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

Catherine Daveloose, Asma Khedher & Michèle Vanmaele

Published online: 19 Mar 2019.

Submit your article to this journal ⬀

Article views: 909

View related articles ⬀

View Crossmark data ⬀

Taylor & Francis
Taylor & Francis Group

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

Catherine Daveloose[a], Asma Khedher[b], and Michèle Vanmaele[a]

[a]Department of Applied Mathematics, Computer Science and Statistics, Ghent University, Gent, Belgium; [b]Korteweg-de Vries Institute for Mathematics, University of Amsterdam, Amsterdam, The Netherlands

ABSTRACT
In this article, 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.

1. Introduction

Popular models for the fluctuation of the underlying process in finance and energy markets are determined by Brownian motions and are therefore continuous in the sense that they do not generate jumps in the sample paths. Their popularity follows from the fact that they are easy to handle. However, as it turns out from observed data, the absence of jumps is not realistic. Therefore, the use and study of discontinuous models have seen an important boost in the last decades.

In this article, we consider the problem of computing conditional expectations of functionals of Lévy processes and jump diffusions which are involved in pricing and hedging procedures as, e.g. of American options, and which usually require numerical evaluation techniques. In general, it is not possible to obtain analytical expressions for conditional expectations and thus numerical methods are called for. Fournié et al. [1] derive expressions for the conditional expectations in terms of regular expectations for models which have a density and for diffusion models using, respectively, the density

CONTACT Asma Khedher ✉ a.khedher@uva.nl ⬚ Korteweg-de Vries Institute for Mathematics, University of Amsterdam, P.O. Box 94248, 1090 GE Amsterdam, The Netherlands.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/lsaa.

method and the Malliavin calculus. In this article, we generalize these two approaches to obtain representations for conditional expectations and their derivatives (with respect to the underlying) in a jump-diffusion setting. The representations we derive are expressed in terms of regular expectations without conditioning but involving a Heaviside step function and some weights. We apply the developed theory to the numerical estimation of American option prices and their deltas.

There is a wide literature on the use of Malliavin calculus in finance, see, e.g. Montero and Kohatsu-Higa [2]. For an overview, we refer to Nualart [3] for continuous processes and Di Nunnio et al. [4] for jump processes. As for the numerical computation of American options, several approaches appeared in this field. von Sydow et al. [5] compared some approaches grouped as Monte Carlo methods, Fourier methods, finite difference methods, and radial basis function methods but only for continuous underlying processes when pricing American options. Broadie and Glasserman [6] built up a tree in order to obtain a discretisation of the underlying diffusion on a grid (see also Barranquand and Martineau [7]). Longstaff and Schwartz [8] 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. Shu et al. [9] and Kovalov et al. [10] use a partial differential equations approach. Zhang and Oosterlee [11] use a Fourier method based on Fourier cosine-series expansions. We follow an approach similar to the one by Mrad et al. [12] and Bally and Pagès [13] based on a *Monte Carlo* method.

Considering a random variable F, a scalar random variable G, and a function f on $\mathbb{R}$, Fournié et al. [1] provide the following representation for the conditional expectation

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

where π is a random variable called weight and H is the Heaviside step function increased with some constant, $H(x) = 1_{\{x\geq 0\}} + c, c \in \mathbb{R}$. The authors use two approaches: the density method and the Malliavin method (MM). 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 second 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 modeled 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 [1] 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. To overcome this shortcoming, we use the *conditional density method* introduced by Benth et al. [14]. 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. The density of the other part is not necessarily known.

For the *Malliavin method*, we work with the Malliavin derivative for jump processes developed by Petrou [15]. 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 in [15] does not require any, as opposed to the Malliavin calculus used in Davis and Johansson [16] or in Benth et al. [17]. We use this approach to compute conditional expectations in our setting as was developed in [1] in a continuous set up.

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 this article, 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.

To perform the numerical experiments, American options are approximated, through a time discretisation, by Bermudan options as in Bally and Pagès [13] or Abbas-Turki and Lapeyre [18]. We make use of a localization technique and a control variate to minimize the variance, see, e.g. Glasserman [19]. 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 Baldeaux [20].

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 Merton [21]. 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. In addition we compute for the prices a confidence interval according to Bouchard and Warin [22].

In their article [22], Bouchard and Warin performed a numerical comparison and discussed the efficiency and the level of complexity of the regression based Longstaff and Schwartz [8] algorithm and the Malliavin approach algorithm in the continuous setting. In our numerical examples, we investigate how the CDM and the MM perform in the discontinuous setting. The fundamental difference between the Longstaff–Schwartz approach and the (conditional) density or Malliavin approach is the way the conditional expectations are approximated. However, the Longstaff–Schwartz method may show rank deficiency for a certain choice of the parameters of the model (see Mostovyi [23]) and 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 article allow the computation of the price of the American option for any parameters choice and lead to representation formulas for the derivative of conditional expectations and consequently provide an estimation of the delta using its own specific method.

The article is organized as follows. In Sections 2 and 3, we develop a representation for conditional expectations via the conditional density method and the MM, 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 article.

2. Conditional expectation via the conditional density method

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space equipped with a filtration $\mathbb{F} := \{\mathcal{F}_t\}_{t \in [0,T]}$ for time horizon $T > 0$, satisfying the *usual conditions* (see Protter [24]). We introduce the generic notation $L = \{L_t\}_{0 \le t \le T}$, 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 $\mathcal{L}(dz)$. Recall that $\mathcal{L}(dz)$ is a σ-finite Borel measure on $\mathbb{R}_0 := \mathbb{R} \setminus \{0\}$.

In this article, we express the realization 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 $\mathbb{F}$-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.

2.1. Representation results

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*

$$F = g_1(X, Y) \quad \text{and} \quad G = g_2(U, V). \tag{2.1}$$

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. *$\log p_{(X,U)}(x, \cdot)$ is differentiable, for all $x \in \mathbb{R}$, and*
3. *$\frac{\partial}{\partial u} \log p_{(X,U)} \in L^2(\mathbb{R}^2, p_{(X,U)})$, i.e. $\mathbb{E}[(\frac{\partial}{\partial u} \log p_{(X,U)}(X, U))^2] < \infty$.*

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

$$g_2(u, v) = h^{-1}(u + g^*(v)), \tag{2.2}$$

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

We apply the *conditional density method* (CDM) as it is developed in Benth et al. [14]. The CDM method is based on the density method (DM) introduced in Glasserman [19] and used in Fournié et al. [1], but allows for more general dynamics than the latter. Indeed, given F and G as in (2.1), the CDM does not require the knowledge of the density of the couple (F, G) as it is the case in the DM but only the knowledge of the density (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

$$H(x) := 1_{\{x \geq 0\}} + c, \tag{2.3}$$

and its distributional derivative, the *Dirac delta function* by δ_0.

In the upcoming theorem, we present the first representation result of this paper. In order to make the proof comprehensible, we follow a formal approach. The formal proof follows similar derivations as in [1, Section 4.1] by Fournié et al. The main difference is that we apply the CDM instead of the DM by conditioning on $\sigma(Y, V)$. A rigorous proof is included in Appendix A.

Theorem 2.2. *Let F and G be as described in Assumptions 2.1 and let f be a Borel measurable function such that $f(F) \in L^2(\Omega)$. Then it holds for any $\alpha \in$ Dom h that*

$$\mathbb{E}[f(F)|G = \alpha] = \frac{\mathbb{E}[f(F)H(G - \alpha)\pi_{(X,U)}]}{\mathbb{E}[H(G - \alpha)\pi_{(X,U)}]}, \tag{2.4}$$

where

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

Proof. Formally, for 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

$$\mathbb{E}[f(F)\delta_0(G - \alpha)] = \mathbb{E}[f(g_1(X, Y))\delta_0(g_2(U, V) - \alpha)]$$
$$= \mathbb{E}[\mathbb{E}[f(g_1(X, Y))\delta_0(g_2(U, V) - \alpha)|\sigma(Y, V)]],$$

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

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

For a function $\phi \in C^1$ with a single root, the composition rule for the Dirac delta function (see Raju [25]) states that

$$\delta_0(\phi(u)) = \frac{\delta_0(u - u_1)}{|\phi'(u_1)|},$$

where u_1 is such that $\phi(u_1) = 0$ and $\phi'(u_1) \neq 0$. Because of relation (2.2), we apply this composition rule to the function $\phi(u) = h^{-1}(u + g^*(V)) - \alpha$, with root $u_1 = h(\alpha) - g^*(V)$, and obtain that

$$\int_{\mathbb{R}} \delta_0(g_2(u, V) - \alpha)p_{(X,U)}(x, u)du = \int_{\mathbb{R}} \frac{\delta_0(u + g^*(V) - h(\alpha))}{(h^{-1})'(h(\alpha))} p_{(X,U)}(x, u)du$$
$$= h'(\alpha)\int_{\mathbb{R}} \delta_0(u + g^*(V) - h(\alpha))p_{(X,U)}(x, u)du.$$

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

$$\int_{\mathbb{R}} \delta_0(g_2(u, V) - \alpha) p_{(X,U)}(x, u) du$$

$$= -h'(\alpha) \int_{\mathbb{R}} H\big(u + g^*(V) - h(\alpha)\big) \frac{\partial}{\partial u} p_{(X,U)}(x, u) du$$

$$= -h'(\alpha) \int_{\mathbb{R}} H(g_2(u, V) - \alpha) \left(\frac{\partial}{\partial u} \log p_{(X,U)}(x, u) \right) p_{(X,U)}(x, u) du.$$

Finally we conclude that

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

$$= \mathbb{E}\left[f(F) H(G - \alpha) \left\{ -\frac{\partial}{\partial u} \log p_{(X,U)}(X, U) \right\} \right] h'(\alpha).$$

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

Note that the weights $\pi_{(X,U)}$ in the representation (2.4) are not unique and one can derive infinitely many of them. This was discussed in the result (4.25)–(4.26) in [1] for random variables admitting a density. In Proposition 2.1 of this latter article, the authors discussed also the optimality of the weights in a minimal variance setting. The extension of these results to our jump-diffusion setting using the CDM is straightforward and follows similar derivations as in the proof of Theorem 2.2. For more details we refer to Daveloose [26].

In the next theorem, we deduce a representation for the sensitivity of (2.4) with respect to α.

Theorem 2.3. *Let F and G be as described in Assumptions 2.1, where* $\log p_{(X,U)}$ *possesses a second order derivative in its second argument and let the Borel measurable function f guarantee* $f(F) \in L^2(\Omega)$. *Then it holds for any* $\alpha \in Dom\ h$ *that*

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

where

$$\mathbb{A}_{F,G}[\cdot](\alpha) = \mathbb{E}\big[\cdot(F) H(G - \alpha) \pi_{(X,U)}\big],$$

$$\mathbb{B}_{F,G}[\cdot](\alpha) = \mathbb{E}\left[\cdot(F) H(G - \alpha)\left(-\pi_{(X,U)}^2 + \pi_{(X,U)}^*\right)\right],$$

$$\pi_{(X,U)} = -\frac{\partial}{\partial u} \log p_{(X,U)}(X, U), \text{ and}$$

$$\pi_{(X,U)}^* = -\frac{\partial^2}{\partial u^2} \log p_{(X,U)}(X, U).$$

Proof. From Theorem 2.2, it follows immediately that

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

For the derivatives in the right hand side, it holds by the square integrability assumption for $f(F)$ that

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

Along the lines of the formal proof of Theorem 2.2, we derive

$$\int_{\mathbb{R}} \delta_0(g_2(u,V)-\alpha)\left(\frac{\partial}{\partial u}\log p_{(X,U)}(x,u)\right)p_{(X,U)}(x,u)du$$

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

$$= -\int_{\mathbb{R}} H\big(u+g^*(V)-h(\alpha)\big)h'(\alpha)\frac{\partial^2}{\partial u^2}p_{(X,U)}(x,u)du$$

$$= -\int_{\mathbb{R}} H(g_2(u,V)-\alpha)h'(\alpha)\left\{\left(\frac{\partial}{\partial u}\log p_{(X,U)}(x,u)\right)^2 + \frac{\partial^2}{\partial u^2}\log p_{(X,U)}(x,u)\right\}p_{(X,U)}(x,u)du,$$

which concludes the proof.

$\square$

2.2. Examples: Jump-diffusion models

In many applications in mathematical finance, one can make grateful use of Theorem 2.2. In fact, we are able to express a realization 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. We will use Monte Carlo simulations to evaluate these conditional expectations using their representations as regular expectations. In the sequel, we consider different jump-diffusion models for the stock price process to illustrate our results with specific examples.

Remark 2.4. **(Processes with independent and identically distributed increments).** In case it holds that the stock price process S is modeled in terms of a process with increments which are independent and identically distributed we remark the following about a conditional expectation of the form $\mathbb{E}[f(S_t)|S_s = \alpha]$. Consider for instance $S = S_0 e^L$ where L denotes a Lévy process. Then, given that $S_s = \alpha$, it holds, due to the independence and the identical distribution of the increments, that for any $0 < s < t < T$

$$S_t = S_s \frac{S_t}{S_s} = \alpha e^{L_t - L_s} \overset{d}{=} \alpha e^{L_{t-s}}.$$

Therefore

$$\mathbb{E}\big[f(S_t)|S_s = \alpha\big] = \mathbb{E}\big[f(\alpha e^{L_{t-s}})\big]$$

and the conditional expectation equals a regular expectations which can easily be estimated via a Monte Carlo method. Although it is possible to use Monte Carlo simulations exploiting the independence and stationarity property of the increments, the application of the obtained representations avoids nested simulations in the pricing of American options, as explained further on in Remark 5.1. Hence, this allows to obtain estimations based on a smaller number of simulated paths.

2.2.1. Jump-diffusion model

The next proposition considers a representation of conditional expectations in terms of unconditional expectations, when 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 (American) option prices (which we will illustrate in Section 5 with numerical experiments).

Proposition 2.5. *Consider a price process S defined by $S_t = e^{L_t}, \forall t \in [0, T]$, where L is a Lévy jump diffusion process with decomposition $L_t = \mu t + \beta W_t + \tilde{N}_t$. Here W is a standard Brownian motion, $\tilde{N}$ is a compound Poisson process independent of W, and μ and β are constant parameters. Then, for any Borel measurable function f fulfilling $f(S_t) \in L^2(\Omega)$, any strictly positive number α, and $0 < s < t < T$, it holds that*

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

where

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

Proof. Following the notation of Theorem 2.2, we set $F = S_t$ and $G = S_s$. The random variables $X = \beta W_t, Y = \mu t + \tilde{N}_t, U = \beta W_s$ and $V = \mu s + \tilde{N}_s$ and the functions $g_i(x, y) = e^{x+y}, i \in \{1, 2\}, g^*(v) = v$, and $h(\alpha) = \log \alpha$, with $\alpha \in \text{Dom } h = \mathbb{R}_0^+$, satisfy Assumptions 2.1. The random variables X and U (resulting from a scaled Brownian motion) have a joint normal distribution with density function

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

To determine the weight in (2.4) 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

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

$\square$

Example: Merton model: In the light of the numerical experiments that we perform in Section 5, we derive the representation (2.4) for the following Merton model given by

$$S_t = s_0 \exp\left(\left(r - \frac{\beta^2}{2}\right)t + \beta W_t + \sum_{i=1}^{N_t} Y_i\right), \tag{2.6}$$

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 the Merton model is a special case of the exponential Lévy jump-diffusion model, the representations obtained via the conditional density method, i.e. through Proposition 2.5 and Theorem 2.3, are as follows, for $0 < s < t < T$ and $\alpha \in \mathbb{R}_0^+$,

$$\mathbb{E}\big[f(S_t)|S_s = \alpha\big] = \frac{\mathbb{E}\big[f(S_t)H(S_s - \alpha)\pi_{s,t}\big]}{\mathbb{E}\big[H(S_s - \alpha)\pi_{s,t}\big]} =: \frac{\mathbb{A}_{t,s}[f](\alpha)}{\mathbb{A}_{t,s}[1](\alpha)}, \tag{2.7}$$

$$\frac{\partial}{\partial \alpha}\mathbb{E}\big[f(S_t)|S_s = \alpha\big] = \frac{\mathbb{B}_{t,s}[f](\alpha)\mathbb{A}_{t,s}[1](\alpha) - \mathbb{A}_{t,s}[f](\alpha)\mathbb{B}_{t,s}[1](\alpha)}{\mathbb{A}_{t,s}[1](\alpha)^2}\frac{1}{\alpha}, \tag{2.8}$$

where

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

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

2.2.2. Additive model

Now we observe a model which is often used to price energy products (see for example Benth et al. [27]). The price process is given by an *additive model*

$$S_t = X_t + Y_t, \forall t \in [0, T] \quad \text{with } S_0 > 0. \tag{2.9}$$

The process Y is adapted to the filtration $\mathbb{F}$ 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 Sato [28]. Namely, it is a process following the dynamics

$$dX_t = -\lambda X_t dt + dL_t, \quad X_0 = S_0, \tag{2.10}$$

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 (2.10) equals

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

An interesting property of this type of non-Gaussian OU-processes is the fact that the autocorrelation is independent of the stationary distribution, see, e.g. Barndorff–Nielsen and Shephard [29] or Sato [28], it equals

$$\text{Corr}(X_t, X_s) = e^{\lambda(s-t)}, \quad \forall 0 < s < t. \tag{2.11}$$

Proposition 2.6. *Let us observe the additive model described by (2.9) and (2.10). Then it holds for any Borel measurable function f satisfying $f(S_t) \in L^2(\Omega), 0 < s < t < T$, and $\alpha \in \mathbb{R}$, that*

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

where

$$\pi = \frac{1-a}{X_s} + \frac{b}{1-\rho} - \frac{\nu(X_t, X_s)}{2X_s}\frac{\mathcal{I}_a(\nu(X_t, X_s))}{\mathcal{I}_{a-1}(\nu(X_t, X_s))}.$$

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

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

Proof. As in Theorem 2.2, 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 that X_t and X_s are both $\Gamma(a, b)$ distributed and by (2.11) we know that $\text{Corr}(X_t, X_s) = e^{\lambda(s-t)} =: \rho$. According to Brewer et al. [30], the density function of this bivariate gamma distribution with non-zero correlation equals

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

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

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

For the function $v(x, u) = 2\sqrt{\rho b^2 xu}/(1-\rho)$, it holds that $\frac{\partial v}{\partial u}(x, u) = \sqrt{(\rho b^2 x/u)}/(1-\rho)$ and $\frac{\partial v}{\partial u}(x, u) = v(x, u)/(2u)$. Using the recurrence formulas for modified Bessel functions (see Bowman [31]), we get

$$\frac{\partial}{\partial u} \log \left(\mathcal{I}_{a-1}(v(x, u))\right) = \frac{1}{\mathcal{I}_{a-1}(v(x, u))} \mathcal{I}'_{a-1}(v(x, u)) \frac{\partial v}{\partial u}(x, u)$$

$$= \frac{1}{\mathcal{I}_{a-1}(v(x, u))} \frac{1}{2} \left(\mathcal{I}_{a-2}(v(x, u)) + \mathcal{I}_a(v(x, u))\right) \frac{v(x, u)}{2u}$$

$$= \frac{1}{\mathcal{I}_{a-1}(v(x, u))} \frac{1}{2} \left(\{\mathcal{I}_{a-2}(v(x, u)) - \mathcal{I}_a(v(x, u))\} + 2\mathcal{I}_a(v(x, u))\right) \frac{v(x, u)}{2u}$$

$$= \frac{1}{\mathcal{I}_{a-1}(v(x, u))} \frac{1}{2} \left(\frac{2(a-1)}{v(x, u)} \mathcal{I}_{a-1}(v(x, u)) + 2\mathcal{I}_a(v(x, u))\right) \frac{v(x, u)}{2u}$$

$$= \left(\frac{(a-1)}{v(x, u)} + \frac{\mathcal{I}_a(v(x, u))}{\mathcal{I}_{a-1}(v(x, u))}\right) \frac{v(x, u)}{2u} = \frac{a-1}{2u} + \frac{v(x, u)}{2u} \frac{\mathcal{I}_a(v(x, u))}{\mathcal{I}_{a-1}(v(x, u))}.$$

According to (2.4) we conclude the statement. $\square$

3. Conditional expectation via MM

Fournié et al. [1] used Malliavin calculus defined for functions on the Wiener space to obtain representations for the conditional expectations. Therefore, their approach is applied to *continuous diffusions*. In this section, we extend their method to allow for the computation of conditional expectations in a *Lévy* and a *jump-diffusion* framework. For this purpose we use a Malliavin derivative of the combination of Gaussian and pure jump Lévy noises, see, e.g. Di Nunno et al. [4] and Solé et al. [32]. In our setting, we use the Malliavin derivative developed by Petrou [15].

Notice that the MM allows for more general dynamics than the conditional density method (CDM). Indeed, no density is required to be known when using the MM. However, the CDM is easier to apply and follows using basic knowledge in probability theory.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space in which Lévy processes are well-defined. The *Malliavin derivative in the Brownian direction* is defined by Petrou [15] 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 [15] and denoted by $\delta^{(0)}$. In this section we make use of some notations, definitions, computational rules and properties which are summarized in Appendix B. Recall the function H defined in (2.3).

3.1. Representation results

The ideas and concepts of the proofs of the results presented in this section are based on the proofs of Theorem 4.1 and Corollary 4.1 in Fournié et al. [1] and follow similar lines of derivations. The difference in our approach is that we use a Malliavin derivative defined for Lévy processes. This allows us to generalize the results in [1] to processes with jumps.

In the following theorem we derive a first representation result for the conditional expectation $\mathbb{E}[f(F)|G = \alpha]$, where the function f possesses a bounded and continuous derivative.

Theorem 3.1. *Denote by C_b^l the space of continuously differentiable functions with bounded derivative. Moreover, Let $f \in C_b^l$ be a Borel measurable function, F and G be in $\mathbb{D}^{(0)}$, u be in $\mathrm{Dom}\ \delta^{(0)}$ such that $f(F)u$ is in $L^2(\Omega \times [0, T])$, and*

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

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

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

Proof. We note that

$$\mathbb{E}[f(F)|G = \alpha] = \lim_{\varepsilon \to 0+} \mathbb{E}[f(F)|G \in]\alpha - \varepsilon, \alpha + \varepsilon[] = \lim_{\varepsilon \to 0+} \mathbb{E}\left[\frac{f(F)\mathbb{1}-\varepsilon, \varepsilon[(G-\alpha)]}{\mathbb{E}[\mathbb{1}-\varepsilon, \varepsilon[(G - \alpha)].}\right]$$

As a first step we show that for any $\varepsilon > 0$,

$$\frac{1}{2}\mathbb{E}\left[f(F)\mathbb{1}-\varepsilon, \varepsilon[(G-\alpha)]\right] = \mathbb{E}\left[H_\varepsilon(G-\alpha)\delta^{(0)}(f(F)u)\right], \tag{3.3}$$

where

$$H_\varepsilon(x) = \begin{cases} \varepsilon c, & x < -\varepsilon, \\ \dfrac{1}{2}(x + \varepsilon) + \varepsilon c, & -\varepsilon \leq x < \varepsilon, \\ \varepsilon + \varepsilon c, & x \geq \varepsilon. \end{cases}$$

Hereto we approximate the function $\frac{1}{2}1_{]-\varepsilon,\varepsilon[}$ by a sequence of bounded continuous functions $(\Phi_{\varepsilon,n})_{n\geq 1}$. Specifically, $\Phi_{\varepsilon,n}$ has support $]-\varepsilon-\varepsilon/n, \varepsilon+\varepsilon/n[$ and equals $\frac{1}{2}$ on $[-\varepsilon+\varepsilon/n, \varepsilon-\varepsilon/n]$. Moreover the graph of $\Phi_{\varepsilon,n}$ connects the points $(-\varepsilon-\varepsilon/n, 0)$ and $(-\varepsilon+\varepsilon/n, 1/2)$ and the points $(\varepsilon-\varepsilon/n, 1/2)$ and $(\varepsilon+\varepsilon/n, 0)$ via straight lines. Besides, we define $\Psi_{\varepsilon,n}(x) := \int_{-\infty}^{x} \Phi_{\varepsilon,n}(y)dy + \varepsilon c$. Then we have by the duality formula, the chain rule, and relation (3.1) that

$$\mathbb{E}\left[\Psi_{\varepsilon,n}(G-\alpha)\delta^{(0)}\big(f(F)u\big)\right] = \mathbb{E}\left[\int_0^T f(F)u_t D_t^{(0)}\big(\Psi_{\varepsilon,n}(G-\alpha)\big)dt\right]$$

$$= \mathbb{E}\left[f(F)\Phi_{\varepsilon,n}(G-\alpha)\int_0^T u_t D_t^{(0)}Gdt\right]$$

$$= \mathbb{E}\left[\mathbb{E}\left[f(F)\Phi_{\varepsilon,n}(G-\alpha)\int_0^T u_t D_t^{(0)}Gdt\Big|\sigma(F,G)\right]\right]$$

$$= \mathbb{E}\left[f(F)\Phi_{\varepsilon,n}(G-\alpha)\right].$$

Because of the facts that $|\Phi_{\varepsilon,n}| \leq \frac{1}{2}$ and $f(F) \in L^2(\Omega)$, the latter expression converges to $\frac{1}{2}\mathbb{E}[f(F)1_{]-\varepsilon,\varepsilon[}(G-\alpha)]$ by the definition of the sequence $(\Phi_{\varepsilon,n})_n$. Besides $|\Psi_{\varepsilon,n}| \leq \varepsilon + \varepsilon c, \delta^{(0)}(f(F)u) \in L^2(\Omega)$, and the sequence $(\Psi_{\varepsilon,n})_n$ approximates H_ε. Hence this proves (3.3).

Next, by the integration by parts formula we find

$$\frac{1}{2}\mathbb{E}[f(F)1] - \varepsilon, \varepsilon[(G-\alpha)] = \mathbb{E}\left[H_\varepsilon(G-\alpha)\left\{f(F)\delta^{(0)}(u) - \int_0^T u_r D_r^{(0)}f(F)dr\right\}\right]$$

$$= \mathbb{E}\left[H_\varepsilon(G-\alpha)\left\{f(F)\delta^{(0)}(u) - f'(F)\int_0^T u_r D_r^{(0)}Fdr\right\}\right].$$

Then applying the latter result for $f \equiv 1$ too, shows that

$$\mathbb{E}\big[f(F)|G=\alpha\big] = \lim_{\varepsilon\to 0+}\frac{\mathbb{E}\left[\frac{1}{\varepsilon}H_\varepsilon(G-\alpha)\left\{f(F)\delta^{(0)}(u) - f'(F)\int_0^T u_r D_r^{(0)}Fdr\right\}\right]}{\mathbb{E}\left[\frac{1}{\varepsilon}H_\varepsilon(G-\alpha)\delta^{(0)}(u)\right]}.$$

Since it holds that $|\frac{1}{\varepsilon}H_\varepsilon(G-\alpha)| \leq 1 + c, \mathbb{E}[|\delta^{(0)}(f(F)u)|] < \infty, \delta^{(0)}(u) \in L^2(\Omega)$, and $\frac{1}{\varepsilon}H_\varepsilon(x) \to H(x)$ for ε tending to zero, this concludes the proof.   □

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. In order to deal with the potential non-smoothness of f, we include an additional assumption on the process u introduced in Theorem 3.1 leading to the following theorem.

Theorem 3.2. *Let F and G be in $\mathbb{D}^{(0)}$ and f be a Borel measurable function such that $f(F) \in L^2(\Omega)$. Consider a process u in $\mathrm{Dom}\,\delta^{(0)}$, guaranteeing $f(F)u$ is in $L^2(\Omega \times [0,T])$, satisfying (3.1) and, in addition,*

$$\mathbb{E}\left[\int_0^T u_t D_t^{(0)} F dt \big| \sigma(F, G)\right] = 0. \tag{3.4}$$

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

$$\mathbb{E}\big[f(F)|G = \alpha\big] = \frac{\mathbb{E}\Big[f(F)H(G-\alpha)\delta^{(0)}(u)\Big]}{\mathbb{E}\Big[H(G-\alpha)\delta^{(0)}(u)\Big]}. \tag{3.5}$$

Proof. i) First, let us consider a Borel measurable function $f \in C_b^1$, such that we can apply Theorem 3.1. Because of the properties of conditional expectations and relation (3.4), we have in representation (3.2) that

$$\mathbb{E}\left[f'(F)H(G-\alpha)\int_0^T u_t D_t^{(0)} F dt\right] = \mathbb{E}\left[\mathbb{E}\left[f'(F)H(G-\alpha)\int_0^T u_t D_t^{(0)} F dt | \sigma(F, G)\right]\right]$$

$$= \mathbb{E}\left[f'(F)H(G-\alpha)\mathbb{E}\left[\int_0^T u_t D_t^{(0)} F dt | \sigma(F, G)\right]\right] = 0.$$

Thus we obtain representation (3.5).

ii) Now we observe a Borel measurable function f for which $f(F) \in L^2(\Omega)$. Let $C_K^\infty(\mathbb{R})$ and $C_K^\infty(\Omega)$ be the space of infinitely differentiable functions with compact support respectively defined in $\mathbb{R}$ and Ω. Since $C_K^\infty(\Omega)$ is dense in $L^2(\Omega)$ there exists a sequence of functions f_n in $C_K^\infty(\mathbb{R})$, such that $f_n(F)$ converges to $f(F)$ in $L^2(\Omega)$. In part i) we concluded that for any function f_n in this sequence representation (3.5) holds. By convergence arguments, we conclude that expression (3.5) also holds for the limit function f as follows. For any f_n we denote

$$g_n(\alpha) := \mathbb{E}\big[f_n(F)|G = \alpha\big] = \frac{\mathbb{E}\Big[f_n(F)H(G-\alpha)\delta^{(0)}(u)\Big]}{\mathbb{E}\Big[H(G-\alpha)\delta^{(0)}(u)\Big]}.$$

Besides we define

$$g(\alpha) := \frac{\mathbb{E}\Big[f(F)H(G-\alpha)\delta^{(0)}(u)\Big]}{\mathbb{E}\Big[H(G-\alpha)\delta^{(0)}(u)\Big]}.$$

Via the Cauchy–Schwarz inequality, we derive that

$$|g(\alpha)-g_n(\alpha)| \leq \frac{\mathbb{E}\left[\Big|f(F)-f_n(F)\Big|\Big|H(G-\alpha)\delta^{(0)}(u)\Big|\right]}{\Big|\mathbb{E}\Big[H(G-\alpha)\delta^{(0)}(u)\Big]\Big|}$$

$$\leq \frac{\mathbb{E}\big[|f(F)-f_n(F)|^2\big]^{1/2}\mathbb{E}\Big[\big|H(G-\alpha)\delta^{(0)}(u)\big|^2\Big]^{1/2}}{\Big|\mathbb{E}\Big[H(G-\alpha)\delta^{(0)}(u)\Big]\Big|}.$$

For any $\alpha \in \mathbb{R}$ we have

$$\frac{\mathbb{E}\left[|H(G-\alpha)\delta^{(0)}(u)|^2\right]^{1/2}}{|\mathbb{E}\left[H(G-\alpha)\delta^{(0)}(u)\right]|} < \infty.$$

By the density argument $f_n(F)$ converges to $f(F)$ in L^2-sense, hence we obtain that

$$|g(\alpha)-g_n(\alpha)| \to 0, \text{ for } n \to \infty, \forall \alpha \in \mathbb{R}.$$

Thus, $g_n(\alpha)$ convergences to $g(\alpha)$. Moreover it holds that $g_n(\alpha)$ converges to $\mathbb{E}[f(F)|G = \alpha]$ by the conditional dominated convergence theorem. Therefore, we conclude that $g(\alpha)$ equals the latter conditional expectation. □

Via the MM we also deduce a representation for the *delta* in terms of unconditional expectations.

Theorem 3.3. *Consider the same setting as in Theorem 3.2 and assume that*

$$\mathbb{E}\left[\delta^{(0)}(u)\int_0^T u_t D_t^{(0)}F dt | \sigma(F, G)\right] = 0. \tag{3.6}$$

Then the delta is given by

$$\frac{\partial}{\partial\alpha}\mathbb{E}[f(F)|G = \alpha] = \frac{\mathbb{B}_{F,G}[f](\alpha)\mathbb{A}_{F,G}[1](\alpha)-\mathbb{A}_{F,G}[f](\alpha)\mathbb{B}_{F,G}[1](\alpha)}{\mathbb{A}_{F,G}[1](\alpha)^2}, \tag{3.7}$$

where

$$\mathbb{A}_{F,G}[\cdot](\alpha) = \mathbb{E}\left[\cdot(F)H(G-\alpha)\delta^{(0)}(u)\right],$$

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

Proof. The structure of formula (3.7) follows clearly from the derivation of representation (3.5). Now we focus on the derivative

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

$$= -\lim_{\varepsilon\to0+}\frac{1}{2\varepsilon}\mathbb{E}\left[f(F)1] - \varepsilon, \varepsilon\left[(G-\alpha)\delta^{(0)}(u)\right].$$

i) First, we consider a Borel measurable function $f \in C_b^1$. It can be shown, as before in Theorem 3.1, that

$$\frac{1}{2}\mathbb{E}\left[f(F)1] - \varepsilon, \varepsilon\left[(G-\alpha)\delta^{(0)}(u)\right] = \mathbb{E}\left[H_\varepsilon(G-\alpha)\delta^{(0)}\left(f(F)\delta^{(0)}(u)u\right)\right],$$

and therefore

$$\mathbb{B}_{F,G}[f](\alpha) = -\lim_{\varepsilon\to0+}\frac{1}{\varepsilon}\mathbb{E}\left[H_\varepsilon(G-\alpha)\delta^{(0)}\left(f(F)\delta^{(0)}(u)u\right)\right]$$

$$= -\mathbb{E}\left[H(G-\alpha)\delta^{(0)}\left(f(F)\delta^{(0)}(u)u\right)\right].$$

By the chain rule and the integration by parts formula, we obtain

$$
\mathbb{E}\left[H(G-\alpha)\delta^{(0)}\left(f(F)\delta^{(0)}(u)u\right)\right]
$$

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

$$
= \mathbb{E}\left[H(G-\alpha)\left\{f(F)\left\{\delta^{(0)}(u)\delta^{(0)}(u)-\int_0^T u_r D_r^{(0)}\delta^{(0)}(u)dr\right\}-\delta^{(0)}(u)f'(F)\int_0^T u_r D_r^{(0)}Fdr\right\}\right]
$$

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

$$
-\mathbb{E}\left[f'(F)H(G-\alpha)\delta^{(0)}(u)\int_0^T u_r D_r^{(0)}Fdr\right].
$$

By expression (3.6), the latter expectation equals

$$
\mathbb{E}\left[f'(F)H(G-\alpha)\delta^{(0)}(u)\int_0^T u_r D_r^{(0)}Fdr\right]
$$

$$
= \mathbb{E}\left[\mathbb{E}\left[f'(F)H(G-\alpha)\delta^{(0)}(u)\int_0^T u_r D_r^{(0)}Fdr|\sigma(F,G)\right]\right]
$$

$$
= \mathbb{E}\left[f'(F)H(G-\alpha)\mathbb{E}\left[\delta^{(0)}(u)\int_0^T u_r D_r^{(0)}Fdr|\sigma(F,G)\right]\right] = 0.
$$

Hence, we conclude that

$$
\mathbb{E}\left[f(F)\delta_0(G-\alpha)\delta^{(0)}(u)\right] = \mathbb{E}\left[f(F)H(G-\alpha)\left\{\delta^{(0)}(u)^2-\int_0^T u_r D_r^{(0)}\delta^{(0)}(u)dr\right\}\right]. \tag{3.8}
$$

ii) Via a density argument which follows the same lines as the proof of Theorem 3.2, we conclude that equation (3.8) also holds for a Borel measurable function f such that $f(F) \in L^2(\Omega)$.

$\square$

Remark 3.4. Remark that condition (3.6) is fulfilled by a combination of condition (3.4) and $\delta^{(0)}(u)$ being $\sigma(F,G)$-measurable, since in this case

$$
\mathbb{E}\left[\delta^{(0)}(u)\int_0^T u_t D_t^{(0)}Fdt|\sigma(F,G)\right] = \delta^{(0)}(u)\mathbb{E}\left[\int_0^T u_t D_t^{(0)}Fdt|\sigma(F,G)\right] = 0.
$$

Next, some practical examples motivated from financial applications are given.

3.2. Examples: Jump-diffusion models

We start by considering a stock price process which is modeled by a general stochastic differential equation (SDE). For this model we derive representations based on the

results developed in Subsection 3.1. Further on, we illustrate these by looking into some specific types of SDEs such as exponential Lévy and stochastic volatility models. Let S satisfy the following *stochastic differential equation*

$$\begin{cases} dS_t = \mu(t, S_{t-})dt + \beta(t, S_{t-})dW_t + \int_{\mathbb{R}_0}\gamma(t, S_{t-}, z)\tilde{N}(dt, dz), \\ S_0 = s_0 > 0, \end{cases} \tag{3.9}$$

where W is a Wiener process and $\tilde{N}$ is a compensated Poisson random measure with Lévy measure $\mathcal{L}$. We assume that $\beta(t, x) > 0$ for all $(t, x) \in [0, T] \times \mathbb{R}$. The coefficient functions $\mu(t, x), \beta(t, x), \gamma(t, x, z) \in C_b^1$ are 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

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

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 Ikeda and Watanabe [33].

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

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

We assume that the coefficients are such that V is strictly positive. 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 Petrou [15])

$$D_s^{(0)}S_t = V_t(V_{s-})^{-1}\beta(s, S_{s-})1_{\{s \leq t\}}. \tag{3.10}$$

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 [1, Lemma 4.1] to our setting.

Lemma 3.5. *It holds that*

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

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

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

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 3.1 and later on we apply Theorem 3.2.

Proposition 3.6. *Let $f \in C_b^1$, $0 < s < t < T$, and $\alpha \in \mathbb{R}$. In the setting described by the stochastic differential equation (3.9) we assume that*

$$\mathbb{E}\left[\int_0^T \left(\frac{V_{r-}}{\beta(r, S_{r-})}\right)^2 dr\right] < \infty \quad and \quad \mathbb{E}\left[\int_0^T \left(\frac{1}{sV_s}\frac{V_{r-}}{\beta(r, S_{r-})}\right)^2 dr\right] < \infty. \tag{3.11}$$

Then the following representation holds for the conditional expectation

$$\mathbb{E}\big[f(S_t)|S_s = \alpha\big] = \frac{\mathbb{E}\big[f(S_t)H(S_s - \alpha)\pi_1 - f'(S_t)H(S_s - \alpha)\pi_2\big]}{\mathbb{E}[H(S_s - \alpha)\pi_1]}, \tag{3.12}$$

where the Malliavin weights equal

$$\pi_1 = \frac{1}{sV_s}\left(\int_0^s \frac{V_{r-}}{\beta(r, S_{r-})}dW_r + s\frac{\zeta_s}{V_s} + \int_0^s \left[\frac{\beta_x(r, S_{r-})}{\beta(r, S_{r-})}V_{r-} - \frac{\zeta_{r-}}{V_{r-}}\right]dr\right) \quad and \quad \pi_2 = \frac{V_t}{V_s}. \tag{3.13}$$

Proof. We apply Theorem 3.1 and to fulfill condition (3.1) we define

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

Note that the process $V_-/\beta(\cdot, S_-)$ is predictable. By the first condition in (3.11) it turns out that this process is in 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 (3.11) allows us to conclude that $\tilde{u}$ is in Dom $\delta^{(0)}$.

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

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

Because of the chain rule we 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

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

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

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

which is the mentioned expression for π_1.

The second weight in (3.2) is

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

□

Theorem 3.2 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 (3.9). For any Borel measurable function f for which $f(S_t) \in L^2(\Omega), 0 < s < t < T$, and $\alpha \in \mathbb{R}$ it holds, under conditions (3.11), that*

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

where the Malliavin weight π differs from π_1 in (3.13) as follows

$$\pi = \pi_1 - \frac{1}{t-s}\frac{1}{V_s}\int_s^t \frac{V_{r-}}{\beta(r,S_{r-})}dW_r. \tag{3.14}$$

Proof. For the application of Theorem 3.2, we need the process

$$\hat{u}_r = \frac{V_{r-}}{V_s\beta(r,S_{r-})}\left\{\frac{1}{s}1_{\{r\leq s\}} - \frac{1}{t-s}1_{\{s\leq r\leq t\}}\right\} = \tilde{u}_r - \frac{V_{r-}}{V_s\beta(r,S_{r-})}\frac{1}{t-s}1_{\{s\leq r\leq t\}}. \tag{3.15}$$

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

$$\delta^{(0)}(\hat{u}) = \pi_1 + \delta^{(0)}\left(-\frac{V_{r-}}{V_s\beta(r,S_{r-})}\frac{1}{t-s}1_{\{s\leq r\leq t\}}\right)$$

$$= \pi_1 - \frac{1}{V_s}\frac{1}{t-s}\int_0^T \frac{V_{r-}}{\beta(r,S_{r-})}1_{\{s\leq r\leq t\}}dW_r + \frac{1}{t-s}\int_0^T \frac{V_{r-}}{\beta(r,S_{r-})}1_{\{s\leq r\leq t\}}D_r^{(0)}\frac{1}{V_s}dr,$$

where π_1 is defined in (3.13). The last term equals zero, since by Lemma 3.5 the Malliavin derivative $D_r^{(0)}(1/V_s)$ introduces a factor $1_{\{r\leq s\}}$. This concludes the proof. □

In the sequel, we present some models to illustrate our results from Propositions 3.6 and 3.7. The first two are defined by a linear SDE and the third one concerns stochastic volatility models.

3.2.1. Exponential Lévy model

We consider a stock price process S modeled by a stochastic exponential of a Lévy process, therefore let S satisfy the following linear SDE

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

where μ and $\beta > 0$ are constants. We assume that $\int_{\mathbb{R}_0}(e^z-1)^2\mathcal{L}(dz) < \infty$. All assumptions imposed on model (3.9) are satisfied. In this particular example, the first variation process V equals $V = S/s_0$ and $\zeta \equiv 0$ and conditions (3.11) are fulfilled. From Proposition 3.6, we find that the expression (3.12) holds with

$$\pi_1 = \frac{s_0}{sS_s}\left(\int_0^s \frac{1}{s_0\beta}dW_r + \int_0^s \frac{1}{s_0}dr\right) = \frac{s_0}{sS_s}\left(\frac{W_s}{s_0\beta} + \frac{s}{s_0}\right) = \frac{1}{S_s}\left(\frac{W_s}{s\beta} + 1\right), \tag{3.16}$$

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 (3.12) leads to

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

where $f \in C_b^1, 0 < s < t < T$, and $\alpha \in \mathbb{R}$.

On the other hand, we 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 (3.15), only by the second term in (3.14). In the present setting, this term equals

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

Hence combining this with (3.16) gives

$$\pi = \frac{1}{S_s}\left(\frac{W_s}{s\beta} - \frac{1}{\beta}\frac{W_t-W_s}{t-s} + 1\right) = \frac{1}{S_s}\left(\frac{tW_s-sW_t}{s(t-s)\beta} + 1\right). \tag{3.17}$$

For any Borel measurable function f guaranteeing that $f(S_t) \in L^2(\Omega), 0 < s < t < T$, and $\alpha \in \mathbb{R}$ the conditional expectation can be rewritten as

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

3.2.2. Merton model

In the light of the numerical experiments, we consider again the Merton model (2.6) which is a special case of the exponential Lévy model. The representations obtained through the MM, thus via Proposition 3.7 and Theorem 3.3, are as follows, for $0 < s < t < T$ and $\alpha \in \mathbb{R}$,

$$\mathbb{E}\big[f(S_t)|S_s = \alpha\big] = \frac{\mathbb{E}\big[f(S_t)H(S_s-\alpha)\pi_{s,t}\big]}{\mathbb{E}\big[H(S_s - \alpha)\pi_{s,t}\big]} = \frac{\mathbb{A}_{t,s}[f](\alpha)}{\mathbb{A}_{t,s}[1](\alpha)}, \tag{3.18}$$

$$\frac{\partial}{\partial\alpha}\mathbb{E}\big[f(S_t)|S_s = \alpha\big] = \frac{\mathbb{B}_{t,s}[f](\alpha)\mathbb{A}_{t,s}[1](\alpha) - \mathbb{A}_{t,s}[f](\alpha)\mathbb{B}_{t,s}[1](\alpha)}{\mathbb{A}_{t,s}[1](\alpha)^2}, \tag{3.19}$$

with

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

where the weight $\pi_{s,t}$ is given by (3.17) and

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

is obtained through similar computations as above in this section.

3.2.3. Stochastic volatility models

Let us consider the following model

$$\begin{cases} dS_t = \mu S_{t-}dt + v(Y_{t-})S_{t-}dW_t^{(1)} + \int_{\mathbb{R}_0}(e^z-1)S_{t-}\tilde{N}(dt,dz), \\ dY_t = a(t,Y_{t-})dt + b(t,Y_{t-})dW_t^{(2)} + \int_{\mathbb{R}_0}\psi(z)\tilde{N}(dt,dz), \end{cases} \quad (3.20)$$

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

$$dW_t^{(1)}dW_t^{(2)} = \rho dt, \quad \rho \in (-1,1). \quad (3.21)$$

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 $\mathbb{R}$, $\int_{\mathbb{R}_0}(e^z-1)^2\mathcal{L}(dz) < \infty$, and ψ is a function on $\mathbb{R}$ such that $\int_{\mathbb{R}_0}\psi^2(z)\mathcal{L}(dz) < \infty$. The process S may then perform the role of the stock price process, while $v(Y)$ is 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 [34]) and the Ornstein–Uhlenbeck stochastic volatility model (see [29, 35]).

From (3.21) we know there exists a Brownian motion $\tilde{W}$, independent of $W^{(2)}$, such that we express $W^{(1)}$ in terms of $\tilde{W}$ and $W^{(2)}$ by

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

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

$$V_t = \frac{S_t}{s_0}, \quad \beta(t,S_{t-}) = v(Y_{t-})S_{t-}\sqrt{1-\rho^2}, \quad \text{and} \quad \zeta_t = 0. \quad (3.22)$$

Applying Proposition 3.6, we find for the weights in representation (3.12)

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

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

$$\pi = \pi_1 - \frac{1}{S_s} \frac{1}{(t-s)\sqrt{1-\rho^2}} \int_s^t \frac{d\tilde{W}_r}{v(Y_{r-})}$$

$$= \pi_1 - \frac{1}{S_s} \frac{1}{(t-s)(1-\rho^2)} \left\{ \int_s^t \frac{dW_r^{(1)}}{v(Y_{r-})} - \rho \int_s^t \frac{dW_r^{(2)}}{v(Y_{r-})} \right\}.$$

Considering the model (3.20), we derived a representation for $\mathbb{E}[f(S_t)|S_s = \alpha]$. In the sequel we observe the conditional expectation

$$\mathbb{E}[w(Y_T)|S_T = \alpha], \tag{3.23}$$

for a certain Borel measurable function $w : \mathbb{R} \mapsto \mathbb{R}_0$. Our motivation to consider the latter expression comes from a paper by Martynov and Rozanova [36], where the authors are interested in the computation of conditional moments of Y. Thus, they consider (3.23) for $w(x) = x$ and $w(x) = x^2$. Moreover, in [37] van der Stoep et al. consider (3.23) for $w(x) = v^2(x)$, which is interesting for the study of stochastic local volatility.

We consider model (3.20) 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 $\tilde{W}$. Thus condition (3.4) is satisfied for any process u in Dom $\delta^{(0)}$. Thus when condition (3.1) is fulfilled, the conditional expectation can be written in the form (3.5). From expression (3.10) and previous derivations (3.22) we deduce that

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

Therefore the process satisfying condition (3.1) is given by

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

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

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

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

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

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 Moreni [38], a variance reduction technique based on importance sampling is proposed for the Monte Carlo pricing of American options via the Longstaff–Schwartz algorithm. For our conditional expectation representations, we study in subsection 4.1 the *localization* technique. This technique was used in Bally et al. [39] but we adapt it here to our setting. Moreover we include *control variates* to reduce the variance. We handle this approach in subsection 4.2.

4.1. Localisation

We adapt the localization technique of Bally et al. [39] for both methods; the conditional density method and the MM. For the proofs of the two following propositions, we refer the reader to Appendix C.

Proposition 4.1. *Assume the setting of Theorem 2.2. Then for any function* $\psi :$ $\mathbb{R} \mapsto [0, \infty)$ *satisfying* $\int_{\mathbb{R}} \psi(t)dt = 1$ *and for all* $\alpha \in$ Dom h, *we have*

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

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

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

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

Proposition 4.2. Assume the setting of Theorem 3.2, then for any continuous function with bounded derivative $\psi : \mathbb{R} \mapsto [0, \infty)$ satisfying $\int_{\mathbb{R}} \psi(t)dt = 1$ and for all $\alpha \in \mathbb{R}$, we have

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

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

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

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

Once we have introduced the localized versions of the representation formulas for the conditional expectation, one natural question arises, namely what is the optimal choice of the localizing function ψ. To find this optimal function, we assume that the additional constant c in the function H is zero, i.e. $H(x) = 1_{\{x \geq 0\}}$. 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 MM is considered. Then, practically speaking, an expectation of the form

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

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

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

In order to reduce the variance, the idea is to minimize the integrated mean squared error with respect to the localizing function ψ. Thus, we have to solve the following optimization problem (this criterion has been introduced by Kohatsu-Higa and Petterson [40])

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

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)

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

The choice of the optimal localizing function ψ is given in the following proposition. It is obvious that the optimal localization function will be different for the numerator and denominator since the optimization problem is different. (The proof in Bally et al. [39] can easily be extended to the current setting.)

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

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

The localizing 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, Bouchard and Warin [22] considered the exponential localizing function

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

where λ^* is given by (4.3). In paragraph 5.2.4 we show how the use of this function reduces the computational effort. We perform numerical experiments for both localizing functions in Section 5.

The representations for the derivatives in Theorems 2.3 and 3.3 have a localized version too. We state the localized versions as well as the choice of the optimal localizing 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 2.3, then for any function $\psi : \mathbb{R} \mapsto [0, \infty)$ satisfying $\int_{\mathbb{R}} \psi(t)dt = 1$ and for all $\alpha \in$ Dom h, we have*

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

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 3.3, then for any continuous function with bounded derivative* $\psi : \mathbb{R} \mapsto [0, \infty)$ *satisfying* $\int_{\mathbb{R}} \psi(t) dt = 1$ *and for all* $\alpha \in \mathbb{R}$, *we have*

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

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 localizing functions minimize the integrated variance

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

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

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

4.2. Control variate

Another approach to obtain variance reduction (besides localization) is to include a *control variate*, see, e.g., [19, Section 4.1]. The advantage of adding a control variate 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 variate 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 $P^{\text{Eu}}(t, \alpha)$. Let us define the function $P_\gamma(t, \alpha) := P(t, \alpha) - \gamma P^{\text{Eu}}(t, \alpha)$, for a real number γ close to 1. Then it holds that

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

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 $P(0, s_0) = P_\gamma(0, s_0) + \gamma P^{\text{Eu}}(0, s_0)$ and its delta equals $\Delta(0, s_0) = \Delta_\gamma(0, s_0) + \gamma \Delta^{\text{Eu}}(0, s_0)$. We rewrite this formula for the American option price as

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

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

Example 4.7. **(Merton model).** The European option price in the Merton model is derived in Merton [21]. In the setting described in paragraph 2.2.1, where the price process is as described in (2.6), the European put option price is given by the series

$$P^{\text{ Eu}}(t, S_t) = P^{\text{ Me}}(t, S_t) = \sum_{n=0}^{\infty} \frac{e^{-\mu(T-t)}\big(\mu(T-t)\big)^n}{n!} P_n^{\text{ BS}}(t, S_t). \tag{4.6}$$

Herein μ is the jump intensity of the Poisson process N introduced in (2.6) and $P_n^{\text{ BS}}(t, S_t)$ 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. Numerical experiments

In this section, we apply our results to estimate the price and delta of American options at time zero. We illustrate our methods with numerical results in a specified jump-diffusion model.

5.1. 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 $iT/n, i = 1, ..., n$. When the number of time periods increases, the Bermudan option converges to the American option (see Bally and Pagès [13]). Bermudan options can be priced through a Bellman dynamic programing principle, see Bellman [41] and Bally et al. [39]. 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

$$P(n\varepsilon, S_{n\varepsilon}) = \Phi(S_{n\varepsilon}) = \Phi(S_T),$$
$$P(k\varepsilon, S_{k\varepsilon}) = \max\left\{\Phi(S_{k\varepsilon}), e^{-r\varepsilon}\mathbb{E}\big[P\big((k+1)\varepsilon, S_{(k+1)\varepsilon}\big)|S_{k\varepsilon}\big]\right\}, \quad k = n-1, ..., 1, 0. \tag{5.1}$$

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

$$\Delta(\varepsilon, S_\varepsilon) = \begin{cases} e^{-r\varepsilon}\partial_\alpha\mathbb{E}\big[P(2\varepsilon, S_{2\varepsilon})|S_\varepsilon = \alpha\big]\big|_{\alpha=S_\varepsilon} & \text{if } P(\varepsilon, S_\varepsilon) > \Phi(S_\varepsilon), \\ \partial_\alpha\Phi(\alpha)\big|_{\alpha=S_\varepsilon} & \text{if } P(\varepsilon, S_\varepsilon) = \Phi(S_\varepsilon), \end{cases}$$

$$\Delta(0, s_0) = e^{-r\varepsilon}\mathbb{E}\big[\Delta(\varepsilon, S_\varepsilon)|S_\varepsilon = s_0\big]\big|_{s_0=S_\varepsilon}.$$

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 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.

Remark 5.1. **(Nested simulations).** To estimate the values for $P(k\varepsilon, S_{k\varepsilon})$ for $k \in \{0, ..., n-1\}$, the estimation of the conditional expectation requires the simulation of N values of $S_{(k+1)\varepsilon}$ starting from the value of $S_{k\varepsilon}$. In total this results into N^k used simulated values for $S_{k\varepsilon}$ for any k. The Bellman dynamic programing principle requires nested simulations to estimate a function of conditional expectations similarly to the compound option case in [19, Example 1.1.3] and illustrated in Figure 1.3 of that book. However, by the use of the obtained representations, the number of simulated values can be considerably reduced. The resulting algorithm needs one single set of simulated paths. For any $k \in \{0, ..., n-1\}$ only N simulated values for $S_{k\varepsilon}$ will be required.

5.2. Implementations for American put options in a Merton model

We consider American put options on a stock price process S defined by the Merton model (2.6). The put payoff function equals $\Phi(x) = (K-x)^+$. Since we want to compare our results to the analysis in Amin [42], 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 (4.6).

5.2.1. Representations
The conditional expectations and their derivatives in (5.1) and (5.2) can be estimated based on the representations we developed in the previous sections. In particular, in the present Merton setting, the representations are presented in paragraphs 2.2.1 and 3.2.2. Throughout this section we consider $H(x) = 1_{\{x\geq 0\}}$.

The regular expectations appearing in the representations (2.7), (2.8), (3.18), and (3.19) can easily be estimated by a Monte Carlo simulation. For example, consider the estimation of the numerator of representation (2.7). 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

$$\mathbb{E}\big[\cdot(S_t)H(S_s-\alpha)\pi_{s,t}\big] \approx \frac{1}{N}\sum_{j=1}^{N}\cdot\big(S_t^j\big)H(S_s^j-\alpha)\pi_{s,t}^j. \tag{5.3}$$

5.2.2. Variance reduction techniques

As discussed in paragraph 4.1, we include the localizing technique. The estimation (5.3) is then replaced by

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

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 include a control variate, see subsection 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 variate. In the current setting, the European option price and delta can be obtained through Merton's approach, see Example 4.7. Consider the algorithm for the price of Bermudan options (5.1). To introduce the control variate, we proceed in two steps. First we replace the put payoff function at each time $k\varepsilon, k=1,...,n$, by

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

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

5.2.3. Backward simulation

As remarked at the end of subsection 5.1 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 Bally et al. [39]). To simulate the compound Poisson process backwards in time, we base our method on results of Karatzas and Shreve [43] and Baldeaux [20]. 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 [20, Lemma 3.1].

Proposition 5.2. *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, see Karatzas and Shreve [43].

Proposition 5.3. *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 $N((j/k)z, (j/k)(k-j)\nu)$ distribution.

The backward simulation technique is interesting in numerical applications and following Baldeaux [20], this technique can also be derived for the Kou model, see Kou [44].

5.2.4. Reduction of computational effort

Bouchard and Warin [22] observed that the computational effort to estimate the American option prices by a MM is reduced by sorting the estimated stock prices. Consider the Bermudan dynamic programing algorithm (5.1). For a fixed k in $\{n-1, ..., 1\}$ we estimate the conditional expectations for $q = 1, ..., N$ by our representations, including localization, as follows

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

where

$$\mathbb{J}[\cdot]\left(S_{k\varepsilon}^{(q)}\right) = \frac{1}{N}\sum_{j=1}^{N} \cdot^{(j)}\left(\psi\left(S_{k\varepsilon}^{(j)} - S_{k\varepsilon}^{(q)}\right)Z_k^{(j)} + \pi_k^{(j)}\left(H\left(S_{k\varepsilon}^{(j)} - S_{k\varepsilon}^{(q)}\right) - \Psi\left(S_{k\varepsilon}^{(j)} - S_{k\varepsilon}^{(q)}\right)\right)\right).$$

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

$$\mathbb{J}[\cdot]\left(S_{k\varepsilon}^{(q)}\right) = \frac{1}{N}\sum_{j=1}^{N} \cdot^{(j)} H\left(S_{k\varepsilon}^{(j)} - S_{k\varepsilon}^{(q)}\right)e^{\lambda^* S_{k\varepsilon}^{(q)}} e^{-\lambda^* S_{k\varepsilon}^{(j)}}\left(\lambda^* Z_k^{(j)} + \pi_k^{(j)}\right).$$

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

$$\mathbb{E}\left[P\big((k+1)\varepsilon, S_{(k+1)\varepsilon}\big)|S_{k\varepsilon} = S_{k\varepsilon}^{s,(q)}\right]$$

$$= \frac{e^{\lambda_P^* S_{k\varepsilon}^{s,(q)}} \sum_{j=q}^{N} P\big((k+1)\varepsilon, S_{(k+1)\varepsilon}^{s,(j)}\big)e^{-\lambda_P^* S_{k\varepsilon}^{s,(j)}}\left(\lambda_P^* Z_k^{s,(j)} + \pi_k^{s,(j)}\right)}{e^{\lambda_1^* S_{k\varepsilon}^{s,(q)}} \sum_{j=q}^{N} e^{-\lambda_1^* S_{k\varepsilon}^{s,(j)}}\left(\lambda_1^* Z_k^{s,(j)} + \pi_k^{s,(j)}\right)}.$$

Thus for q going from N to 1, the sums in the numerator and denominator get only one additional term. Hence to estimate $\mathbb{E}[P((k+1)\varepsilon, S_{(k+1)\varepsilon})|S_{k\varepsilon} = S_{k\varepsilon}^{s,(q)}]$ for each q, we 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 Bermudan options. We compare our results for the prices to those reported by Amin [42] and to the regression based method introduced by Longstaff and Schwartz [8]. For the deltas, we compare our results to those of Hilliard and Schwartz [45]. To evaluate the accuracy of our representations, we consider European options

Table 1. Estimates of European put option prices $P^{Eu}(s, \alpha)$ (5.5) via the CDM and MM approach, with relative errors to the Merton prices in percentages.

time	$\alpha = 35$		$\alpha = 40$		$\alpha = 42$	
	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)

Parameters are given in (5.4), we fix $K = 45$.

since there are analytic formulas at hand in the Merton model, whereas there are non for Bermudan and American options. To overcome this problem, we compute confidence intervals for the prices as discussed in Section 3.1 in Bouchard and Warin [22].

The following parameter set for a put option on the underlying stock price process S is used,

$$S \text{ modelled by } (2.6): \quad s_0 = 40, r = 0.08, \beta^2 = 0.05, \mu = 5, \delta^2 = 0.05,$$
$$\text{put option}: \quad T = 1, K \in \{30, 35, 40, 45, 50\}. \tag{5.4}$$

5.3.1. Accuracy of the method

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

$$P^{Eu}(t, \alpha) = e^{-r(T-t)} \mathbb{E}[\Phi(S_T)|S_t = \alpha]. \tag{5.5}$$

The delta at time t equals

$$\Delta^{Eu}(t, \alpha) = e^{-r(T-t)} \frac{\partial}{\partial \alpha} \mathbb{E}[\Phi(S_T)|S_t = \alpha]. \tag{5.6}$$

The conditional expectations and their derivatives appearing in (5.5) and (5.6) are estimated following the techniques described in subsection 5.2 (except the control variate). As an example, we estimate the prices $P^{Eu}(t, \alpha)$ and deltas $\Delta^{Eu}(t, \alpha)$ of a European put option with maturity $T = 1$ and strike $K = 45$ on the underlying S described in (5.4), at times $t \in \{0.1, 0.2, ..., 0.9\}$ and for $\alpha \in \{35, 36, ..., 45\}$. 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 localizing technique. Each estimate results from the same set of $N = 5000000$ 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 Example 4.7. Similar results were obtained for the other values of $\alpha \in \{35, 36, ..., 45\}$. Table 2 shows the corresponding results for the

Table 2. Estimates of European put option deltas $\Delta^{Eu}(s, \alpha)$ (5.6) via the CDM and MM approach, with relative errors to the Merton deltas in percentages.

| time | $\alpha = 35$ | | $\alpha = 40$ | | $\alpha = 42$ | |
	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)

Parameters are given in (5.4), we fix $K = 45$.
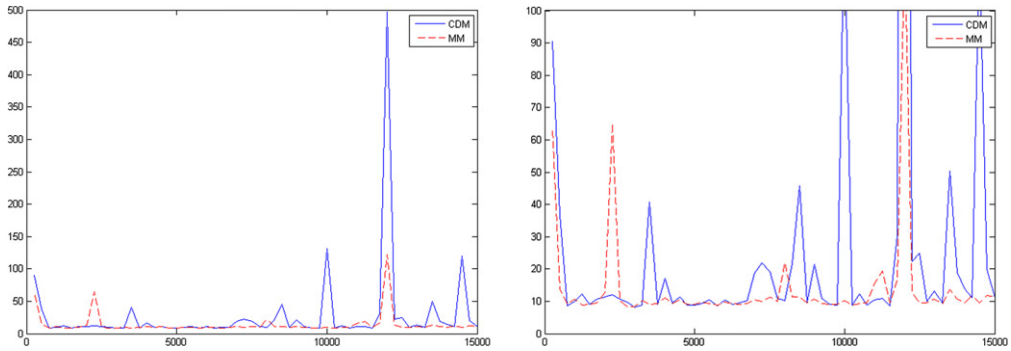
Figure 1. Estimates for the Bermudan put option price obtained through the CDM and MM representations *without* control variate and *without* localization technique, against the number of simulated paths. In the right graph the vertical axis is restricted to $[0, 100]$.

deltas. It turns out that the relative errors when comparing our approach to the one of Merton [21] are very small. Hence the algorithm we developed based on our representations is accurate for European options.

5.3.2. Results for Bermudan option prices and deltas

We consider a Bermudan put option on the stock price process S with parameters given in (5.4), the strike is fixed at $K = 45$. Amin [42] and Hilliard and Schwartz [45] developed a tree method to estimate Bermudan and American option prices. In the current setting their estimate for the option price equals 9.954. The Merton European option price at time zero equals 9.422. We choose $n = 10$. The dynamic programing algorithm presented in subsection 5.1 and our representations are used to estimate $P(0, s_0)$ and $\Delta(0, s_0)$.

Figures 1 and 2 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. Notice the difference in scale between the left and the right graph on the vertical axis. For $N = 250i, i = 1, ..., 60$, we simulated N paths of the underlying at the discrete time points jT/n, $j = 1, ..., n$, and we estimated the option price at time zero through the CDM and the MM, with and without control variate and optimal
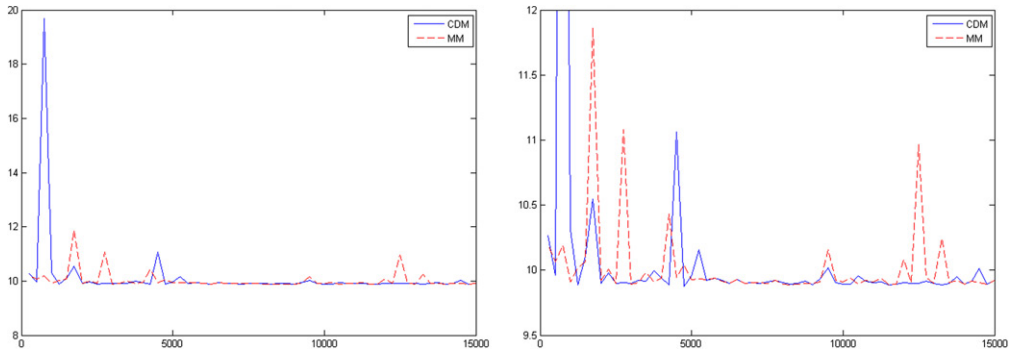
Figure 2. Estimates for the Bermudan put option price obtained through the CDM and MM representations *with* control variate and *with* localization technique, against the number of simulated paths. In the right graph the vertical axis is restricted to $[9.5, 12]$.

Table 3. Estimates of Bermudan put option prices and confidence intervals for parameter set (5.4), obtained through the sorted CDM and MM approach with control variate and exponential localization and by the Longstaff–Schwartz method [8] with a control variate (LSM). $n = 10$ and $N = 100000$.

	Price					Confidence Interval	
Strike	European	Amin	LSM	CDM	MM	CDM	MM
30	2.621	2.720	2.706	2.716	2.726	$[2.6765, 2.8609]$	$[2.6873, 2.9422]$
35	4.412	4.603	4.580	4.591	4.598	$[4.5223, 4.7555]$	$[4.5430, 4.8369]$
40	6.696	6.995	6.995	7.008	7.011	$[6.9054, 7.1821]$	$[6.9362, 7.2675]$
45	9.422	9.954	9.907	9.920	9.922	$[9.7799, 10.0857]$	$[9.8106, 10.1876]$
50	12.524	13.318	13.254	13.270	13.271	$[13.1457, 13.4084]$	$[13.1309, 13.5312]$

European prices computed via the Merton approach. Bermudan option price estimates from [42] (Amin).

localization technique. In case the European option is included as a control variate, 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.

Table 3 presents the estimated prices of the Bermudan put option with strikes 30, 35, 40, 45, and 50, obtained through the sorted CDM and MM approach including the control variate and the exponential localization function. For these estimates, a time discretisation is performed for $n = 10$ and 100000 paths were simulated. We include the estimates for the prices obtained respectively by Amin [42] and by using a regression based method with control variate as in Longstaff and Schwartz [8].

To evaluate the accuracy of our method, we include confidence intervals for the prices. The computation of these confidence intervals is discussed in [22, Section 3.1]. In our case, we include two confidence intervals which we compute using respectively the CDM and the MM. It turns out that the prices we find lay within the confidence intervals and thus they are in an acceptable range.

As described in paragraph 5.2.4, the computational effort is reduced when we perform a sorted algorithm. Table 4 presents the CPU time in seconds for the LSM, CDM, and MM, respectively. We do not compare the time to Amin [42] and Hilliard and Schwartz [45] methods since there is no clear indication about how long their

Table 4. CPU time in seconds for the different methods to compute the estimates of Bermudan put option prices for parameter set (5.4), obtained through the sorted CDM and MM approach with control variate and exponential localization and by the Longstaff–Schwartz method [8] with control variate (LSM). $n = 10$ and $N = 100000$.

	CPU time in seconds		
Strike	LSM	CDM	MM
30	1.56	2.20	2.18
35	1.74	2.17	2.15
40	2.01	2.15	2.17
45	2.29	2.09	2.16
50	2.45	2.09	2.17

Table 5. Estimates of Bermudan put option deltas for parameter set (5.4), obtained through the sorted CDM and MM approach with control variate and exponential localization. $n = 10$ and $N = 500000$.

	Delta			
Strike	European	H-S	CDM	MM
30	−0.1645	−0.1744	−0.1772	−0.1745
35	−0.2474	−0.2654	−0.2670	−0.2652
40	−0.3357	−0.3644	−0.3651	−0.3644
45	−0.4227	−0.4656	−0.4656	−0.4653
50	−0.5035	−0.5626	−0.5626	−0.5625

European deltas computed via the Merton approach. Bermudan option delta estimates from Hilliard and Schwartz [45] (H-S).

algorithms take. It turns out that when considering 10 time steps, our algorithms are comparably fast to the regression based algorithm. As analyzed in Bouchard and Warin [22], the complexity of our algorithms for both the MM and the CDM, when considering the sorted localizing method is of order $O(N \ln (N)^{(d-1)\vee 1})$, where d is the dimension of the underlying factor. The complexity of the LSM method depends on the choice of the basis of polynomials as well as the number and power of these polynomials. However, this is not the aim of the study of this paper and we refer to many articles in the literature that studied this issue for LSM. In, e.g. the article by Zhou [46], the author investigated whether there exists an optimal regression complexity in the LSM framework for options pricing. See also the analysis in Bouchard and Warin [22] and in the thesis of Plavsic [47]. Notice that in this latter thesis, the author pointed out that the computational time for LSM is proportional to the average number of in-the-money paths (i.e. the option's moneyness). That explains why for increasing values of the strike K, the CPU time for LSM increases in Table 4.

Table 5 presents the estimated deltas of the Bermudan put option with strikes 30, 35, 40, 45, and 50, obtained through the sorted CDM and MM approach including the control variate and the exponential localization function. For these estimates a time discretisation is performed for $n = 10$ and 500000 paths were simulated.

We conclude that the extension that we provided to the geometric Brownian motion model observed in Bally et al. [39] by adding normally distributed jumps to the driving process leads to numerical results which are in line with those found in the literature. As is the case in Bouchard and warin [22], the Monte Carlo method showed to be very promising in the context of jump-diffusions and could be further improved.

Table 6. Estimates of American put option prices for parameter set (5.8), obtained through the sorted CDM approach with exponential localization. $n = 10$ and $N = 10000$.

Strike	Price		CPU time in seconds	
	LSM	CDM	LSM	CDM
5.5	0.4847	0.5163	0.10	0.16
5.25	0.2439	0.2534	0.10	0.16
5	0.0191	0.0081	0.12	0.16
4.5	0	0	0.09	0.16

American option price estimates using the regression based method as in Longstaff and Schwartz [8].

5.4. Numerical results for the additive model

In this subsection, we consider the additive model (2.9) with X as described in (2.10) and Y given by

$$dY_t = -\lambda' Y_t dt + dL_t', \tag{5.7}$$

where L' is a compound Poisson process with intensity μ and exponentially distributed jumps with parameter ν.

The aim is to compute the price of Bermudan options as time-discrete approximations to American options written on such models using the conditional density representation. For the numerical example to be relevant for the energy markets, we note that Benth et al. [48] showed empirically that the spot model fits the Phelix Base electricity price index at the European Power Exchange (EEx) very well. In their article, they estimated the parameters in the suggested model and found the following values

$$
\begin{array}{ll}
X \text{ is as in } (2.10) \text{ with} & \lambda = 0.2008, a = 13.3009, b = 8.5689, \\
Y \text{ is as in } (5.7) \text{ with} & \lambda' = 0.3333, \mu = 20/250, \nu = 0.2, \\
\text{Put option} & T = 10/365, K \in \{5.5, 5.25, 5, 4.5\}.
\end{array} \tag{5.8}
$$

We compare our conditional density method to the regression based method as in Longstaff and Schwartz [8]. Table 6 represents the estimated prices of the American put option with strikes 5.5, 5.25, 5, 4.5 obtained through the sorted conditional density method including the exponential localization function. For these estimates, the number of exercise dates $n = 10$ and 10000 paths were simulated. For the same parameters, we include the estimates for the prices obtained by the regression method which we denote in the table by LSM and also the cost in seconds for each method.

In the case, the strike $K = 5$ and $n = 10$ the LSM indicates rank deficiency. That could explain the difference in the prices given respectively by the LSM and the CDM in the third row of Table 6. This is due to the fact that the LSM algorithm for pricing American options is unstable when the time parameter is small. See for example the discussion in Mostovyi [23].

The CDM with sorting as shown in Table 6 is comparably fast to the LSM. It can always be applied to compute the price of the American option in contrary to the LSM.

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 Fournié et al. [1] 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 MM. We applied the developed theory to the estimation of Bermudan and American option prices and their deltas. We used a localization technique and a control variate 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. Moreover, one can study parallelization for the Monte Carlo method with Malliavin calculus to improve the numerical results as discussed in Abbas-Turki and Lapeyre [18].

Finally, one can also study the MM for the computation of conditional expectations where conditioning is considered w.r.t. two random variables. One might think for example of the hedging of Asian options or spread options. In this context, the Longstaff–Schwartz method is not suited since the regression with two-dimensional explanatory data becomes numerically involved. The MM can be promising in this case.

Disclosure statement

Funding

References

[1]   Fournié, E., Lasry, J.-M., Lebuchoux, J., Lions, P.-L. (2001). Applications of Malliavin calculus to Monte Carlo methods in finance II. *Finance Stochast.* 5(2):201–236. DOI: 10.1007/PL00013529.

[2]   Montero, M., Kohatsu-Higa, A. (2003). Malliavin calculus applied to finance. *Physica A Statis. Mech. Appl.* 320:548–570. DOI: 10.1016/S0378-4371(02)01531-5.

[3]   Nualart, D. (2006). *The Malliavin Calculus and Related Topics.* Berlin, Germany: Springer.

[4] Di Nunno, G., Meyer-Brandis, T., Øksendal, B., Proske, F. (2005). Malliavin calculus and anticipative itô formulae for lévy processes. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* 8(2):235–258. DOI: 10.1142/S0219025705001950.

[5] von Sydow, L., Josef Höök, L., Larsson, E., Lindström, E., Milovanović, S., Persson, J., Shcherbakov, V., Shpolyanskiy, Y., Sirén, S., Toivanen, J., et al., Zhao, Y. (2015). BENCHOP-the BENCH marking project in option pricing. *Int. J. Comput. Math.* 92(12): 2361–2379. DOI: 10.1080/00207160.2015.1072172.

[6] Broadie, M., Glasserman, P. (1997). Pricing American-style securities using simulation. *J. Econ. Dynam. Control* 21(8-9):1323–1352. DOI: 10.1016/S0165-1889(97)00029-8.

[7] Barraquand, J., Martineau, D. (1995). Numerical valuation of high dimensional multivariate American securities. *J. Finance Quantitative Anal.* 30(3):383–405. DOI: 10.2307/2331347.

[8] Longstaff, F. A., Schwartz, E. S. (2001). Valuing American options by simulations: a simple least squares approach. *Rev. Finan. Stud.* 14(1):113–148. DOI: 10.1093/rfs/14.1.113.

[9] Shu, J., Gu, Y., Zheng, W. (2002). A novel numerical approach of computing American option. *Int. J. Found. Comput. Sci.* 13(05):685–693. DOI: 10.1142/S0129054102001394.

[10] Kovalov, P., Linetsky, V., Marcozzi, M. (2007). Pricing multi-asset American options: a finite element method-of-lines with smooth penalty. *J. Sci. Comput.* 33(3):209–237. DOI: 10.1007/s10915-007-9150-z.

[11] Zhang, B., Oosterlee, C. W. (2012). Fourier cosine expansions and put–call relations for Bermudan options. In: Carmona, A. R., Del Moral, P., Hu, P., and Oudjane, N., (eds.). *Numerical Methods in Finance: Bordeaux.* Berlin: Springer, pp. 323–350.

[12] Mrad, M., Touzi, N., Zeghal, A. (2006). Monte Carlo estimation of a joint density using Malliavin calculus, and application to American options. *Comput. Econ.* 27(4):497–531. DOI: 10.1007/s10614-005-9005-3.

[13] Bally, V., Pagès, G. (2003). Error analysis of the quantization algorithm for obstacle problems. *Stochastic Process. Appl.* 106(1):1–40. DOI: 10.1016/S0304-4149(03)00026-7.

[14] Benth, F. E., Di Nunno, G., Khedher, A. (2010). 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.

[15] Petrou, E. (2008). Malliavin calculus in lévy spaces and applications to finance. *Electron. J. Probab.* 13(0):852–879. DOI: 10.1214/EJP.v13-502.

[16] Davis, M. H. A., Johansson, M. P. (2006). Malliavin Monte Carlo Greeks for jump-diffusions. *Stochastic Process. Appl.* 116(1):101–129. DOI: 10.1016/j.spa.2005.08.002.

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

[18] Abbas-Turki, L. A., Lapeyre, B. (2012). American options by Malliavin calculus and non-parametric variance and bias reduction methods. *SIAM J. Finan. Math.* 3(1):479–510. DOI: 10.1137/11083890X.

[19] Glasserman, P. (2004). *Monte Carlo Methods in Financial Engineering.* Berlin: Springer.

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

[21] Merton, R. (1976). Option pricing when underlying stock returns are discontinuous. *J. Finan. Econ.* 3(1-2):125–144. DOI: 10.1016/0304-405X(76)90022-2.

[22] Bouchard, B., Warin, X. (2012). Monte-Carlo valuation of American options: Facts and new algorithms to improve existing methods. *Num. Meth. Finan.* Springer Proceedings in Mathematics 12:215–255.

[23] Mostovyi, O. (2013). On the stability the least squares Monte Carlo. *Optim. Lett.* 7(2): 259–265. DOI: 10.1007/s11590-011-0414-z.

[24] Protter, P. (2005). *Stochastic Integration and Differential Equations.* 2nd edn, Version 2.1. No. 21 in Stochastic Modelling and Applied Probability. Berlin: Springer.

[25] Raju, C. K. (1982). Products and compositions with the Dirac Delta function. *J. Phys. A Math. Gen.* 15(2):381–396. DOI: 10.1088/0305-4470/15/2/011.

[26] Daveloose, C. (2016). *Essays on Pricing and Hedging in Markets with Imperfections*. Ph.D. thesis, Ghent University.

[27] Benth, F. E., Kallsen, J., Meyer-Brandis, T. (2007). A non-Gaussian Ornstein-Uhlenbeck process for electricity spot price modeling and derivatives pricing. *Appl. Math. Finance* 14(2):153–169. DOI: 10.1080/13504860600725031.

[28] Sato, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions*. No. 68 in Cambridge Studies in Advanced Mathematics. Cambridge: Cambridge University Press.

[29] Barndorff-Nielsen, O. E., Shephard, N. (2001). Modelling by Lévy processes for financial econometrics. Barndorff-Nielsen, O., Mikosch, T., and Resnick, S. (eds.), *Lévy Processes: Theory and Applications*. Basel, Switzerland: Birkhäuser, pp. 283–318.

[30] Brewer, D. W., Tubbs, J. D., Smith, O. E. (1987). A differential equations approach to the modal location for a family of bivariate gamma distributions. *J. Multivariate Anal.* 21(1): 53–66. DOI: 10.1016/0047-259X(87)90098-4.

[31] Bowman, F. (1958). *Introduction to Bessel Functions*. New York: Dover Publications Inc.

[32] Solé, J. L., Utzet, F., Vives, J. (2007). Canonical Lévy process and Malliavin calculus. *Stochastic Process. Appl.* 117(2):165–187. DOI: 10.1016/j.spa.2006.06.006.

[33] Ikeda, N., Watanabe, S. (1981). *Stochastic Differential Equations and Diffusion Processes*. New York: Elsevier Science.

[34] Bates, D. (1996). Jumps and stochastic volatility: exchange rate processes implicit in Deutsche Mark options. *Rev. Finan. Stud.* 9(1):69–107. DOI: 10.1093/rfs/9.1.69.

[35] Barndorff-Nielsen, O. E., Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *J. Royal Stat. Soc. B* 64(2):253–280. DOI: 10.1111/1467-9868.00336.

[36] Martynov, M., Rozanova, O. (2013). On dependence of volatility on return for stochastic volatility models. *Stochastics* 85(5):917–927. DOI: 10.1080/17442508.2012.673616.

[37] van der Stoep, A. W., Grzelak, L. A., Oosterlee, C. W. (2014). The Heston stochastic-local volatility model: efficient Monte Carlo simulation. *Int. J. Theor. Appl. Finan.* 17(7): 14050045.

[38] Moreni, N. (2004). A variance reduction technique for American option pricing. *Physica A Statis. Mech. Appl.* 338(1-2):292–295. DOI: 10.1016/j.physa.2004.02.055.

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

[40] Kohatsu-Higa, A., Pettersson, R. (2002). Variance reduction methods for simulation of densities on wiener space. *SIAM J. Numer. Anal.* 40(2):431–450. DOI: 10.1137/S0036142901385507.

[41] Bellman, R. (1957). *Dynamic Programming*. Princeton, NJ: Princeton University Press.

[42] Amin, K. I. (1993). Jump diffusion option valuation in discrete time. *J. Finan.* 48(5): 1833–1863. DOI: 10.2307/2329069.

[43] Karatzas, I., Shreve, S. (1991). *Brownian Motion and Stochastic Calculus*. Berlin: Springer.

[44] Kou, S. (2002). A jump-diffusion model for option pricing. *Manag. Sci.* 48(8):1086–1101. DOI: 10.1287/mnsc.48.8.1086.166.

[45] Hilliard, J., Schwartz, A. (2005). Pricing European and American derivatives under a jump-diffusion process: a bivariate tree approach. *J. Finan. Quant. Anal.* 40(03):671–691.

[46] Zhou, Y. (2004). *On the existence of an optimal regression complexity in the Least-Squares Monte Carlo (LSM) framework for options pricing.* 39th Actuarial Research Conference, Iowa Memorial Union, The University of Iowa, Iowa City, April.

[47] Plavsic, M. (2011). *Pricing American options - aspects of computation*. Ph.D. thesis, Imperial College London.

[48] Benth, F. E., Kallsen, J., Meyer-Brandis, T. (2007). A critical empirical study of two electricity spot price modeling and derivatives pricing. *Appl. Math. Finance* 14(2):153–169. DOI: 10.1080/13504860600725031.

Appendix A. Proof in Section 2

Proof of Theorem 2.2. We consider the following approximation

$$\mathbb{E}\big[f(F)|G=\alpha\big] = \lim_{\varepsilon\to 0+}\mathbb{E}\big[f(F)|G\in]\alpha-\varepsilon,\alpha+\varepsilon[\big]$$

$$= \lim_{\varepsilon\to 0+}\frac{\mathbb{E}\big[f(F)1_{]-\varepsilon,\varepsilon[}(G-\alpha)\big]}{\mathbb{E}\big[1_{]-\varepsilon,\varepsilon[}(G-\alpha)\big]},$$

and know by the conditional density method and Assumptions 2.1(1) that

$$\mathbb{E}[f(F)1_{]-\varepsilon,\varepsilon[}(G-\alpha)] = \mathbb{E}[\mathbb{E}[f(g_1(X,Y))1_{]-\varepsilon,\varepsilon[}(g_2(U,V)-\alpha)|\sigma(Y,V)]]$$

$$= \mathbb{E}\left[\int_{\mathbb{R}^2}f(g_1(x,Y))1_{]-\varepsilon,\varepsilon[}(g_2(u,V)-\alpha)p_{(X,U)}(x,u)dxdu\right].$$

We introduce two functions

$$\Phi(x;\varepsilon,\alpha,h) := 1_{]-\varepsilon,\varepsilon[}\big(h^{-1}(x)-\alpha\big),$$

$$\Psi(x;\varepsilon,\alpha,h) := \int_{-\infty}^{x}\Phi(z;\varepsilon,\alpha,h)dz + 2\varepsilon c', \quad\text{where } c' = ch'(\alpha).$$

By relation (2.2) and integration by parts we have

$$\int_{\mathbb{R}}1_{]-\varepsilon,\varepsilon[}(g_2(u,V)-\alpha)p_{(X,U)}(x,u)du$$

$$= \int_{\mathbb{R}}1_{]-\varepsilon,\varepsilon[}\big(h^{-1}\big(u+g^*(V)\big)-\alpha\big)p_{(X,U)}(x,u)du$$

$$= \int_{\mathbb{R}}\Phi\big(u+g^*(V);\varepsilon,\alpha,h\big)p_{(X,U)}(x,u)du$$

$$= -\int_{\mathbb{R}}\Psi\big(u+g^*(V);\varepsilon,\alpha,h\big)\frac{\partial}{\partial u}p_{(X,U)}(x,u)du.$$

Combining the previous steps leads to the observation that

$$\mathbb{E}\big[f(F)|G=\alpha\big] = \lim_{\varepsilon\to 0+}\frac{\mathbb{E}\left[\int_{\mathbb{R}^2}f(g_1(x,Y))\frac{1}{2\varepsilon}\Psi\big(u+g^*(V);\varepsilon,\alpha,h\big)\left(-\frac{\partial}{\partial u}p_{(X,U)}(x,u)\right)dxdu\right]}{\mathbb{E}\left[\int_{\mathbb{R}^2}\frac{1}{2\varepsilon}\Psi\big(u+g^*(V);\varepsilon,\alpha,h\big)\left(-\frac{\partial}{\partial u}p_{(X,U)}(x,u)\right)dxdu\right]}. \qquad (A.1)$$

Because of the facts that $\frac{1}{2\varepsilon}\Psi(u+g^*(V);\varepsilon,\alpha,h)$ is bounded in ε, that $\mathbb{E}[f^2(F)]<\infty$, and that $\mathbb{E}[\pi^2_{(X,U)}]<\infty$, the limit can be brought inside the integrals in the numerator and denominator. Thus we determine

$$\lim_{\varepsilon\to 0+}\frac{1}{2\varepsilon}\Psi\big(u+g^*(V);\varepsilon,\alpha,h\big) = \lim_{\varepsilon\to 0+}\frac{1}{2\varepsilon}\int_{-\infty}^{u+g^*(V)}\Phi(z;\varepsilon,\alpha,h)dz + c'$$

$$= \lim_{\varepsilon\to 0+}\frac{1}{2\varepsilon}\int_{-\infty}^{u+g^*(V)}1_{]-\varepsilon,\varepsilon[}\big(h^{-1}(z)-\alpha\big)dz + c' = \int_{-\infty}^{u+g^*(V)}\delta_0\big(h^{-1}(z)-\alpha\big)dz + c'$$

$$= \int_{-\infty}^{u+g^*(V)}\delta_0(z-h(\alpha))h'(\alpha)dz + c' = h'(\alpha)1_{\{h(\alpha)\in]-\infty,u+g^*(V)]\}} + c'$$

$$= h'(\alpha)1_{\{\alpha\le h^{-1}(u+g^*(V))\}} + c' = h'(\alpha)1_{\{g_2(u,V)-\alpha\ge 0\}} + c' = h'(\alpha)H(g_2(u,V)-\alpha).$$

The numerator of (A.1) therefore equals

$$\mathbb{E}\left[\int_{\mathbb{R}^2} f(g_1(x, Y)) H(g_2(u, V) - \alpha)\left\{-\frac{\partial}{\partial u}\log p_{(X,U)}(x, u)\right\} p_{(X,U)}(x, u)\,dx\,du\right] h'(\alpha)$$

$$= \mathbb{E}\left[\mathbb{E}\left[f(F)H(G - \alpha)\left\{-\frac{\partial}{\partial u}\log p_{(X,U)}(X, U)\right\}|\sigma(Y, V)\right]\right] h'(\alpha)$$

$$= \mathbb{E}\left[f(F)H(G - \alpha)\left\{-\frac{\partial}{\partial u}\log p_{(X,U)}(X, U)\right\}\right] h'(\alpha).$$

Simiatolarly the denominator of (A.1) equals

$$\mathbb{E}\left[H(G - \alpha)\left\{-\frac{\partial}{\partial u}\log p_{(X,U)}(X, U)\right\}\right] h'(\alpha).$$

This concludes the proof of Theorem 2.2. □

Appendix B. Malliavin calculus

In this article, we make use of the Malliavin calculus as defined in Petrou [15]. 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 $f \in C_b^1$. 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

  $$\left|\mathbb{E}\left[\int_0^T u_t D_t^{(0)} F\,dt\right]\right| \leq C\|F\|_{L^2(\Omega)}$$

  for all $F \in \mathbb{D}^{(0)}$, is the domain of $\delta^{(0)}$, denoted by $\text{Dom }\delta^{(0)}$. For every $u \in \text{Dom }\delta^{(0)}$ we define $\delta^{(0)}(u)$ as the Skorohod integral in the Wiener direction of u by

  $$\mathbb{E}\left[F\delta^{(0)}(u)\right] = \mathbb{E}\left[\int_0^T u_t D_t^{(0)} F\,dt\right],$$

  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 $u \in \text{Dom }\delta^{(0)}$. Then $Fu \in \text{Dom }\delta^{(0)}$ and

  $$\delta^{(0)}(Fu) = F\delta^{(0)}(u) - \int_0^T u_t D_t^{(0)} F\,dt$$

  if and only if the right hand side of the latter 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 $u \in \text{Dom }\delta^{(0)}$ and the Skorohod integral coincides with the Ito-integral

  $$\delta^{(0)}(u) = \int_0^T u_t\,dW_t.$$

Appendix C. Proofs localization

Proof of Proposition 4.1. Adding and subtracting the same expression, keeping in mind that $\psi = \Psi'$, we get for the numerator of representation (2.4), the following equality

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

The last term equals

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

Using the integration by parts formula, we get

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

and the result follows.

□

Proof of Proposition 4.2. Adding and subtracting the same expression, keeping in mind that $\psi = \Psi'$, we get for the numerator of representation (3.5) that

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

Applying the same arguments as in Theorem 3.2, we show that the last term equals

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

and the result follows.

□