}$. We do not consider European option prices or deltas at time zero since they do not involve conditional expectations. The estimation of the prices or deltas based on the CDM or MM approach includes the localising technique. Each estimate results from the same set of $N = 5\,000\,000$ simulated paths. In Table 1 we present the CDM and the MM estimates for the option prices for $\alpha \in \{35, 40, 42\}$. We also report the relative errors in percentages to the Merton option prices, see paragraph 5.2.3. Similar results were obtained for the other values of $\alpha \in \{35, 36, \dots, 45\}$. Table 2 shows the corresponding results for the deltas. It turns out that the relative errors when comparing our approach to the approach of [29] are very small. Hence the algorithm we developed is accurate.

	$\alpha = 35$		$\alpha = 40$		$\alpha = 42$	
time	CDM (r.e. %)	MM (r.e. %)	CDM (r.e. %)	MM (r.e. %)	CDM (r.e. %)	MM (r.e. %)
0.1	11.6447 (0.10)	11.6438 (0.09)	9.1820 (-0.02)	9.1816 (-0.03)	8.3520 (0.02)	8.3520 (0.01)
0.2	11.4615 (-0.04)	11.4597 (-0.05)	8.9133 (-0.07)	8.9133 (-0.07)	8.0584 (0.01)	8.0589 (0.02)
0.3	11.2838 (0.02)	11.2820 (0.01)	8.6180 (-0.05)	8.6181 (-0.05)	7.7308 (0.03)	7.7307 (0.03)
0.4	11.0765 (-0.03)	11.0730 (-0.06)	8.2936 (0.08)	8.2946 (0.09)	7.3569 (0.02)	7.3573 (0.03)
0.5	10.8633 (0.02)	10.8615 (0.01)	7.9014 (-0.04)	7.9011 (-0.04)	6.9254 (-0.02)	6.9243 (-0.03)
0.6	10.6319 (0.03)	10.6272 (-0.01)	7.4534 (-0.13)	7.4576 (-0.07)	6.4253 (0.00)	6.4269 (0.03)
0.7	10.3970 (0.05)	10.3894 (-0.02)	6.9561 (0.14)	6.9576 (0.16)	5.8288 (0.08)	5.8305 (0.10)
0.8	10.1765 (0.02)	10.1730 (-0.02)	6.3418 (0.07)	6.3398 (0.03)	5.0875 (0.15)	5.0906 (0.21)
0.9	10.0221 (-0.00)	10.0198 (-0.03)	5.6409 (0.05)	5.6392 (0.02)	4.1176 (-0.04)	4.1161 (-0.07)

Table 1: Estimates of European put option prices $P^{\mathrm{Eu}}(s, \alpha)$ (5.12) via the CDM and MM approach, with relative errors to the Merton prices in percentages. Parameters are given in (5.11), we fix $K = 45$.

	$\alpha = 35$		$\alpha = 40$		$\alpha = 42$	
time	CDM (r.e. %)	MM (r.e. %)	CDM (r.e. %)	MM (r.e. %)	CDM (r.e. %)	MM (r.e. %)
0.1	-0.5473 (0.36)	-0.5469 (0.29)	-0.4372 (0.15)	-0.4368 (0.03)	-0.3959 (-0.36)	-0.3955 (-0.44)
0.2	-0.5703 (0.35)	-0.5703 (0.35)	-0.4525 (0.06)	-0.4534 (0.24)	-0.4095 (-0.08)	-0.4098 (-0.03)
0.3	-0.5936 (-0.21)	-0.5943 (-0.10)	-0.4715 (0.24)	-0.4703 (-0.01)	-0.4266 (0.59)	-0.4260 (0.45)
0.4	-0.6243 (-0.28)	-0.6233 (-0.45)	-0.4927 (0.21)	-0.4933 (0.33)	-0.4423 (0.40)	-0.4417 (0.26)
0.5	-0.6662 (0.38)	-0.6660 (0.35)	-0.5194 (0.29)	-0.5173 (-0.12)	-0.4603 (-0.02)	-0.4588 (-0.34)
0.6	-0.7147 (0.65)	-0.7137 (0.51)	-0.5538 (0.29)	-0.5546 (0.44)	-0.4863 (0.11)	-0.4841 (-0.34)
0.7	-0.7749 (0.90)	-0.7766 (1.12)	-0.6028 (0.36)	-0.6038 (0.53)	-0.5227 (0.28)	-0.5211 (-0.01)
0.8	-0.8403 (0.11)	-0.8355 (-0.46)	-0.6782 (0.24)	-0.6793 (0.40)	-0.5801 (0.26)	-0.5797 (0.19)
0.9	-0.9156 (-0.44)	-0.9171 (-0.27)	-0.8123 (0.11)	-0.8141 (0.33)	-0.7012 (0.57)	-0.7006 (0.48)

Table 2: Estimates of European put option deltas $\Delta^{\mathrm{Eu}}(s, \alpha)$ (5.13) via the CDM and MM approach, with relative errors to the Merton deltas in percentages. Parameters are given in (5.11), we fix $K = 45$.

5.3.2. Results for American option prices and deltas

We consider an American put option on the stock price process S with parameters given in (5.11), the strike is fixed at $K = 45$. [1] and [22] developed a tree method to estimate the American option price. In the current setting their estimate for the option price equals 9.954. The Merton European option price at time zero equals 9.422. As mentioned before we approximate an American option by a Bermudan option using a time discretisation, we choose $n = 10$. The dynamic programming algorithm presented in paragraph 5.1.2 and our representations are used to estimate $P(0, s_0)$ and $\Delta(0, s_0)$.

Figures 1-4 illustrate the influence of the variance reduction techniques on the estimates for the price. The graphs on the right hand side are obtained by zooming in on the left graphs. For $N = 250i$, $i = 1, \ldots, 30$, we simulated N paths of the underlying at the discrete time points jT/n, $j = 1, \ldots, n$, and we estimated the option price at time zero through the CDM and the MM, with and without control variable and with and without the optimal localisation technique. In case the European option is included as a control variable, we put $\gamma = 0.9$.

The variance reduction techniques have a remarkable improvement on the results obtained via the CDM and MM approaches. It appears that the CDM results show some more variation than the MM results.

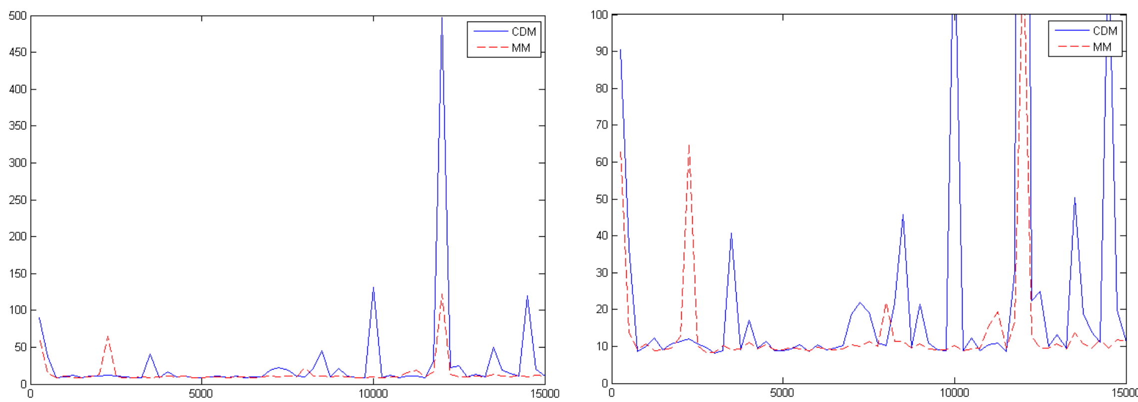

Figure 1: Estimates for the American put option price obtained through the CDM and MM representations *without* control variable and *without* localisation technique, against the number of simulated paths. In the right graph the vertical axis is restricted to $[0, 100]$.

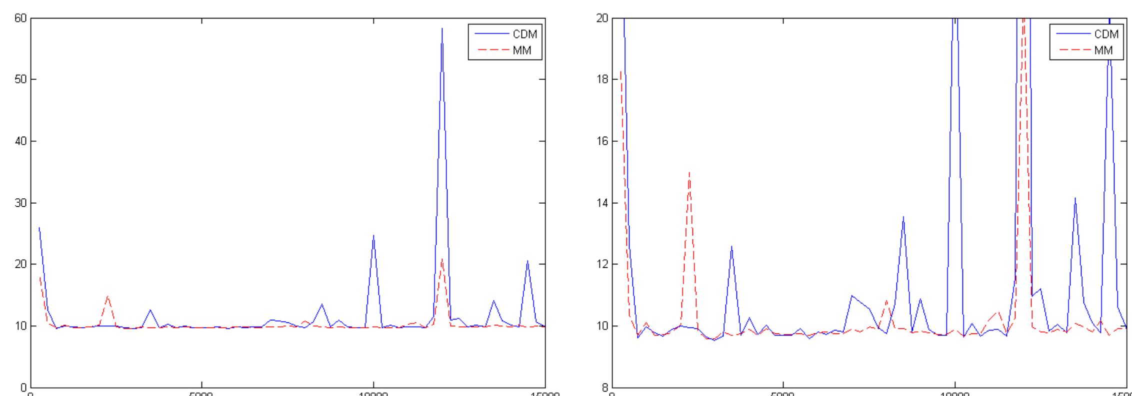

Figure 2: Estimates for the American put option price obtained through the CDM and MM representations *with* control variable and *without* localisation technique, against the number of simulated paths. In the right graph the vertical axis is restricted to $[8, 20]$.

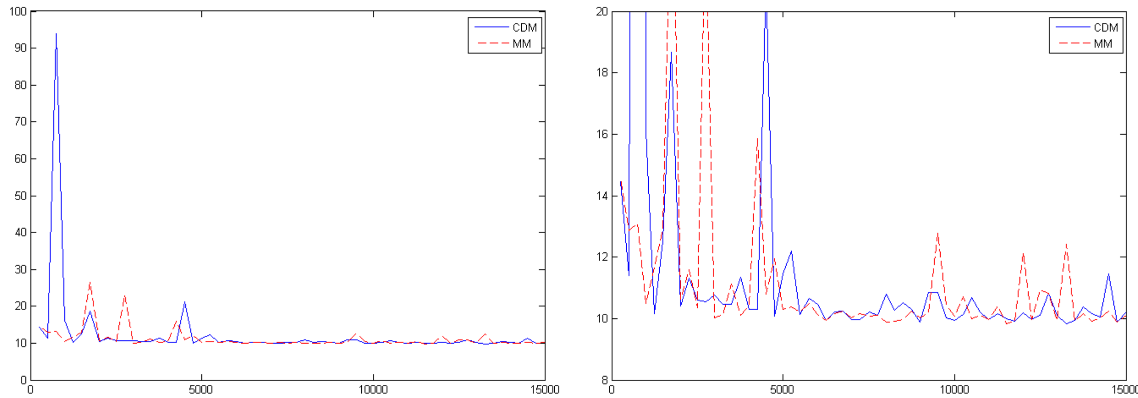

Figure 3: Estimates for the American put option price obtained through the CDM and MM representations *without* control variable and *with* localisation technique, against the number of simulated paths. In the right graph the vertical axis is restricted to $[8, 20]$.

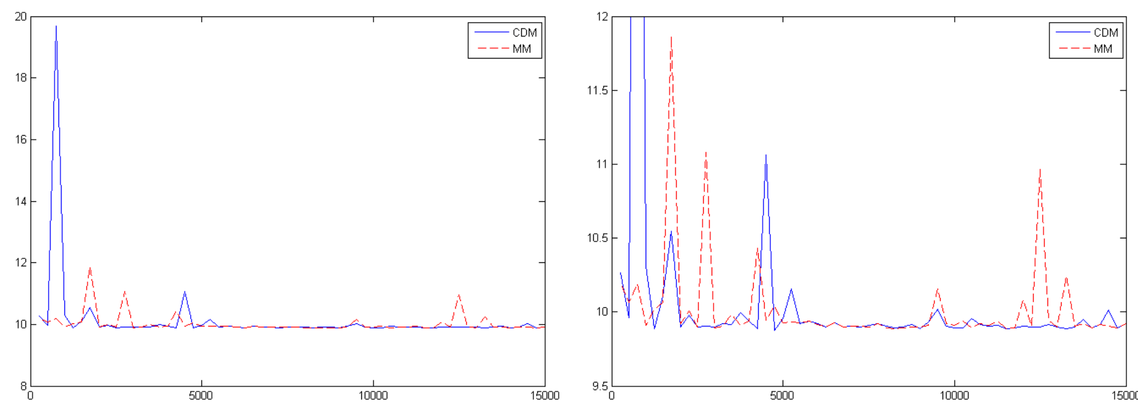

Figure 4: Estimates for the American put option price obtained through the CDM and MM representations *with* control variable and *with* localisation technique, against the number of simulated paths. In the right graph the vertical axis is restricted to $[9.5, 12]$.

Figure 5 presents the required time to estimate the American option price and delta at time zero through the backward dynamic programming algorithm combined with our Malliavin method including a control variable and the optimal localising technique. The paths are simulated backwards in time and the option prices and deltas are estimated simultaneously. The CPU time is given in terms of the number of simulated paths N, for discretisation parameter $n \in \{10, 20\}$. The conditional density method performs more or less at the same rate as the Malliavin method. The CPU time is in the line of [4] and is quadratic in terms of the number of simulated paths. In [1] and [22] there is no clear indication about how long their algorithms take.

As described in paragraph 5.2.5 the computational effort is reduced when we consider an exponential localising function and perform a sorted algorithm. For example for $n = 10$ this method only needs 18 seconds for $N = 10\,000$ and 41 seconds for $N = 20\,000$.

Table 3 presents the estimated prices and deltas of the American put option with strikes $30, 35, 40, 45$, and 50, obtained through the sorted CDM and MM approach including the control variable and the exponential localisation function. For these estimates a time discretisation is performed for $n = 10$ and $500\,000$ paths where simulated. We include the estimates for the prices and deltas obtained by [1] and [22] respectively.

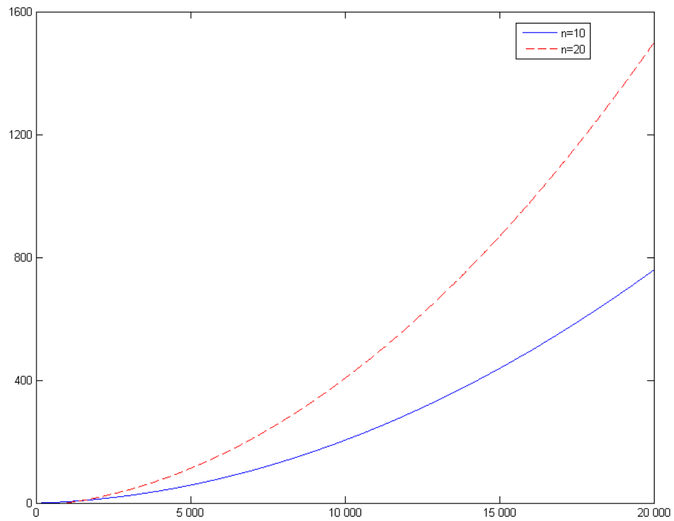

Figure 5: CPU time in seconds to estimate American option price and delta against the number of simulated paths. Malliavin method including a control variable and localisation.

	Price				Delta			
Strike	European	Amin	CDM	MM	European	H-S	CDM	MM
30	2.621	2.720	2.729	2.708	-0.1645	-0.1744	-0.1772	-0.1745
35	4.412	4.603	4.601	4.585	-0.2474	-0.2654	-0.2670	-0.2652
40	6.696	7.030	7.013	7.001	-0.3357	-0.3644	-0.3651	-0.3644
45	9.422	9.954	9.921	9.913	-0.4227	-0.4656	-0.4656	-0.4653
50	12.524	13.318	13.270	13.265	-0.5035	-0.5626	-0.5626	-0.5625

Table 3: Estimates of American put option prices and deltas for parameter set (5.11), obtained through the sorted CDM and MM approach with control variable and exponential localisation. $n = 10$ and $N = 500\,000$. European prices and deltas computed via the Merton approach. American option price estimates from [1] (Amin) and delta estimates from [22] (H-S).

We conclude that the extension that we provided to the geometric Brownian motion model observed in Bally et al. (2005) by adding normally distributed jumps to the driving process leads to numerical results which are in line with those found in the literature.

6. Conclusion

Conditional expectations play an important role in the pricing and hedging of financial derivatives. In the literature there exist several methods for the numerical computations of conditional expectations. One of the methods is to use Monte Carlo by rewriting conditional expectations in terms of expectations without conditioning but involving weights. This was first discussed in [20] in a continuous framework. In this paper we extended this latter approach to include Lévy and jump-diffusion processes. For this purpose we used two approaches: the conditional density method and the Malliavin method. We applied the developed theory to the estimation of American option prices and their deltas. We used a localisation technique and a control variable to improve the estimation of the involved expectations. Moreover, we illustrated our results with different examples and found accurate numerical results.

As far as further investigations are concerned, one may study other choices of the weights in the representation of the conditional expectations. Notice that there are infinitely many possibilities for the weights and thus infinitely many representations of the conditional expectations, see e.g. Proposition 2.6.

Appendix A. Malliavin Calculus

In this paper we make use of the Malliavin calculus as defined in [30]. The following properties and definitions concerning the Malliavin derivative in the direction of the Wiener process are applied.

- The Malliavin derivative in the direction of the Brownian motion is denoted by $D^{(0)}$. The space $\mathbb{D}^{(0)}$ contains all the random variables in $L^2(\Omega)$ that are differentiable in the Wiener direction.

- **Chain rule**
  Let $F \in \mathbb{D}^{(0)}$ and f be a continuously differentiable function with bounded derivative. Then it holds that $f(F) \in \mathbb{D}^{(0)}$ and
$$D^{(0)} f(F) = f'(F) D^{(0)} F.$$

- **Skorohod integral** $\delta^{(0)}$
  Let $\delta^{(0)}$ be the adjoint operator of the directional derivative $D^{(0)}$. The operator $\delta^{(0)}$ maps $L^2(\Omega \times [0,T])$ to $L^2(\Omega)$. The set of processes $u \in L^2(\Omega \times [0,T])$ such that
$$\Big|\mathbb{E}\Big[\int_0^T u_t D_t^{(0)} F dt\Big]\Big| \le C \, \|F\|_{L^2(\Omega)}$$
  for all $F \in \mathbb{D}^{(0)}$, is the domain of $\delta^{(0)}$, denoted by $\operatorname{Dom} \delta^{(0)}$. For every $u \in \operatorname{Dom} \delta^{(0)}$ we define $\delta^{(0)}(u)$ as the Skorohod integral in the Wiener direction of u by
$$\mathbb{E}[F \delta^{(0)}(u)] = \mathbb{E}[\int_0^T u_t D_t^{(0)} F dt],$$
  for any $F \in \mathbb{D}^{(0)}$. The equation above is called the **duality formula**.

- **Integration by parts**
  Let $Fu \in L^2(\Omega \times [0,T])$, where $F \in \mathbb{D}^{(0)}$, and $u \in \operatorname{Dom} \delta^{(0)}$. Then $Fu \in \operatorname{Dom} \delta^{(0)}$ and
$$\delta^{(0)}(Fu) = F\delta^{(0)}(u) - \int_0^T u_t D_t^{(0)} F dt$$
  if and only if the second part of the equation is in $L^2(\Omega)$.

- **Predictable processes**
  Let u be a predictable process such that $\mathbb{E}[\int_0^T u_t^2 dt] < \infty$. Then $u \in \operatorname{Dom} \delta^{(0)}$ and the Skorohod integral coincides with the Ito-integral
$$\delta^{(0)}(u) = \int_0^T u_t dW_t.$$

Acknowledgements The Financial support from the Agency for Innovation by Science and Technology in Flanders (IWT) is gratefully acknowledged by Catherine Daveloose. Asma Khedher thanks the KPMG Center of Excellence in Risk Management for the financial support. Part of the research by Asma Khedher and Michèle Vanmaele was carried out during their stay at the CAS - Centre of Advanced Study of the Norwegian Academy of Science and Letter with the support from the research program SEFE. Michèle Vanmaele also acknowledges the Research Foundation Flanders (FWO) and the Special Research Fund (BOF) of the Ghent University for providing the possibility to go on sabbatical leave to CAS.

References

[1] K.I. Amin. Jump diffusion option valuation in discrete time. *The Journal of Finance*, 48(5): 1833–1863, 1993.

[2] J. Baldeaux. Quasi-Monte Carlo methods for the Kou model. *Monte Carlo Methods and Applications*, 14(4):281–302, 2008.

[3] V. Bally and G. Pagès. Error analysis of the quantization algorithm for obstacle problems. *Stochastic Processes and their Applications*, 106:1–40, 2003.

[4] V. Bally, L. Caramellino, and A. Zanette. Pricing American options by Monte Carlo methods using a Malliavin calculus approach. *Monte Carlo Methods and Applications*, 11:97–133, 2005.

[5] O.E. Barndorff-Nielsen and N. Shephard. Modelling by Lévy processes for financial econometrics. In O.E. Barndorff-Nielsen, T. Mikosch, and S. Resnick, editors, *Lévy processes: Theory and Applications*. Birkhäuser, Boston, 2001.

[6] O.E. Barndorff-Nielsen and N. Shephard. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64:253–280, 2002.

[7] J. Barraquand and D. Martineau. Numerical valuation of high dimensional multivariate American securities. *Journal of Finance and Quantitative Analysis*, 30:383–405, 1995.

[8] D. Bates. Jumps and stochastic volatility: exchange rate processes implicit in Deutsche Mark options. *Review of Financial Studies*, 9(1):69–107, 1996.

[9] R. Bellman. *Dynamic Programming*. Princeton University Press, 1957.

[10] F.E. Benth, J. Kallsen, and T. Meyer-Brandis. A non-Gaussian Ornstein-Uhlenbeck process for electricity spot price modeling and derivatives pricing. *Applied Mathematical Finance*, 14 (2):153–169, 2007.

[11] F.E. Benth, G. Di Nunno, and A. Khedher. Lévy models robustness and sensitivity. *In QP-PQ: Quantum Probability and White Noise Analysis, Proceedings of the 29th Conference in Hammamet, Tunisia, 1318 October 2008. H. Ouerdiane and A Barhoumi (eds.)*, 25:153–184, 2010.

[12] F.E. Benth, G. Di Nunno, and A. Khedher. Robustness of option prices and their deltas in markets modelled by jump-diffusions. *Communications on Stochastic Analysis*, 5(2):285–307, 2011.

[13] B. Bouchard and X. Warin. Monte-Carlo valuation of American options: Facts and new algorithms to improve existing methods. *Numerical Methods in Finance, Springer Proceedings in Mathematics*, 12:215–255, 2012.

[14] F. Bowman. *Introduction to Bessel functions*. Dover, 1958.

[15] D.W. Brewer, J.D. Tubbs, and O.E. Smith. A differential equations approach to the modal location for a family of bivariate gamma distributions. *Journal of Multivariate Analysis*, 21 (1):53–66, 1987.

[16] M. Broadie and P. Glasserman. Pricing American-style securities using simulation. *Journal of Economic Dynamics and Control*, 21(8-9):1323–1352, 1997.

[17] M.H.A. Davis and M.P. Johansson. Malliavin Monte Carlo Greeks for jump-diffusions. *Stochastic Processes and their Applications*, 116(1):101–129, 2006.

[18] G. Di Nunno, T. Meyer-Brandis, B. Øksendal, and F. Proske. Malliavin calculus and anticipative Itô formulae for Lévy processes. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 8(2):235–258, 2005.

[19] E. Fournié, J.M. Lasry, J. Lébucheux, P.L. Lions, and N. Touzi. Applications of Malliavin calculus to Monte Carlo methods in finance. *Finance and Stochastics*, 3:391–412, 1999.

[20] E. Fournié, J.M. Lasry, J. Lébucheux, and P.L. Lions. Applications of Malliavin calculus to Monte Carlo methods in finance II. *Finance and Stochastics*, 5:201–236, 2001.

[21] P. Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer, 2004.

[22] J. Hilliard and A. Schwartz. Pricing European and American derivatives under a jump-diffusion process: a bivariate tree approach. *Journal of Financial and Quantitative Analysis*, 40(3), 2005.

[23] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*. Elsevier Science, 1981.

[24] I. Karatzas and S. Shreve. *Brownian motion and stochastic calculus*. Springer-Verlag, New York, 1991.

[25] A. Kohatsu-Higa and R. Pettersson. Variance reduction methods for simulation of densities on Wiener space. *SIAM Journal on Numerical Analysis*, 40:431–450, 2002.

[26] S. Kou. A jump-diffusion model for option pricing. *Management Science*, 48:1086–1101, 2002.

[27] F.A. Longstaff and E.S. Schwarz. Valuing American options by simulations: a simple least squares approach. *The Review of Financial Studies*, 14:113–148, 2001.

[28] M. Martynov and O. Rozanova. On dependence of volatility on return for stochastic volatility models. *Stochastics: An International Journal of Probability and Stochastic Processes*, 85(5): 917–927, 2013.

[29] R.C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144, 1976.

[30] E. Petrou. Malliavin calculus in Lévy spaces and applications to finance. *Electronic Journal of Probability*, 13:852–879, 2008.

[31] P. Protter. *Stochastic integration and differential equations, Second Edition, Version* 2.1. Number 21 in Stochastic Modelling and Applied Probability. Springer, Berlin, 2005.

[32] C.K. Raju. Products and compositions with the Dirac delta function. *Journal of Physics A: Mathematical and General*, 15(2):381–396, 1982.

[33] K.-I. Sato. *Lévy processes and infinitely divisible distributions*. Number 68 in Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1999.

[34] J.L. Solé, F. Utzet, and J. Vives. Canonical Lévy process and Malliavin calculus. *Stochastic Processes and their Applications*, 117:165–187, 2007.

[35] A.W. van der Stoep, L.A. Grzelak, and C.W. Oosterlee. The Heston stochastic-local volatility model: efficient Monte Carlo simulation. *International Journal of Theoretical and Applied Finance*, 17(7), 2014.