Finance Stochast. 5, 201–236 (2001)

Finance and Stochastics

Applications of Malliavin calculus to Monte-Carlo methods in finance. II

Eric Fournié,[1,3] **Jean-Michel Lasry,**[1] **Jérôme Lebuchoux,**[1] **Pierre-Louis Lions**[2]

[1] PARIBAS Capital Markets, 10, Harewood Avenue, NW1 6AA London, England
[2] Ceremade, UMR 9534, Université Paris-Dauphine, Place Maréchal de Lattre de Tassigny, 75775 Paris Cedex 16, France (e-mail: mazzella@ceremade.dauphine.fr)
[3] Laboratoire de Probabilités, Université Pierre et Marie Curie, 4, Place Jussieu, 75252 Paris Cédex 05, France

Abstract. This paper is the sequel of Part I [1], where we showed how to use the so-called Malliavin calculus in order to devise efficient Monte-Carlo (numerical) methods for Finance. First, we return to the formulas developed in [1] concerning the "greeks" used in European options, and we answer to the question of optimal weight functional in the sense of minimal variance. Then, we investigate the use of Malliavin calculus to compute conditional expectations. The integration by part formula provides a powerful tool when used in the framework of Monte Carlo simulation. It allows to compute everywhere, on a single set of trajectories starting at one point, solution of general options related PDEs.

Our final application of Malliavin calculus concerns the use of Girsanov transforms involving anticipating drifts. We give an example in numerical Finance of such a transform which gives reduction of variance via importance sampling.

Finally, we include two appendices that are concerned with the PDE interpretation of the formulas presented in [1] for the delta of a European option and with the connections between the functional dependence of some random variables and their Malliavin derivatives.

Key words: Monte Carlo methods, Malliavin calculus, hedge ratios and greeks, conditional expectations, PDE, anticipative Girsanov transform, functional dependence

JEL Classification: G13

Mathematics Subject Classification (1991): 60H07, 60J60, 65C05

1 Introduction

As is well-known, the valuation and the hedging of most financial products involve expected values, and their differentials, of functionals of Brownian motions

Manuscript received: February 1999; final version received: January 2000

and related stochastic processes. A general and natural approach to the practical (numerical) computation of such quantities is the Monte-Carlo method. The advantages of Monte-Carlo simulations on other numerical approaches, which are essentially numerical partial differential equations approaches (we exclude explicit formula which are limited to a few simple examples), are the flexibility and in particular the possibility of dealing with path-dependent products, and the potential for computations in higher dimensions. The drawbacks of Monte-Carlo approaches are i) the relative slow convergence and the lack of precision, ii) the inadequacy for the treatment of american options.

Our goal in this series of papers (Part I [1], this article and forthcoming publications) is to provide some cures for these drawbacks. More precisely, we emphasize the usefulness of tools of modern probability theory, in particular of the Malliavin calculus that we briefly presented in Part I (for more details on Malliavin calculus, the reader may consult [5], [6], [7], [10]). Let us also mention in passing the possible use of large deviations theory (see [3], [2]). In Part I [1], we showed how one can use Malliavin calculus to write down explicit probabilistic formulas for the greeks – i.e. differentials of the values of various European options with respect to various parameters – which are needed for hedging. The formula turn out to be expectations of the original pay off multiplied by some explicit weights. Such representations of greeks allow for a direct and efficient Monte-Carlo simulation and we presented in [1] various experiments showing the improvements in speed and precision of such an approach compared to a traditional "difference quotients" approach.

In this article, we present several other applications of Malliavin calculus. First of all, we analyze the question of the weights appearing in the representations of greeks mentioned above : indeed, see [1], those weights are non unique and thus we need to choose the "best" one. In [1], we chose in some sense the simplest one. We study, in Sect. 2 below, the selection of an optimal weight where optimal means minimal variance. Next, in Sect. 3, we develop an idea which is briefly mentioned in Part I [1] which consists in applying the Malliavin calculus, or integration by parts, on a piece of the functional whose expectation is to be computed. This allows to write several probabilistic expressions that share the same expectation and thus to look for a reduced variance.

Section 4 is concerned with what we believe to be one of the most promising applications of Malliavin calculus to numerical Finance namely the representation of conditional expectations. Indeed, conditional expectations are extremely difficult to compute by traditional Monte-Carlo simulations for obvious reasons – "almost all" paths generated by the simulation will miss the event involved in the conditional expectation. However, the Malliavin integration by parts allows to obtain different representations that can be computed by Monte-Carlo simulations ! In other words, we can use paths that do not go through or to the right zone. We illustrate this general observation by two applications which are relevant for Finance. First, we show how one generation of paths emanating from a fixed initial position of the underlying asset can be used to compute the price of a European option at any later time and at any position. This should clearly pave the way to a full Monte-Carlo approach to American options and we shall come back to this issue in a forthcoming publication. Next, as an example, we consider a stochastic volatility model and the price of a European option conditioned by the value of the volatility at same time(s). Once more, the aforementioned integration by parts allows us to compute efficiently such conditional values by a traditional Monte-Carlo simulation.

Our final application of Malliavin calculus concerns the use of Girsanov transforms involving anticipating drifts. This is described in Sect. 5 and we just allude here to the interest of such transforms. Girsanov transforms with non-anticipating drifts are well-known to be useful in various contexts for numerical Finance : reduction of variance via importance sampling for instance, change of numeraire. We show here that very general drifts can be used and we present one simple application to the pricing of European options. Indeed, it is natural to use a drift that depends on the terminal position in order to drive the process to the zone of interest for the computation of the payoff. One can think of a simple call where all paths leading to a terminal position below the strike do not contribute to the expectation.

Finally, we also include two rather more mathematical appendices very much related to the facts shown in [1] and here, that are concerned with the PDE interpretation of the formula presented in [1] for the delta of a European option and with the connections between the functional dependence of some random variables and their Malliavin derivative.

We conclude this long introduction with a few notations that we keep throughout the article. We denote by $(W_t)_{t\geq 0}$ an N-dimensional Brownian motion defined on a complete probability space $(\Omega, \mathscr{F}, P)$ and by $(\mathscr{F}_t)_{t\geq 0}$ the augmentation with respect to P of the filtration generated by W. Finally, we shall write indifferently X_t or $X(t)$ for any process X defined on $(\Omega, \mathscr{F}, \mathscr{F}_t, W_t, P)$.

2 Optimal weights for greeks

As recalled in the introduction, we obtained in Part I [1] representation formulas for the so-called greeks of European options. All these formula can be summarized as follows : let $T > 0$, let F be a "smooth" $\mathscr{F}$-measurable random variable with values in $\mathbb{R}^m$ where $m \geq 1$, and we shall need to make precise what we really mean by "smooth". Furthermore, we assume that F depends upon a real-valued parameter λ, $F = F(0)$, and $\lambda^{-1}(F(\lambda) - F(0)) \rightarrow G$ as $\lambda \rightarrow 0$ $(\lambda \neq 0)$ in $L^1(\Omega)$ (for instance). We then consider

$$\frac{\partial}{\partial \lambda} \mathbf{E}\left[\Phi(F)\right]\bigg|_{\lambda=0} = \mathbf{E}\left[\Phi'(F).G\right] \tag{2.1}$$

where Φ is an arbitrary smooth function over $\mathbb{R}^m$.

For financial applications, F is the position at time T of the underlying asset assumed to be a diffusion process defined, for instance, by

$$dX_t = \sigma(X_t)dW_t + b(X_t)dt, \quad X_0 = x \in \mathbb{R}^d \tag{2.2}$$

where $d \geq 1$, σ is $d \times N$ matrix-valued function, b is a function from $\mathbb{R}^d$ into $\mathbb{R}^d$ and we assume that σ and b are Lipschitz on $\mathbb{R}^d$. For instance, in the usual Black-Scholes model, $d = 1$, $\sigma(S) = \sigma S$ and σ is the volatility, $b(S) = bS$. But the general formulation (2.2) obviously allows for more general models or even for "extended" procesus such as $X_t = (S_t, M_t)$ where

$$\begin{aligned} dS_t &= \sigma S_t dW_t + bS_t dt \\ dM_t &= S_t dt \end{aligned}$$

which are relevant for Asian options.
Finally, Φ is the pay-off and various choices of Φ correspond to various options (call, put, digital, knock out, etc). Let us also mention in passing that we could consider as well non homogeneous models where σ and b depend upon t.

2.1 Minimal weights

We showed in Part I [1] various situations in which one can rewrite (2.1) as

$$\mathbf{E}\left[\Phi'(F).G\right] = \mathbf{E}\left[\Phi(F)\pi\right] \tag{2.3}$$

for any smooth Φ, where π is a weight that we computed explicitly in various cases.

The existence of such a weight is straightforward if F is smooth i.e. if the law of $F(\lambda)$ admits a smooth, positive density f_λ on $\mathbb{R}^m$ which depends smoothly on λ. Indeed, under natural conditions that do not need to be made precise, we have

$$\begin{aligned} \frac{\partial}{\partial\lambda}\mathbf{E}\left[\Phi(F(\lambda))\right] &= \frac{\partial}{\partial\lambda}\int_{\mathbb{R}^m} \Phi(z) f_\lambda(z)dz \\ &= \int_{\mathbb{R}^m} \Phi(z)\left(\frac{\partial}{\partial\lambda}\log f_\lambda\right)(z)\; f_\lambda(z)dz \\ &= \mathbf{E}\left[\Phi(F(\lambda))\pi\right] \end{aligned}$$

with $\pi = \left(\frac{\partial}{\partial\lambda}\log f_\lambda\right)(F(\lambda)$. Hence, (2.3) holds with $\pi_0 = \left(\frac{\partial}{\partial\lambda}\log f_\lambda\right)\Big|_{\lambda=0}(F)$. Let us notice, for future use, that this construction not only yields a weight π but also a weight which is measurable with respect to the σ-field generated by F.

However, the preceding weight is theoretical since, in general the density f_λ is not known (and not even available numerically even though we shall see in Sect. 4 a general approach which can be applied in particular to the computation of $f_\lambda(z)$). Let us also observe that there are many weights π such that (2.3) holds for all Φ : more precisely, the set of such π is exactly given by

$$\mathscr{W} = \left\{\pi \mid \mathbf{E}\left[\pi|\sigma(F)\right] = \pi_0\right\} \tag{2.4}$$

always assuming that all random variables π are integrable, and where $\sigma(F)$ denote the σ-field generated by F.

Obviously, for any practical application, one needs to determine an element of $\mathscr{W}$ which is explicit or that can be computed by a standard Monte-Carlo procedure. We presented in Part I [1] several examples relevant to Financial Engineering where this can be done using Malliavin calculus. These examples can be incorporated in the previous abstract framework as follows. We now assume that F is "smooth in Malliavin sense" i.e. admits a Malliavin derivative $(D_tF,\ t \in [0,T])$ in $L^2(\Omega \times (0,T))^N$ (for instance) and that D_tF is "not degenerate" and more precisely that we can find at least one "smooth" process u_t with values in $\mathbb{R}^N$ such that

$$\mathbf{E}\left[\int_0^T D_tF.u_t dt \,|\, \sigma(F)\right] = \mathbf{E}\left[G \,|\, \sigma(F)\right] \tag{2.5}$$

Then, we have for any such u (and for any Φ)

$$\begin{aligned}
\mathbf{E}\left[\Phi'(F)G\right] &= \mathbf{E}\left[\Phi'(F)\mathbf{E}(G|\sigma(F))\right] \\
&= \mathbf{E}\left[\Phi'(F)\mathbf{E}\left(\int_0^T D_tF.u_t dt \,|\, \sigma(F)\right)\right] \\
&= \mathbf{E}\left[\phi'(F)\int_0^T D_tF.u_t dt\right] \\
&= \mathbf{E}\left[\int_0^T D_t\Phi(F).u_t\, dt\right] \\
&= \mathbf{E}\left[\phi(F)\pi\right]
\end{aligned}$$

where $\pi = \delta(u_t)$ is the Skorohod integral on $[0,T]$ of u. In other words, $\delta(u_t)$ belongs to $\mathscr{W}$. Conversely, if $\pi \in \mathscr{W} \cap \mathbb{H}^1$ – we denote as is customary in stochastic analysis by $\mathbb{H}^1$ the space of L^2 random variables X such that $D_tX \in L^2(\Omega \times (0,T))^N$ – then $\pi = \delta(u_t)$ for some adapted u_t satisfying (2.5) and this is the case in particular if $\pi = \pi_0$. Indeed, taking $\Phi \equiv 1$ in (2.3), we see that $\mathbf{E}(\pi) = 0$ and thus

$$\pi = \int_0^T \mathbf{E}\left[D_t\pi|\mathscr{F}_t\right].dW_T = \delta(u_t)$$

with $u_t = \mathbf{E}\left[D_t\pi|\mathscr{F}_t\right]$. In addition, we have for any Φ using the same string of equalities as before

$$E\left[\Phi'(F)\mathbf{E}(G|\,\sigma(F))\right] = \mathbf{E}\left[\Phi'(F)\mathbf{E}\left(\int_0^T D_tF.u_t dt|\sigma(F)\right)\right]$$

hence (2.5) holds.

Our final observation consists in remarking that π_0 is, among all weights, the one which yields the mininum variance i.e. π_0 is a minimum over all $\pi \in \mathscr{W}(\cap\mathbb{H}^1)$ of the convex functional

$$\mathscr{V}(\pi) = \mathbf{E}\left[|\Phi(F)\pi - \mathbf{E}[\Phi'(F)G]\,|^2\right] \tag{2.6}$$

Indeed, we have trivially

$$\begin{aligned}\mathscr{V}(\pi) &= \mathbf{E}\left[\Phi^2(F)\pi^2\right] - \mathbf{E}\left[\Phi(F)\pi_0\right]^2 \\ &= \mathbf{E}\left[\Phi^2(F)(\pi-\pi_0)^2\right] + \mathscr{V}(\pi_0)\end{aligned}$$

In conclusion, we have shown the

Proposition 2.1 *Under the above assumptions, the set* $\mathscr{W} \cap \mathbb{H}^1$ *is equal to*

$$\left\{\pi = \delta(u_t) \in \mathbb{H}^1 \mid u_t \text{ satisfies } (2.5)\right\}$$

and π_0 *is a minimum of* $\mathscr{V}(\pi)$ *over that set.*

2.2 Euler-Lagrange equation

Obviously, the set of minima of (2.6) is given by $\{\pi \in \mathscr{W} \mid \pi = \pi_0$ a.s. on $\{\Phi(F) \neq 0\}\}$. For any such minimum $\overline{\pi} = \delta(\overline{u})$ in $\mathscr{W} \cap \mathbb{H}^1$, the Euler-Lagrange equation holds for $\overline{\pi}$ or equivalently for $\overline{u}$ namely we have for all "smooth " v, such that

$$\mathbf{E}\left[\int_0^T D_tF.v_t dt \mid \sigma(F)\right] = 0 \tag{2.7}$$

the following equality

$$\mathbf{E}\left[\left\{\Phi(F)\overline{\pi} - \mathbf{E}\left(\Phi'(F)G\right)\right\}\Phi(F)\delta(v)\right] = 0 \tag{2.8}$$

or equivalently by integrating by parts

$$\mathbf{E}\left\{\int_0^T v_t.D_tF\Phi'(F)\left[2\Phi(F)\overline{\pi} - \mathbf{E}(\Phi'(F)G)\right] + v_t.D_t\overline{\pi}\Phi^2(F)\,dt\right\} = 0$$

or in view of (2.7)

$$\mathbf{E}\left\{\int_0^T v_t.D_tF\overline{\pi}(\Phi^2)'(F) + v_t.D_t\overline{\pi}\Phi^2(F)\,dt\right\} = 0 \tag{2.9}$$

Next, we remark that, at least formally, we may choose for each "smooth" w_t

$$v_t = w_t - D_tF\,\mathbf{E}\left[\int_0^T D_tF.w_t dt \,|\, \sigma(F)\right]\mathbf{E}\left[\int_0^T |D_tF|^2 dt \,|\, \sigma(F)\right]^{-1}$$

at least if $\int_0^T |D_tF|^2 dt$ does not "vanish too much". Inserting this choice in (2.9), we find

$$\mathbf{E}\left[\int_0^T w_t.D_tF\overline{\pi}(\Phi^2)'(F) + w_t.D_t\overline{\pi}\Phi^2(F)dt\right] =$$

$$= \mathbf{E}\left[\int_0^T w_t.D_tF\left(\mathbf{E}\left[\overline{\pi}\int_0^T |D_tF|^2 dt \,|\, \sigma(F)\right](\Phi^2)'(F) + \right.\right.$$

$$+ \ \mathbf{E}\left[\int_0^T D_tF.D_t\overline{\pi}dt \,|\, \sigma(F)\right]\Phi^2(F)\Big) \ \left[\mathbf{E}\left[\int_0^T |D_tF|^2 dt\,|\,\sigma(F)\right]\right\}^{-1}\Bigg]$$

Hence, we have

$$\begin{cases} D_tF\,\overline{\pi}(\Phi^2)'(F) + D_t\overline{\pi}\Phi^2(F) = D_tF\,\dfrac{\mathbf{E}\left[\left(\int_0^T |D_tF|^2 dt\right)\overline{\pi}\,|\,\sigma(F)\right]}{\mathbf{E}\left[\int_0^T |D_tF|^2 dt\,|\,\sigma(F)\right]}(\Phi^2)'(F)+ \\ +D_tF\,\dfrac{\mathbf{E}\left[\int_0^T D_tF.D_t\overline{\pi}dt\,|\,\sigma(F)\right]}{\mathbf{E}\left[\int_0^T |D_tF|^2 dt\,|\,\sigma(F)\right]}\ \Phi^2(F) \end{cases} \tag{2.10}$$

Obviously, (2.10) holds if $\overline{\pi}$ is $\sigma(F)$ measurable i.e. is a function of F and thus coincides with π_0 in view of (2.4). Conversely, (2.10) implies at least when Φ^2 does not vanish

$$D_t\overline{\pi} \ = \ A\ D_tF \quad \text{in} \quad \Omega\times[0,T]$$

for some random variable A independent of $t\in[0,T]$. This condition should imply under appropriate nondegeneracy and smoothness conditions that $\overline{\pi}$ is a function of F exactly as the fact that the gradients of two scalar functions on $\mathbb{R}^N$ are parallel at each point “implies” that these functions are functionally dependent. We are able to answer positively this question in a few rather specific cases (see Appendix B below) and we mention this open problem not only for its own intrinsic interest but also because it is highly suggestive of a potential use of the Malliavin derivatives in order to measure the “nonlinear correlations” of random variables.

2.3 Financial applications

Now, we go back and comment some of the formula established in Part I [1] for the greeks. Our goal is to show that, in most of them, the derived weights are optimal in the sense of Proposition 2.1 above. We shall always assume that the contingent asset S_t is one-dimensional and satisfies (2.2) with either $\sigma(S)=\sigma S$, $b(S)=bS$, $\sigma>0$, $b\in\mathbb{R}$, the usual Black-Scholes model (BS), or $\sigma(S)=\sigma>0$, $b(S)=b\in\mathbb{R}$, the Brownian model that we only consider for a pedagogical purpose. We begin with an option of the form $V=E[\phi(S_T)]$, for some $T>0$ and some measurable function Φ with at most polynomial growth at infinity. This option could be, for instance, a call ($\Phi(S)=(S-K)_+$), or a digital ($\Phi(S)=1_{(S>K)}$), where $K>0$ is fixed. Then, as shown in Part I [1], we have

(delta) $\quad \dfrac{\partial V}{\partial S} = \mathbf{E}\left[\Phi(S_t)\pi\right],\ \pi=\dfrac{W_T}{\sigma ST}(BS \text{ model},\ \pi=\dfrac{W_T}{\sigma T}$ (Brownian model) ,

(vega) $\quad \dfrac{\partial V}{\partial \sigma} = \mathbf{E}\left[\Phi(S_t)\pi\right],\ \pi=\dfrac{W_T^2}{\sigma ST}-W_T-\dfrac{1}{\sigma}$ (in both cases)

Since obviously π is a function of S_T, these weights are optimal.

The case of the Asian option is more interesting. We thus consider now $V = \mathbf{E}\left[\Phi\left(\int_0^T S_t dt\right)\right]$. This corresponds to the case $d = 2$, where the second component of $X_t = (S_t, I + \int_0^t S_s ds)$ satisfies $dX_t^2 = S_t dt$. Then, the delta is given by

$$\frac{\partial V}{\partial S} = \mathbf{E}\left[\Phi(\int_0^T S_t dt)\pi\right] \tag{2.11}$$

where $\pi = \frac{6}{T^2}\int_0^T W_t dt - 2\frac{W_T}{T}$ in the Brownian case and

$$\pi = \frac{2}{S\sigma^2}\frac{S_T - S}{\int_0^T S_t dt} + \frac{1}{S}\left(1 - \frac{2b}{\sigma^2}\right) \quad \text{for the BS Model .} \tag{2.12}$$

Once more, these weights are functions of $\left(S_t, \int_0^T S_t dt\right)$ and thus are optimal.

We emphasize the fact that, in all the previous examples but the last one, those explicit weights may be deduced directly from the explicitly known laws of the processes. However, in the last case (BS model, Asian option), even though the law of $\int_0^T S_t dt$ is not known explicitly, Malliavin calculus yields the optimal weight!

Finally, if we consider a general model (2.2) with σ and b of class C^1 (say) with σ' and b' bounded and if we assume that σ is uniformly invertible (to simplify), then, as shown in [1], we have

$$(\nabla_x \mathbf{E}[\Phi(X_T)])^* = \mathbf{E}\left[\Phi(X_T)\left(\frac{1}{T}\int_0^T (\sigma^{-1}(X_t)Y_t)^* dW_t\right)\right]$$

where we denote by M^* the transpose of a matrix M and Y_t is the linearized flow given by

$$DY_t = (\sigma'(X_t).Y_t).dW_t \ , \quad Y_0 = I$$

In general, this representation, deduced from Malliavin calculus, does not provide an optimal weight since it is not, in general, a function of X_T.

Remark 2.1 Let us observe that, in cases when we do not know how to obtain an explicit optimal weight, one may use a combination of Monte-Carlo simulations and of an iterative procedure based upon the variance functional $\mathscr{V}$ and its gradient (which can be computed along the lines developed in the Euler-Lagrange section). Starting with an explicit π_1, we compute by a Monte-Carlo simulation an approximation for $A = \mathbf{E}\left[\Phi'(F)G\right]$ and then use $\mathscr{V}(\pi)$ to “improve” the weight by a gradient descent method, obtaining in this way a new weight π_2 and so forth ...

Remark 2.2 It is worth observing that the existence of a weight π is insured by the smoothness of the laws of the underlying processes, while Malliavin calculus actually yields such weights in case when the Malliavin derivatives of these processes exist and “do not vanish too much ”. At least formally, this second notion of smoothness is more restrictive and there are examples of financial products that do not satisfy it. More precisely, taking $d = 1$, a solution X_t of (2.2) with σ, b smooth, σ positive, we may consider a barrier option (we could as well consider lookbacks . . .) like for instance

$$V = \mathbf{E}\left[\Phi(X_T)\ 1_{(\tau \geq T)}\right]$$

where τ is the first hitting time of the region $\{x \geq B\}$ where $B > 0$ is fixed. The approach developed in Part I [1] breaks down in this case since $D_t\tau$ does not exist. However, all the laws appearing in that expectation are smooth. We shall see in a future publication how this serious difficulty may be circumvented and how one can extend the results and methods of Part I [1] (and of this paper) to the cases of general barrier options and of options of maxima (or minima) such as lookback options.

3 Identities and variance reduction

In this section, our goal is not to provide a general theory but to emphasize an observation that was sketched in Part I [1] and which is crucial for the practical implementation of Monte-Carlo simulations based upon our use of Malliavin calculus.

Let us begin by explaining how one can localize the integration by parts provided by Malliavin calculus around the singularity points, in the context of greeks. Also, in order to simplify the presentation, we shall simply consider the case when the underlying asset is a pure Brownian motion i.e. $S_t = S_0 + \sigma W_t,\ t \geq 0$. Then, we consider the delta of a digital, or equivalently the gamma of a call, namely

$$A = \frac{\partial}{\partial S}\mathbf{E}\left(\mathbf{1}_{(S_t>K)}\right)\Big|_{S=S_0} = \frac{\partial^2}{\partial S^2}\mathbf{E}\left((S_t-K)^+\right)\Big|_{S=S_0} \tag{3.13}$$

Then, as can be checked by a direct inspection or through a (very!) particular case of Malliavin integration by parts, we certainly have

$$A = \mathbf{E}\left(\mathbf{1}_{(S_t>K)}\ \frac{W_t}{\sigma t}\right) \tag{3.14}$$

or

$$A = \mathbf{E}\left((S_t-K)_+\ \frac{(W_t^2-t)}{\sigma^2 t^2}\right) \tag{3.15}$$

Next, if we wish to compute A by a Monte-Carlo simulation, it is important to compare the variance of the two random variables whose mean yield A namely $\mathbf{1}_{(S_t>K)}\frac{W_t}{\sigma t}$ and $(S_t-K)_+\left(\frac{W_t^2-t}{\sigma^2 t^2}\right)$. In other words, we wish to compare

$$\mathbf{E}\,(\mathbf{1}_{(S_t>K)}\ W_t^2) \quad \text{and} \quad \mathbf{E}\,\left(\mathbf{1}_{(S_t>K)}(S_t-K)^2\frac{(W_t^2-t)^2}{\sigma^2t^2}\right)$$

These expressions are difficult to compare but a limit on their behavior is provided by computing their behavior as t goes to $+\infty$ or as t goes to 0_+. This is indeed straightforward by the scaling properties of the Brownian motion and we find

$$\begin{cases} \frac{1}{t}\mathbf{E}\left(1_{(S_t>K)}W_t^2\right) \longrightarrow \frac{1}{2} \text{ as } \ t\to+\infty\ , \\ \frac{1}{t}\mathbf{E}\left[1_{(S_t>K)}(S_t-K)^2\left(\frac{(W_t^2-t)^2}{\sigma^2t^2}\right)\right] \longrightarrow \mathbf{E}\left(1_{(W_1>0)}\left(W_1^6-2W_1^4+W_1^2\right)\right) = \\ \qquad = \ 5 \text{ as } \ t\to+\infty \end{cases}$$

$$\begin{cases} \frac{1}{t}\mathbf{E}\left(1_{(S_t>K)}W_t^2\right) \longrightarrow 1 \text{ if } S_0>K,\ \frac{1}{2} \text{ if } \ S_0=K,\ \text{as } \ t\to 0_+, \\ \mathbf{E}\left((S_t-K)_+\frac{(W_t^2-t)^2}{\sigma^2t^2}\right) \longrightarrow 2\frac{(S_0-K)^2}{\sigma^2} \text{ if } \ S_0>K \text{ as } \ t\to 0_+, \\ \frac{1}{t}\mathbf{E}\left((S_t-K)_+\frac{(W_t^2-t)^2}{\sigma^2t^2}\right) \longrightarrow 5 \text{ if } \ S_0=K \text{ as } \ t\to 0_+ \end{cases}$$

and all these expectations are exponentially small if $S_0 < K$.
In other words, the variance of the second variable is asymptotically 10 times larger than the variance of the first one. The explanation is that each integration by parts creates a weight of order $\frac{W_t}{t}$ which builds up the variance. Let us record for future purposes that the variance or more precisely the second moment of the random variables entering the expectations in (3.14), (3.15) respectively behaves like $\frac{c}{\sigma^2 t}$ as t goes to $+\infty$ where $c = \frac{1}{2}, 5$ respectively as t goes to $+\infty$ or as t goes to 0_+ and $S_0 = K$.

These considerations together with realistic Monte-Carlo simulations (see [1] for more details) led us to propose in [1] a localized integration by parts. For instance, in the case of the digital, we write $\chi_0(x) = 1_{(x>0)}$ and we introduce a "smooth" function (say Lipschiz) χ such that we have for some $a > 0$,

$$\chi(x) = 0 \text{ if } \ x \le -a\ , \quad \chi(x) = 1 \text{ if } \ x \ge a\ . \tag{3.16}$$

For instance, take $a = 1$, $\chi(x) = \frac{1}{2}(x+1)$ if $-1 \le x \le 1$.

Then, using again the integration by parts we find

$$A = \frac{\partial}{\partial S}\mathbf{E}\left(\chi_0(S_t-K)\right)\Big|_{S=S_0} = \mathbf{E}\left[\chi'(S_t-K)\right] + \frac{\partial}{\partial S}\mathbf{E}\left((\chi_0-\chi)(S_t-K)\right)\Big|_{S=S_0}$$

$$A \ = \ \mathbf{E}\left[\chi'(S_t-K) - \phi(S_t-K)\frac{W_t}{\sigma t}\right] \tag{3.17}$$

where $\phi = \chi - \chi_0$. Notice that, because of (3.16), ϕ and χ' are supported in $[-a, +a]$. We may now estimate the second moment of $\chi'(S_t-K)-\phi(S_t-K)\frac{W_t}{\sigma t}$ as t goes to 0_+ or as t goes to $+\infty$. We begin with the analysis of the case when t goes to 0_+. And we immediately observe that this second moment is exponentially small provided we choose χ in such a way that χ' and ϕ vanish near $S_0 - K$. This is obviously possible if $S_0 \neq K$ (choose $a < |S_0 - K|$ for instance ...). If $S_0 = K$, we choose χ such that $\chi(0) = 1/2$ and we obtain easily

$$\mathbf{E}\left(\left\{\chi'(S_t - K) - \phi(S_t - K)\frac{W_t}{\sigma t}\right\}^2\right) = \frac{1}{4\sigma^2 t} + o\left(\frac{1}{t}\right)$$

thus dividing by two the size of the second moment.

We now turn to the case when $t \to +\infty$. We then pick $a = 1$, χ_1 satisfying (3.16) with $a = 1$ such that $\chi_1(0) = \frac{1}{2}$ and we choose $a = t$, $\chi = \chi_1\left(\frac{\cdot}{t}\right)$. Then, one checks easily that

$$\mathbf{E}\left[\left\{\chi'(S_t - K) - \phi(S_t - K)\frac{W_t}{\sigma_t}\right\}^2\right] t \to \frac{1}{4\sigma^2} \text{ as } t \text{ goes to } +\infty .$$

Let us emphasize the fact that the preceding choices for a and χ are simple illustrations of the variance reduction induced by the localization approach i.e. by the use of the integration by parts around the singularity (here, $S_0 - K$). In terms of variance, the above choices lead to an asymptotic variance reduction by a factor of the order 4. It is precisely given by $\frac{4(\pi-1)}{2\pi-4}$.

One can perform a similar analysis for the gamma of a call. We now assume that χ is at least of class $C^{1,1}$ and satisfies in addition to (3.16)

$$\int_{-a}^{+a} \chi dx = a . \tag{3.18}$$

We then introduce $\phi = \displaystyle\int_{-\infty}^{x} \chi - \chi_0 \; dy$ and we write

$$A = \frac{\partial^2}{\partial S^2}\mathbf{E}\left((S_t - K)_+\right)\Big|_{S=S_0} = \mathbf{E}\left(\chi''(S_t - K)\right) - \frac{\partial^2}{\partial S^2}\mathbf{E}\left(\phi(S_t - k)\right)\Big|_{S=S_0}$$

hence

$$A = \mathbf{E}\left(\chi''(S_t - K) - \phi(S_t - k)\frac{W_t^2 - t}{\sigma^2 t^2}\right) . \tag{3.19}$$

Then, we immediately observe that $B = \mathbf{E}\left[\left(\chi''(S_t - K) - \phi(S_t - K)\frac{W_t^2 - t}{\sigma^2 t^2}\right)^2\right]$ is exponentially small as t goes to 0_+ if χ'' and ϕ vanish near $S_0 - K$ and this is certainly possible if $S_0 \neq K$ since ϕ and χ'' are supported in $[-a, +a]$ in view of (3.16) and (3.18). Next, if $S_0 = K$, we again choose χ such that $\chi(0) = \frac{1}{2}$ and we find that tB goes to $\frac{S}{2\sigma^2}$ as t goes to 0_+. And one can make a similar analysis in the case when t goes to $+\infty$.

The first general observation that we wish to make concerns the usefulness of this localization method. We have seen how the variance can be significantly reduced and this theoretical (by a simple asymptotic analysis) evidence is confirmed by the numerical examples in [1] - that show in fact an even more dramatic speed up of the Monte-Carlo simulations.

We also wish to point out that, although the preceding elementary analysis was performed in the trivial case of a Brownian model for the underlying asset, this localization trick, in conjunction with Malliavin integration by parts, should always be used for practical Monte-Carlo simulations and for general models. In particular, in the case of Black-Scholes model that is when the underlying asset is a lognormal process, the above considerations immediately adapt and we have for instance if S_+ solves

$$dS_t = \sigma S_t dW_t + bS_t dt \ , \ \text{for } t \geq 0 \ ,$$

the following expressions

$$\frac{\partial}{\partial S}\mathbf{E}\left(1_{(S_t>K)}\right)\bigg|_{S=S_0} = \mathbf{E}\left(\chi'(S_t-K)-(\chi-\chi_0)(S_t-K)\frac{W_t}{\sigma t S_0}\right) \tag{3.20}$$

$$\begin{aligned}&\frac{\partial^2}{\partial S^2}\mathbf{E}\left((S_t-K)_+\right)\bigg|_{S=S_0}\\ &= \mathbf{E}\left(\chi''(S_t-K)-\phi(S_t-K)\frac{1}{S_0^2\sigma t}\left(\frac{W_t^2}{\sigma t}-W_t-\frac{1}{\sigma}\right)\right)\end{aligned} \tag{3.21}$$

Let us also mention in passing that the previous asymptotic evaluation also adapts trivially to that case.

We conclude this section by a general remark on Brownian expectations.

Remark 3.3 We wish to point out here that if we are interested in computing by a Monte-Carlo simulation an expression of the form $\mathbf{E}\left(\varphi(W_t)\right)$ for some function φ, then it may be of interest, in order to reduce the variance of the random variable whose expectation is the desired quantity, to rewrite $\mathbf{E}\left(\varphi(W_t)\right)$ using the Malliavin integration by parts at least on pieces of $\varphi(W_t)$. More precisely, we may write formally (without bothering about the mathematical conditions that may be needed) if $\varphi(x) = \varphi_1(x) + x\varphi_2(x) + \varphi_3(x), \varphi_3 = \phi_3'$

$$\begin{aligned}\mathbf{E}\left(\varphi(W_t)\right) &= \mathbf{E}\left[\varphi_1(W_t)\right] + \mathbf{E}\left[\varphi_2(W_t)W_t\right] + \mathbf{E}\left[\varphi_3(W_t)\right]\\ &= \mathbf{E}\left[\varphi_1(W_t)\right] + t\mathbf{E}\left[\varphi_2'(W_t)\right] + \mathbf{E}\left[\phi_3(W_t)\frac{W_t}{t}\right]\\ &= \mathbf{E}\left[\varphi_1(W_t) + t\varphi_2'(W_t) + \phi_3(W_t)\frac{W_t}{t}\right].\end{aligned}$$

And we may want or need to reiterate such manipulations on each piece.

4 Conditional expectations

In this section, we present a new application of Malliavin calculus to the representation of conditional expectations. We shall derive representation formulas which are explicit enough to be computed by Monte-Carlo simulations in a straightforward way. In Sect. 4 below, we first show why such formula should exist and in order to do so we follow the same line of arguments as in Sect. 2. Next, we show in Sect. 4.2 how it is possible to obtain the desired formula from Malliavin calculus and we work out some examples. Sections 4.3 and 4.4 are then devoted to some particular applications that are relevant for numerical Finance: in Sect. 4.3, we explain how one can use a single set of sample paths emanating from a fixed position to compute the price of a European option at all points and all intermediate times between the initial time and the maturity. We present some numerical examples. Finally, in Sect. 4.4, we develop conditional expectations involving one component of a multidimensional "state process" and apply this approach to the computation of quantities that are relevant for the calibration of stochastic volatility models.

4.1 Conditional expectations via the densities

Exactly as we did in Sect. 2, we consider general conditional expectations of the form

$$\mathbf{E}\left[\phi(F)|G=0\right] \tag{4.22}$$

where F is a m-dimensional $\mathscr{F}_T$-measurable random variable, G is a scalar $\mathscr{F}_T$-measurable random variable and ϕ is a Borel-measurable function on $\mathbb{R}^N$ with appropriate growth at infinity, say at most polynomial if we assume that all moments of F (and G) are bounded. Of course, the above expression does not make sense for arbitrary random variables (F,G) and we need to assume, in general, that the law of (F,G) admits some smoothness (which is made precise below). At least formally, the above expression may be written as $\mathbf{E}\left[\phi(F)\delta_0(G)\right]$ $\mathbf{E}\left[\delta_0(G)\right]^{-1}$ and we wish to explain now why there exists in general a weight π such that we have for all ϕ

$$\mathbf{E}\left[\phi(F)|G=0\right] = \frac{\mathbf{E}[\phi(F)H(G)\pi]}{\mathbf{E}[H(G)\pi]} \tag{4.23}$$

where $H=1_{(x>0)}+c$, $c\in\mathbb{R}$ is arbitrary. Of course, (4.23) is equivalent to

$$\mathbf{E}\left[\phi(F)\delta_0(G)\right] = \mathbf{E}\left[\phi(F)H(G)\pi\right], \text{ for all } \phi\,. \tag{4.24}$$

This is indeed the case as soon as the joint law of (F,G) admits a density $p(x,y)$ ($x\in\mathbb{R}^N$, $y\in\mathbb{R}$) such that its log is C^1 (or even Lipschitz) with a differential which grows at most in a polynomial way at infinity (for instance). Indeed, we may then write

$$\begin{aligned}\mathbf{E}\left[\phi(F)\delta_0(G)\right] &= \iint \phi(x)\delta_0(y)p(x,y)dxdy\\ &= -\iint \phi(x)\ H(y)\frac{\partial p}{\partial y}(x,y)dxdy\\ &= \iint \phi(x)H(y)\ q_0(x,y)\ p(x,y)\ dxdy\end{aligned}$$

hence (4.24) where $\pi=\pi_0=q_0(F,G)$ and $q_0(x,y)=\frac{1}{p}\ \frac{\partial p}{\partial y}=-\frac{\partial}{\partial y}\log p$.

Remark 4.1 Of course, if we assume that G takes its values in $\mathbb{R}^p$ $(p \geq 1)$ instead of being scalar, a similar manipulation may be made replacing the Heaviside function H by any function H such that

$$\delta_0 \;=\; P(D_y)H$$

where $P(D_y)$ denotes, as usual, any differential operator with constant coefficients determined by a polynomial P. Then, $q_0(s,y) = \frac{1}{p}P(-D_y)p$. For instance, we may choose $P(D_y) = \frac{\partial}{\partial y_1} \dots \frac{\partial}{\partial y_p}$ and $H = 1_{(y_i>0, \forall 1 \leq i \leq p)}$.

Remark 4.2 Let a be a C^1 (or even Lipschitz) function on $\mathbb{R}^N \times \mathbb{R}$ such that $a(x,0) = 1$ for all $x \in \mathbb{R}^m$. Then, $\delta_0 = a\delta_0$ and we see that we may replace q_0 in the preceding derivatives by $q = -\frac{1}{p}\frac{\partial(ap)}{\partial y} = aq_0 - \frac{\partial a}{\partial y}$ $(= -a\frac{\partial}{\partial y}\log(ap)$ if a is positive).

The last elementary observation we want to make in this section consists in using the smoothness of ϕ (provided we assume it) and write

$$\mathbf{E}\left(\phi(F)\delta_0(G)\right) = \mathbf{E}\left[\phi(F)h(G)\pi_1 - \phi'(F)h(G)\pi_2\right] \tag{4.25}$$

for some weights π_1, π_2. This is indeed possible at least when we assume, as we did above, some smoothness of $\log p$. Indeed, we look for $\pi_1 = q(F,G)$ and $\pi_2 = r(F,G)$ and we write

$$\begin{aligned} & \mathbf{E}\left[\phi(F)H(G)\pi_1 - \phi'(F)H(G)\pi_2\right] \;= \\ = & \iint \left\{\phi(x)H(y)q(s,y) - \phi'(x)H(y)r(x,y)\right\} p(x,y)dxdy \\ = & \iint \phi(x)H(y)\left\{q + \frac{1}{p}\frac{\partial}{\partial x}(rp)\right\} p \; dxdy \\ = & \mathbf{E}\left[\phi(F)\delta_0(G)\right] \end{aligned}$$

provided we request that q and r satisfy

$$q + \frac{1}{p}\;\frac{\partial}{\partial x}(rp) \;=\; q_0 = -\frac{\partial}{\partial y}(\log p)\;. \tag{4.26}$$

Remark 4.3 Of course, in the above formula, π_0 (resp. π_1 and π_2) may be replaced by more general weights. For instance, in (4.24), π_0 may be replaced by any π such that : $E\left[\pi|\sigma(F,G)\right] = \pi_0$. And, exactly as in Sect. 2, π_0 is among all such weights an optimal one in terms of variance.

Remark 4.4 In view of the multiplicity of choices (of a or of r ...), a natural question is to decide whether there exists a particular choice that minimizes the variance or equivalently the second moment of the random variable $\phi(F)H(G)\pi$ (or $\phi(F)H(g)\pi_1 - \phi'(F)H(g)\pi_2 \dots$). Then, if we consider for instance the question of the selection of an optimal a and we choose $H(y) = \frac{1}{2}\text{sign}(y)$, we are then led to the following minimization problem : minimize, over all $a(x,y)$ with appropriate growth at infinity and such that $a(x,0) = 1$ for all $x \in \mathbb{R}^m$, the expression

$$\mathbf{E}\left(\Phi(F)^2\left[a(F,G)q_0(F,G) - \frac{\partial a}{\partial y}(F,G)\right]^2\right) =$$

$$= \int_{\mathbb{R}^M} \Phi(x)^2 dx \int_{\mathbb{R}} dy \left| a(x,y)q_0(x,y) - \frac{\partial a}{\partial y}(y)\right|^2 p(x,y) \ .$$

We then claim that an (in fact the, at least when Φ does not vanish ...) optimal a is given by

$$\overline{a}(x,y) = b(x)\left(\int_y^{+\infty} p(x,z)dz\right) p(x,y)^{-1}$$

where $b(x) = p(x,0)\left(\int_0^{+\infty} p(x,z)dz\right)^{-1}$. Indeed, one can check easily that $\overline{a}$ solves the associated Euler-Lagrange equation

$$\left(\frac{\partial}{\partial y} + q_0\right) p \left(\frac{\partial}{\partial y} - q_0\right) a = 0 \ , \text{ on } \mathbb{R}^m \times \mathbb{R}$$

and the above minimization problem is convex. Notice that the weight $\overline{\pi} = \overline{q}(F,G)$ associated to $\overline{a}$ is trivial namely $\overline{\pi} = b(F)$ and we have of course

$$\begin{aligned}\mathbf{E}\left[\Phi(F)\delta_0(G)\right] &= \int \Phi(x)p(x,0)dx = \int \Phi(x)1_{(y>0)}b(x)p(x,y)dy \\ &= \mathbf{E}\left[\Phi(F)H(G)\overline{\pi}\right] \ .\end{aligned}$$

However, such a choice heavily depends upon the use of δ_0 and H while (4.24) and ((4.25) with the choices made above hold in fact whenever we replace δ_0 by ψ and H by Ψ with $\Psi' = \psi$.

4.2 Conditional expectations in Malliavin calculus

The elementary considerations developed in the previous section are of course theoretical since, in general, the joint law $p(x,y)$ of (F,G) is not known. Once more, they only explain why weights like π or (π_1,π_2) should exist. We shall now show how explicit or at least computable weights can be obtained using Malliavin calculus. Exactly as in Sect. 2, we then need to assume that F and G are smooth in Malliavin sense i.e. (for instance) $F, G \in L^2(\Omega)$, D_tF, $D_tG \in L^2(\Omega \times [0,T])$ and that D_tG is non degenerate. We thus assume that there exists a smooth process $u_t \in \mathbb{H}^1$ such that

$$\mathbf{E}\left[\int_0^T D_tGu_t \ dt | \sigma(F,G)\right] = 1 \ . \tag{4.27}$$

Whenever this is possible, one may simply choose $u_t = \frac{1}{D_tG}$.

Then, we have, for any C^1 function ϕ (or locally Lipschitz ...) such that ϕ grows, say, at most linearly at infinity and for any Heaviside-like function $H(y) = 1_{(y>0)} + c$ with $c \in \mathbb{R}$, the following representation formula

Theorem 4.1 *Under the above assumptions, we have*

$$\mathbf{E}\left[\phi(F)|G=0\right] = \frac{\mathbf{E}\left[\phi(F)H(G)\delta(u) - \phi'(F).H(G)\int_0^T D_tF.u_t dt\right]}{\mathbf{E}\left[H(G)\delta(u)\right]} \tag{4.28}$$

Corollary 4.1 *If there exists u_t satisfying (4.27) such that*

$$\mathbf{E}\left[\int_0^T D_tFu_t\ dt|\sigma(F,G)\right] = 0\ , \tag{4.29}$$

then we have

$$\mathbf{E}\left[\phi(F)|G=0\right] = \frac{\mathbf{E}\left[\phi(F)H(G)\delta(u)\right]}{\mathbf{E}\left[H(G)\delta(u)\right]} \tag{4.30}$$

and this formula is valid by density for any Borel measurable function ϕ with, say, at most linear growth at infinity.

Proof of Theorem 4.1 The formula (4.28) is at least formally completely obvious. We write $\mathbf{E}\left[\phi(F)|G=0\right] = \mathbf{E}\left[\phi(F)\delta_0(G)\right] \Big/ \mathbf{E}\left[\delta_0(G)\right]$. And we have

$$\begin{aligned}
\mathbf{E}\left[\phi(F)\delta_0(G)\right] &= \mathbf{E}\left[\int_0^T D_t\left\{\phi(F)H(G)\right\}u_t dt\right] \\
&\quad -\mathbf{E}\left[(\phi'(F)H(G).\int_0^T D_tfu_t dt\right] \\
&= \mathbf{E}\left[\phi(F)H(G)\delta(u) - \phi'(F)H(G).\int_0^T D_tFu_t dt\right]\ .
\end{aligned}$$

In order to justify this formal computation, we go back to a definition of such a conditional density and we write

$$\mathbf{E}\left[\phi(F)|G=0\right] = \lim_{\epsilon\to 0_+} \frac{\mathbf{E}\left[\phi(F)1_{(-\epsilon+\epsilon)}(G)\right]}{\mathbf{E}\left[1_{(-\epsilon,+\epsilon)}(G)\right]}$$

Then, we may write

$$\begin{aligned}
\mathbf{E}\left[\phi(F)1_{(-\epsilon,+\epsilon)}(G)\right] &= \mathbf{E}\left[\int_0^T D_t\left\{\phi(F)H_\epsilon(G)\right\}u_t dt\right] \\
&\quad -\mathbf{E}\left[\phi'(F)H_\epsilon(G).\int_0^T D_tFu_t dt\right] \\
&= \mathbf{E}\left[\phi(F)H_\epsilon(G)\delta(u) - \phi'(F)H_\epsilon(G).\int_0^T D_tFu_t dt\right]\ ,
\end{aligned}$$

where $H_\epsilon(y) = c$ if $y \leq -\epsilon$, $= y + \epsilon + c$ if $-\epsilon \leq y \leq \epsilon$, $= 2\epsilon + c$ if $y \geq \epsilon$. We then conclude easily upon letting ϵ go to 0_+ since $\frac{1}{\epsilon} H_\epsilon(G)$ converges a.s. to $2H(G)$ (recall that $P(G = 0) = 0$). □

We shall give in the next sections some specific examples of u_t which allow to compute or determine in a computable way the Skorohod integral $\delta(u)$.

Remark 4.5 The existence of u_t satisfying (4.29) in addition to (4.27) means that $D_t F$ and $D_t G$ are not "parallel" or in other words (see Appendix B) that F and G are not correlated. Indeed, if F is a (smooth) nontrivial (nonconstant) function of G, such a process u_t clearly does not exist since $D_t F$ is proportional to $D_t G$.

Remark 4.6 The above approach can be extended to situations where we condition expectations with respect to vector-valued random variables G. One simply iterates the above argument choosing for instance "integrating" processes u_t^i such that

$$\mathbf{E}\left[\int_0^T D_t G_j u_t^i dt | \sigma(F, G)\right] = \delta_{ij} \quad \text{for } 1 \leq i, j \leq m$$

requires, of course, that $(G_1, \ldots, G_m)$ are not correlated i.e. that $(D_t G_1, \ldots, D_t G_m)$ is a free system or, for instance, that the Malliavin covariance matrix $(D_t G_i D_t G_j)$ is "invertible" enough.

Remark 4.7 With the above choices of u_t, we also have for "any" function ψ

$$\mathbf{E}\left[\phi(F)\psi(G)\right] = \mathbf{E}\left[\phi(F)\Psi(G)\delta(u) - \phi'(F).\Psi(G)\int_0^T D_t F u_t dt\right]$$

where $\Psi' = \psi$, and thus if u_t satisfies (4.29)

$$\mathbf{E}\left[\phi(F)\psi(G)\right] = \mathbf{E}\left[\phi(F)\Psi(G)\delta(u)\right].$$

This equality shows in particular that, if we set $\pi = \delta(u)$

$$\mathbf{E}\left[\pi | \sigma(F, G)\right] = \pi_0$$

where π_0 has been determined in the preceding Sect. (4.1).

And, exactly as in Sect. 2, one may check that $\pi_0 = \delta(u^0)$ where $u_t^0 = \mathbf{E}\left[D_t q_0(F, G) | \mathcal{F}_t\right]$ and that u^0, in general, satisfies (4.27) and (4.29). Indeed, we have for all ϕ and ψ

$$\mathbf{E}\left[\phi(F)\Psi(G)\delta(u^0)\right] = \mathbf{E}\left[\int_0^T D_t\left\{\phi(F)\Psi(G)\right\} u_t^0 dt\right]$$
$$= \mathbf{E}\left[\int_0^T D_t F u_t^0 dt).\phi'(F)\Psi(G) + \phi(F)\psi(G)(\int_0^T D_t G u_t^0 dt)\right]$$

Hence, we have for all ϕ and ψ

$$\mathbf{E}\left\{\phi(F)\psi(G)\left[1 - \mathbf{E}\left[\int_0^T D_t G u_t^0 dt | \sigma(F, G)\right]\right]\right\}$$
$$= \mathbf{E}\left\{\phi'(F).\Psi(G)\mathbf{E}\left[\int_0^T D_t F u_t^0 dt | \sigma(F, G)\right]\right\}$$

or $\frac{\partial}{\partial y}(q, p) = \text{div}_x(rp)$ on $\mathbb{R}^m \times \mathbb{R}$, if $1 - \mathbf{E}\left[\int_0^T D_t G u_t^0 dt | \sigma(F,G)\right] = q(F,G)$ and
$\mathbf{E}\left[\int_0^T D_t F u_t^0 dt | \sigma(F,G)\right] = r(F,G)$. Therefore, if qp and rp vanish at infinity, then $q \equiv r \equiv 0$.

Finally, exactly as in Sect. 2, π_0 is among all weights the one that minimizes the variance of $\phi(F)\Psi(G)\pi$.

Remark 4.8 As mentioned before, our weights allow to replace (δ_0, H) by any couple (ψ, Ψ) such that $\Psi' = \psi$. One may use this fact in order to localize the integration by parts as we explained in Sect. 3. Indeed, we may consider a smooth bump function ψ_0 supported in $[-a, +a]$ for some $a > 0$ such that $\int_{-a}^{+a} \psi dy = 1$: take for instance $\psi_0 = \frac{1}{2a} 1_{[-a,+a]}$. Then, we set $\psi = \delta_0 - \psi_0$ and choose Ψ such that $\Psi' = \psi$ and Ψ has compact support in $[-a, +a]$ $\left(\Psi(y) = -\int_{-\infty}^{y} \psi_0(z)dz + H(y)\right)$. Then, we may write

$$\mathbf{E}[\phi(F)\delta_0(G)] = \mathbf{E}[\phi(F)\psi_0(G)] + \mathbf{E}[\phi(F)\psi(G)]$$

$$= \mathbf{E}[\phi(F)[\psi_0(G) + \Psi(G)\pi]].$$

Remark 4.9 At this stage, we wish to mention in passing that everything we did in Part I [1] or that we are doing here can be adapted to situations involving, in addition to the Brownian motion W, a jump process like a Poisson process provided it is assumed to be independent of W.

4.3 Application : Global pricing and hedging of a European option

In this section, we present the first application of the approach developed in the previous Sects. (4.1–4.2) to Numerical Finance. Roughly speaking, we show that the representations we obtained above allows to determine the price and greeks of a European option for all S (initial position) and for all $t_1 \in (0, T)$ using only the process that starts at $t_0 = 0$, S_0 fixed. In other words, one can determine the solution (and its differentials) of a diffusion parabolic equation at all points and at all positive times using only the trajectories of the process that start initially at a fixed position. In some sense, this completely modifies the general understanding of the use of Monte-Carlo methods for the computation of the solution of diffusion parabolic equation or of the price (and hedges) of a European option. Indeed, one would usually think that Monte-Carlo methods compute the solution at a single point (in space-time) while PDE numerical schemes compute the solution everywhere (at least on a predetermined grid). Our approach shows that this is not the case since one can compute (in a completely parallel way) the solution everywhere by a Monte-Carlo method involving a single generation of trajectories emanating from a given point!

We first explain how one can derive closed form formulas and, at the end of this section, we present a numerical example for the "global" computation of a European option by a Monte-Carlo Method.

We thus introduce some notation. The underlying process (contingent asset) S_t is assumed to solve

$$dS_t = \sigma(S_t)dW_t + b(S_t)dt \ , \quad S_t = S \quad \text{at} \quad t = 0 \ , \tag{4.31}$$

and we consider for all $t \geq 0$, $S \in \mathbb{R}$ (or $S > 0$)

$$V(S,t) \ = \ \mathbf{E}_S \left[\phi(S_t)e^{-rt}\right]$$

where ϕ is the pay-off, a given measurable function (with at most polynomial growth at infinity), $r \geq 0$ is the interest rate (assumed to be constant to simplify the presentation), $\sigma(S)$ is the volatility. And we assume that σ and b are smooth functions of S with σ' and b' bounded. We also assume that σ is non degenerate ($\sigma^2 > 0$). Finally, we denote by E_S the usual expectation involving S_t which starts at $t = 0$ at the point S (i.e. solves (4.31)) in order to emphasize the nature of the trajectories we use.

Specific examples include the Brownian case i.e. σ =constant, $b = 0$ or the lognormal case i.e. $\sigma(S) = \sigma S$ with $\sigma > 0$ (and a trivial degeneracy at $S = 0$), $b(S) = bS$ and $b \in \mathbb{R}$.

We also wish to point out that we only consider here the one-dimensional situation in order to simplify the presentation even though everything we do may be adapted to multidimensional settings.

We next recall that V solves uniquely the following linear partial differential equation

$$\frac{\partial V}{\partial t} - \frac{1}{2}\sigma^2(S)\frac{\partial^2 V}{\partial S^2} - b(S)\frac{\partial V}{\partial S} + rV = 0 \quad \text{for} \quad S \in \mathbb{R}, \ t > 0$$

$$V\Big|_{t=0} \ = \ \phi(S) \quad \text{for} \quad S \in \mathbb{R} \ .$$

Finally, we recall that the Markov property immediately yields the following formula for all $0 < t < T$ and for all S_0, $S \in \mathbb{R}$

$$\mathbf{E}_{S_0}\left[\phi(S_T)|S_t = S\right] \ = \ V(S, T-t)\ e^{r(T-t)} \ .$$

We may thus fix $S_0 \in \mathbb{R}$ ($S_0 > 0$ in the lognormal case) and $T > 0$ (the maturity of the option) and we want to compute or determine in a "computable" way $\mathbf{E}_{S_0}\left[\phi(S_T)|S_t = S\right]$ together with various differential like, for instance, $\frac{\partial}{\partial S}\mathbf{E}_{S_0}\left[\phi(S_T)|S_t = S\right]$ i.e. the delta.

We then begin to apply the method introduced in the preceding section (4.2) in the case when ϕ is smooth. We then choose $u_s = \frac{1}{t}\frac{\xi_s}{\sigma(S_s)\xi_t}\ 1_{(0,t)}(s)$ where ξ_t denotes the tangent flow defined by

$$d\xi_s = \sigma'(S_s)\xi_s dW_s + b'(S_s)\xi_s ds, \ \xi_0 = 1 \ .$$

Recalling that $D_s S_t = \sigma(S_s)\frac{\xi_t}{\xi_s}\ 1_{(0,t)}(s)$, se see that (4.27) holds. Therefore, we find

$$\mathbf{E}_{S_0}\left[\phi(S_T)|S_t = S\right] =$$

$$= \mathbf{E}_{S_0}\left\{\phi(S_T)H(S_t - S)\delta(u) - \phi'(S_T)H(S_t - S). \int_0^T D_s S_T\ u_s ds\right\}$$

$$= \mathbf{E}_{S_0}\left[\phi(S_T)H(S_t - S)\delta(u) - \phi'(S_T)H(S_t - S)\frac{\xi_T}{\xi_t}\right] .$$

and we have

$$\delta(u) = \frac{1}{t\xi_t}\int_0^t \frac{\xi_s}{\sigma(S_s)}dW_S + \frac{1}{t}\int_0^t \frac{\xi_s}{\sigma(S_s)}\ \frac{D_s\xi_t}{\xi_t^2}\ ds .$$

There remains to determine $D_s\xi_t$, a computation which we isolate in the following

Lemma 4.1 *Let $\zeta_t = \frac{\partial^2 S_t}{\partial S^2}$ i.e. the solution of the following SDE*

$$d\zeta_t = \left\{\sigma'(S_t)\zeta_t + \sigma''(S_t)\xi_t^2\right\}dW_t + \left\{b'(S_t)\zeta_t + b''(S_t)\xi_t^2\right\}dt,\ \zeta_0 = 0 . \quad (4.32)$$

Then, $D_s\xi_t = \left\{\frac{\sigma(S_s)}{\xi_s}\zeta_t + \sigma'(S_s)\xi_t - \sigma(S_s)\ \frac{\zeta_s}{\xi_s^2}\ \xi_t\right\}\ 1_{(t\geq s)}$.□.

Admitting temporarily this fact, we finally obtain the

Theorem 4.2 *We have for all* $S,\ t \in (0, T]$

$$\mathbf{E}_{S_0}(\phi(S_T)|S_t = S)$$
$$= \mathbf{E}_{S_0}\left[\phi(S_t)H(S_t - S)\pi - \phi'(S_T)H(S_t - S)\frac{\xi_T}{\xi_t}\right]\left(\mathbf{E}_{S_0}[H(S_t - S)\pi]\right)^{-1} ,$$

where π is given by

$$\pi = \frac{1}{t\xi_t}\left(\int_0^t \frac{\xi_s}{\sigma(S_s)}dW_s + t\frac{\zeta_t}{\xi_t} + \int_0^t \sigma'(S_s)\frac{\xi_s}{\sigma(S_s)} - \frac{\zeta_s}{\xi_s}ds\right) . \quad (4.33)$$

Proof of Lemma 4.1 Applying the usual Malliavin calculus rules, we find

$$\begin{cases} d(D_s\xi_t) = \sigma'(S_t)(D_s\xi_t)dW_t + \sigma''(S_t)(D_s S_t)\xi_t dW_t + \\ + b'(S_t)(D_s\xi_t)dt + b''(S_T)(D_s S_t)\xi_t\ dt,\ \text{for } t \geq s , \end{cases}$$

with $D_s\xi_t\Big|_{t=s} = \sigma'(S_s)\xi_s$. Hence, we have

$$d(D_s\xi_t) = \sigma'(S_T)(D_s\xi_t)dW_t + \sigma''(S_t)\frac{\sigma(S_s)}{\xi_s}\xi_t^2 dW_t +$$

$$+b'(S_t)(D_s\xi_t)dt + b''(S_t)\ \frac{\sigma(S_s)}{\xi_s}\xi_t^2 dt, \quad \text{for } t \geq s$$

Therefore, $\varphi_t = D_s\xi_t - \frac{\sigma(S_s)}{\xi_s}\zeta_t$ solves

$$d\varphi_t = \sigma'(S_t)\varphi_t dW_t + b'(S_t)\varphi_t dt, \quad \text{for} \quad t \geq s$$

and thus

$$\varphi_t = \varphi_s \frac{\xi_t}{\xi_s} = \left(\sigma'(S_s)\xi_s - \frac{\sigma(S_s)}{\xi_s}\zeta_s\right)\frac{\xi_t}{\xi_s}$$

$$= \sigma'(S_s)\ \xi_t \ - \ \frac{\sigma(S_s)\zeta_s\xi_t}{\xi_s^2}\ .$$

□

If we do not want to use the smoothness of ϕ, we need to apply Corollary 4.1 and thus we choose

$$u_s \ = \ \frac{\xi_s}{\sigma(S_s)\xi_t}\ \left\{\frac{1}{t}\ 1_{(0,t)}(s) - \frac{1}{T-t}\ 1_{(t,T}(s)\right\}\ ,$$

so that (4.27) and (4.29) hold with this choice.
Then, we set

$$\left\{\begin{array}{l} \pi = \delta(u) = \frac{1}{\xi_t}\left(\frac{1}{t}\int_0^t \frac{\xi_s}{\sigma(S_s)}dW_s - \frac{1}{T-t}\int_t^T \frac{\xi_s}{\sigma(S_s)}dW_s\right) + \\ \qquad + \frac{\zeta_t}{\xi_t^2} + \frac{1}{t\xi_t}\int_0^t \frac{\sigma'(S_s)}{\sigma(S_s)}\xi_s ds - \frac{1}{t\xi_t}\int_0^t \frac{\zeta_s}{\xi_s}ds. \end{array}\right. \tag{4.34}$$

and we deduce from Corollary 4.1 the

Theorem 4.3 *We have for all* $S, b \in (0, T)$

$$\mathbf{E}_{S_0}\left(\phi(S_T)|S_t = S\right) \ = \ \frac{\mathbf{E}_{S_0}(\phi(S_T)H(S_t - S)\pi)}{\mathbf{E}_{S_0}(H(S_t - S)\pi)} \tag{4.35}$$

with π *given by (4.34).*

Remark 4.10 If we wish to compute the delta namely $\frac{\partial}{\partial S}\mathbf{E}_{S_0}(\phi(S_T)|S_t = S)$, we see from the previous formula that we only need to compute

$$\mathbf{E}_{S_0}\left[\phi(S_T)\delta_S(S_t)\pi\right] = \mathbf{E}_{S_0}\left[\int_0^T D_s\left\{\phi(S_T)H(S_t - S)\right\}.u_s\pi ds\right]$$

$$= \ \mathbf{E}_{S_0}\left[\phi(S_T)H(S_t - S)\pi^2 - \phi(S_T)H(S_t - S)\int_0^T u_s D_s\pi ds\right]$$

and one can compute $D_s\pi$ from the explicit formula (4.34). It involves however tedious but straightforward computations that we do not wish to include here and we shall only mention them in the context of the examples that follow.

We next detail the preceding formula in the Brownian case, $(S_t = S_0 + \sigma W_t)$, and in the log normal case $\left(S_t = S_0 \exp\left\{\sigma W_t - \left(\frac{\sigma^2}{2} - b\right)t\right\}\right)$.

Example 4.1 The Brownian case.
In that case, $\xi_t \equiv 1$, $\zeta_t \equiv 0$ and we obtain

$$\mathbf{E}_{S_0}(\phi(S_T)|S_t = S) = \frac{\mathbf{E}_{S_0}\left[\phi(S_T)H(S_t - S)\frac{W_t}{t\sigma} - \phi'(S_T)H(S_t - S)\right]}{\mathbf{E}_{S_0}\left[H(S_t - S)\frac{W_t}{t\sigma}\right]}$$

$$= \frac{\mathbf{E}_{S_0}\left[\phi(S_T)W_t - t\sigma\phi'(S_T))H(S_t - S)\right]}{\mathbf{E}_{S_0}\left[H(S_t - S)W_t\right]}$$

as a consequence of Theorem 4.2, and by Theorem 4.2

$$\mathbf{E}_{S_0}(\phi(S_T)|S_t = S) = \frac{\mathbf{E}_{S_0}\left[\phi(S_T)H(S_t - S)\frac{1}{\sigma}(\frac{W_t}{t} - \frac{W_T - W_t}{T-t})\right]}{\mathbf{E}_{S_0}\left[(S_t - S)\frac{1}{\sigma}(\frac{W_t}{t} - \frac{W_T - W_t}{T-t})\right]}$$

$$= \frac{\mathbf{E}_{S_0}\left[\phi(S_T)H(S_t - S)(TW_t - tW_T)\right]}{\mathbf{E}_{S_0}\left[H(S_t - S)(TW_t - tW_T)\right]}.$$

Example 4.2 The lognormal case.
In that case, $\xi_t \equiv \frac{S_t}{S_0}$ and $\zeta_t \equiv 0$. We thus deduce from theorem 4.2

$$\mathbf{E}_{S_0}(\phi(S_T)|S_t = S)$$
$$= \frac{\mathbf{E}_{S_0}\left[\phi(S_T)H(S_t - S)(t\sigma S_t)^{-1}(W_t + t\sigma) - \phi'(S_T)H(S_t - S)\frac{S_T}{S_t}\right]}{\mathbf{E}_{S_0}\left[H(S_t - S)(t\sigma S_t)^{-1}(W_t + t\sigma)\right]}$$
$$= \frac{\mathbf{E}_{S_0}\left[\phi(S_T)H(S_t - S)S_t^{-1}(W_t + t\sigma) - t\sigma\phi'(S_T)H(S_t - S)S_T S_t^{-1}\right]}{\mathbf{E}_{S_0}\left[H(S_t - S)S_t^{-1}(W_t + t\sigma)\right]}.$$

And, similarly, we deduce from theorem 4.3

$$\mathbf{E}_{S_0}(\phi(S_T)|S_t = S) = \frac{\mathbf{E}_{S_0}\left[\phi(S_T)H(S_t - S)\left\{(\sigma S_t)^{-1}(\frac{W_t}{t} - \frac{W_T - W_t}{T-t}) + S_t^{-1}\right\}\right]}{\mathbf{E}_{S_0}\left[H(S_t - S)\left\{(\sigma S_t)^{-1}(\frac{W_t}{t} - \frac{W_T - W_t}{T-t}) + S_t^{-1}\right\}\right]}$$

$$= \frac{\mathbf{E}_{S_0}\left[\phi(S_T)H(S_t - S)S_t^{-1}\left(\frac{TW_t - tW_T}{T-t} + \sigma t\right)\right]}{\mathbf{E}_{S_0}\left[H(S_t - S)S_t^{-1}(\frac{TW_t - tW_T}{T-t} + \sigma t)\right]}.$$

We conclude this section by a brief presentation of a numerical illustration of the preceding formula. We performed simulations for the BS model and computed the conditional expectation by the previous formula on a predefined grid. We compared the results obtained by these simulations and our filtering formula with the exact price given by the Black-Scholes formula for values of the grid up to one standard deviation from the money.

More precisely, we have chosen a grid of size $N \times M$ where N, M stand respectively for the number of steps in time and space. Let $t_1 < \ldots < t_N$ and

$$X = \left\{x_{ij} = S_0 \exp\left(\left(r - \frac{\sigma^2}{2}\right)t_i + j\sigma\Delta x_i\right),\ \Delta x_i = \sqrt{t_i}\right\},$$

the grid on which we evaluate the conditional expectations. Let K be the number of trajectories (starting from S_0) we have simulated, we have computed the following terms and compared them to exact prices given by Black-Scholes formula.

$$\mathbf{E}_{S_0}\left(e^{-r(T-t)}(S_T-K)_+|S_t=x_{ij}\right)\simeq$$

$$\simeq e^{-r(T-t)}\frac{\frac{1}{K}\sum_{k=0}^{K}(S_T(k)-K)_+H(S_t(k)-x_{ij})S_t^{-1}(k)\left(\frac{TW_t(k)-tW_T(k)}{T-t}+\sigma t\right)}{\frac{1}{K}\sum_{k=0}^{K}H(S_t(k)-x_{ij})S_t^{-1}(k)(\frac{TW_t(k)-tW_T(k)}{T-t}+\sigma t)}$$

This is a crude estimator. There are many variations of this estimator which give improvements for this particular case. But we want to show the performance of the formula without any statistical optimization.

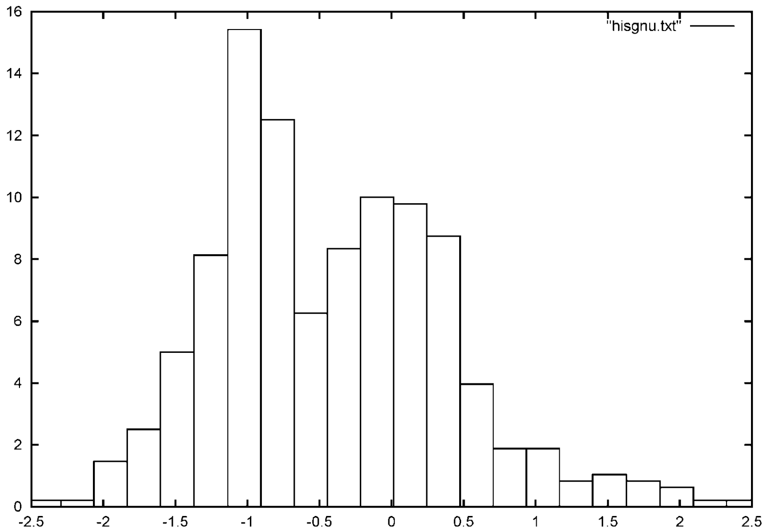

Fig. 1. Histogramme of relative error (%)

For the simulations, we used a very simple method of generation of low discrepancy trajectories, the so-called brownian bridge method using Sobol points at times $T/2, T$.

We present the results in the Figs. 1 and 2.

One sees that the filtering formula is quite accurate close to the initial time, and becomes less efficient when the time goes toward the maturity. The formula is also better around the money. It gives a larger relative error out of the money. This means that the number of trajectories is too small in these areas.

These are crude numerical results. We will show in a forthcoming paper how to improve these simulations in a way which will allow to solve the problem of American options with a satisfying accuracy.

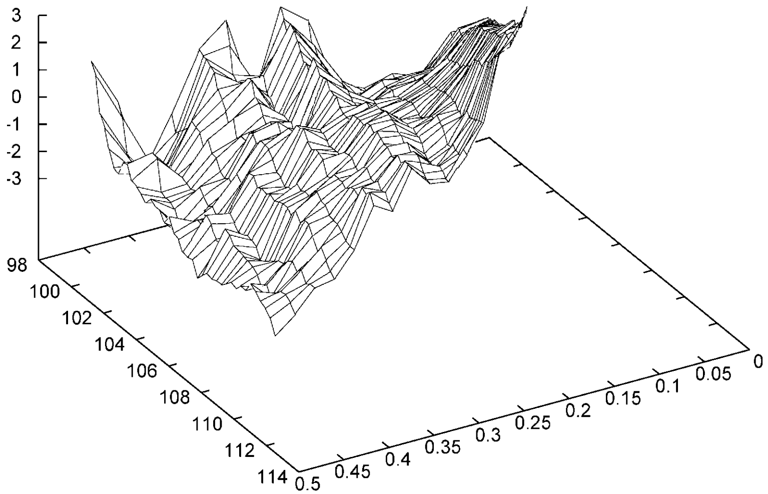

Fig. 2. Relative error of filtering formula for BS model with parameters $S_0 = 100$, rate $r = 0.1$, volatility $\sigma = 0.1$, maturity $T = 1$ year. We use some low discrepancy Monte Carlo generation and 10000 trajectories

4.4 Application to a stochastic volatility model

The application we shall develop here is a prototypical example of the conditioning of one component by another component of a multidimensional diffusion process. We present here a particular example which is significant in numerical Finance for the so-called calibration of stochastic volatility models. We thus consider a simple stochastic volatility model

$$\begin{cases} dS_t = \sigma_t \; S_t \; dW_t^1 + b \; S_t \; dt \;, \quad S_0 > 0 \\ d\sigma_t = \alpha \; \sigma_t \; dW_t^2 + \beta \sigma_t dt \;, \quad \sigma_0 > 0 \end{cases} \tag{4.36}$$

where $\alpha, \beta, b \in \mathbb{R}$ and (W_t^1, W_t^2) are independent Brownian motions. The independence assumption is not essential for the analysis which follows but it certainly simplifies the explicit computations that we perform.

In order to calibrate such models on market data, it is useful to be able to compute

$$\mathbf{E} \; \left[\sigma_T^2 \mid S_T = S\right] \tag{4.37}$$

where $T > 0$ is fixed. And we refer to forthcoming papers for more details on the financial application. This is indeed rather easy thanks to our general approach. First, we notice that $D_t \sigma_T \equiv 0$ since W^1 and W^2 are independent and that (4.29) holds thus for any u. Next, we observe that we have for $t \geq s$

$$\begin{cases} d(D_s S_t) = \sigma_t (D_s S_t) dW_t^1 + b(D_s S_t) dt, \quad \text{for } t \geq s \\ D_s S_t \Big|_{t=s} = \sigma_s S_s . \end{cases}$$

Hence, $D_s S_t = \sigma_s S_t$. Therefore, we may choose $u_t = \frac{1}{T\sigma_t S_t}$ and both conditions (4.27) and (4.29) hold. Finally, we compute

$$\begin{aligned}\pi = \delta(u) &= \frac{1}{TS_T}\int_0^T \frac{1}{\sigma_t}\, dW_t + \frac{1}{T}\int_0^T \frac{1}{\sigma_t}\,\frac{D_t S_T}{S_T^2}dt \\ &= \frac{1}{TS_T}\int_0^T \frac{dW_t}{\sigma_t} + \frac{1}{S_T}.\end{aligned}$$

We have thus shown the following formula

$$\mathbf{E}\left[\sigma_T^2 | S_T = S\right] = \frac{\mathbf{E}\left[\sigma_T^2 H(S_T - S)\; S_T^{-1}\;(T + \int_0^T \frac{dW_t}{\sigma_t})\right]}{\mathbf{E}\left[H(S_T - S)\; S_T^{-1}\;(T + \int_0^T \frac{dW_t}{\sigma_t})\right]} \tag{4.38}$$

We conclude this section by a numerical illustration of the preceding formula.

We performed Monte Carlo simulations in the same spirit as in the previous example. We allowed the maturity T to vary between 0 and 0.2 year and the conditional spot S at time T to vary around the forward of one standard deviation for $\sigma = \sigma_0$ constant. We computed the formula (4.38) by the standard Monte Carlo estimator. We present the results in the Fig. 3.

We see that the conditional volatility is (as it should be) higher when conditioning to values of S_T out of the money.

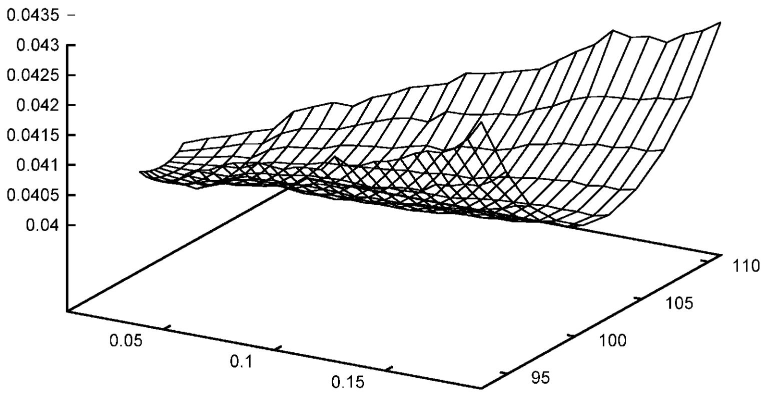

Fig. 3. Values of $\mathbf{E}\left[\sigma_T^2 \mid S_T = S\right]$ for parameters $S_0 = 100$, rate $r = 0.1$, initial volatility $\sigma_0 = 0.2$, $\alpha = 0.1$ and $\beta = 0.2$

5 Anticipating Girsanov transforms

5.1 Introduction

In this section, we develop one more application of Malliavin calculus namely the possibility of using anticipating Girsanov transforms. Before going into the mathematical details, let us first explain the heuristic idea behind our developments.

Let us consider a European option whose pay-off vanishes when the contingent asset at maturity does not belong to a certain interval. Then, all Monte-Carlo trajectories that do not hit that region at maturity do not contribute to the pay-off and are thus wasted. Therefore, it is natural to try to replace the underlying process by another one with a greater probability of hitting the "target zone" at maturity. This however means transforming the process or changing variables and we then need to be able to compute the Jacobian of this transform. Let us also observe that this transform is a priori anticipative since we should use the value of the process at the final time i.e. the maturity.

If the transform were not anticipative, one could and may apply the standard theory of Girsanov transforms that precisely does the job. Examples of applications of this type to Numerical Finance are presented in E. Fournié et al. [2] for out of the money contingent claims (the transform is then determined by the theory of large deviations) and E. Fournié et al. [3] for stochastic volatility models (where the transform is determined by a "constant volatility submodel"). Both examples involve situations where the optimal transform (optimal since it induces a zero variance) associated with importance sampling, namely the ratio (delta / price) at each point and each time, which is of course unknown in general, is approximated using a convenient asymptotic theory.

However, here, as explained above, we need to use the theory of anticipating Girsanov transforms and we refer the interested reader to D. Nualart [6] and S. Kusuoka [4] for a more complete presentation of the theory which involves the "usual" Girsanov exponential together with an extra term which is really a weighted Jacobian. Our main concern in the next Sect. 5.2 will be to have explicit expressions for a class of transforms that covers our needs in the context of Numerical Finance. Then, in Sect. 5.3, we briefly present some applications.

5.2 Anticipating Girsanov theorems for a terminal transform

Let $T > 0$ be fixed. We begin with a simple transformation defined by

$$d\widehat{W_t} \;=\; dW_t + \varphi(\widehat{W_T})dt \quad \text{for} \quad 0 \le t \le T \;, \tag{5.39}$$

in other words, we consider the transformation on paths : $\omega \mapsto \widehat{\omega}$ where $\widehat{\omega}(t) - \varphi(\widehat{\omega}(T))t = \omega(t)$. Here and below, φ is a Lipschitz function such that the mapping $\omega \mapsto \widehat{\omega}$ is well-defined and bijective. This is obviously the case if $T \sup_{z\in\mathbb{R}} \varphi'(z) < 1$. And, in order to simplify the presentation, we shall always assume that this condition holds. Finally, we only consider the one-dimensional case eventhough everything we do adapt immediately to higher dimensions.

Then, we claim that there exists a measure $\widehat{P}$ under which $\widehat{\omega}$ is a Brownian motion and for all $F \in \mathscr{F}_T$

$$\mathbf{E}[F] = \widehat{\mathbf{E}}\left[\widehat{F}\,.\exp\left(\widehat{W_t}\varphi(\widehat{W_t}) - \frac{T}{2}\varphi^2(\widehat{W_T})\right).\left(1 - T\varphi'(\widehat{W_T})\right)\right] \tag{5.40}$$

where $\widehat{F}(\widehat{W}) = F(W)$, and $\widehat{\mathbf{E}}$ denotes expectation under $\widehat{P}$. Let us begin with the trivial case where $F = F(W_T)$.

Then, we have setting $z = y - T\varphi(y)$

$$\begin{aligned}
\mathbf{E}[F] &= \int F(z)\frac{e^{-\frac{|z|^2}{2T}}}{\sqrt{2\pi T}}dz = \int \widehat{F}(y)\frac{e^{\frac{-|y-T\varphi(y)|^2}{2T}}}{\sqrt{2\pi T}}(1-T\varphi')(y)dy \\
&= \int \widehat{F}(y)\frac{e^{-|y|^2/2T}}{\sqrt{2\pi T}}e^{y\varphi(y)-\frac{T}{2}\varphi^2(y)}(1-T\varphi'(y))dy \\
&= \widehat{\mathbf{E}}\left[\widehat{F}(\widehat{W_T}).\exp\left(\widehat{W_T}\varphi(\widehat{W_T})-\frac{T}{2}\varphi^2(\widehat{W_T})\right).(1-T\varphi'(\widehat{W_T}))\right].
\end{aligned}$$

More generally, if $N \geq 1$, $F = F(\Delta_1, \ldots, \Delta_N)$, $\Delta_i = W_{ih} - W_{(i-1)h}$, $h = \frac{T}{N}$, then we write

$$\begin{aligned}
\mathbf{E}(F) &= \int F(z_1, \ldots, z_N)\ \frac{e^{-\frac{|z|^2}{2h}}}{(2\pi h)^{N/2}}dz \\
&= \int \widehat{F}(y_1, \ldots, y_N)(2\pi h)^{-N/2}\exp\left(-\frac{1}{2h}|y-\varphi\left(\sum_{j=1}^N y_j\right)h|^2\right)Ddy
\end{aligned}$$

where D =det $(\delta_{ij} - \varphi'(\sum_{j=1}^N y_j)h\ e_{ij})$, $e_{ij} = 1$ for all $1 \leq i, j \leq N$.

Hence, we deduce

$$\mathbf{E}(F) = \int \widehat{F}(y)\frac{e^{\frac{-|y|^2}{2h}}}{(2\pi h)^{N/2}}exp\left(\left(\sum_{i=1}^N y_i\right)\varphi\left(\sum_{j=1}^N y_j\right)-\frac{T}{2}\varphi^2\left(\sum_{j=1}^N y_j\right)\right)Ddy.$$

There remains to compute D which is a consequence of the “classical” matrix lemma.

Lemma 5.1 *Let $D = \det(\delta_{ij} - a \otimes b)$, where $a, b \in \mathbb{R}^N$. Then, $D = 1 - (a, b)$.*

Admitting temporarily this lemma, we conclude the proof of (46) since

$$D = 1 - N\ h\ \varphi'\left(\sum_{j=1}^N y_j\right) = 1 - T\varphi'\left(\sum_{j=1}^N y_j\right).$$

Proof of Lemma 5.1 Many proofs are possible. A nice one consists in writing $f(b) = 1-\det(\delta_{ij} - a \otimes b)$ where a is fixed. Obviously, f is smooth and by direct inspection we check that $f(0) = 0$, $f'(0) = a$, $f''(0) = 0$. Next, since $(\delta_{ij} - a \otimes b)$ $(\delta_{ij} - a \otimes c) = \delta_{ij} - a \otimes b - a \otimes [c(1-(a,b))]$ we deduce that we have

$$f(b) + f(c) - f(b)f(c) = f(b + c(1-(a,b))).$$

Differentiating twice this relationship with respect to c and setting $c = 0$, we obtain

$$\begin{cases} a(1-f(b)) & = f'(b)(1-(a,b)) \\ 0 & = f''(b)(1-(a,b))^2 \end{cases}$$

hence $f''(b) = 0$ on $\mathbb{R}^N$ and f is linear since $f(0) = 0$.
Therefore, we deduce from the first equation that $f(b) = (a, b)$. □

In fact the above argument extends to a more general transformation defined by $(\omega \mapsto \widehat{\omega})$ where

$$\widehat{\omega}(t) - t\varphi(\widehat{\Delta}_1, \ldots, \widehat{\Delta}_n) = \omega(t) \quad \text{for all} \quad t \in [0, T]$$

and $\widehat{\Delta}_i = \widehat{\omega}(t_i) - \widehat{\omega}(t_{i-1})$, where $t_0 = 0 < t_1 < t_2 < \ldots < t_N = T$ is an arbitrary partition of $[0, T]$ and φ is a smooth function on $\mathbb{R}^N$ such that $\omega \mapsto \widehat{\omega}$ is bijective. And we obtain in place of (5.40)

$$\begin{cases} \mathbf{E}[F] = \widehat{\mathbf{E}}\Big[\widehat{F}.\exp\Big(\widehat{W_T}\varphi(\widehat{\Delta}_1, \ldots, \widehat{\Delta}_N) - \frac{T}{2}\varphi^2(\widehat{\Delta}_1, \ldots, \widehat{\Delta}_N)\Big). \\ \qquad .\Big(1 - \S_{i=1}^N(t_i - t_{i-1})\frac{\partial\varphi}{\partial x_i}(\widehat{\Delta}_i, \ldots, \widehat{\Delta}_n)\Big)\Big] \end{cases}$$

or in other words

$$\begin{cases} \mathbf{E}[F] = \widehat{\mathbf{E}}\Big[\widehat{F}.\exp\Big(\widehat{W_T}\varphi(\widehat{\Delta}_1, \ldots, \widehat{\Delta}_N) + \\ \quad - \frac{T}{2}\varphi^2(\widehat{\Delta}_1, \ldots, \widehat{\Delta}_N)\Big)\Big(1 - \int_0^T \widehat{D}_t\varphi dt\Big)\Big] \end{cases} \tag{5.41}$$

where we denote by $\widehat{D}$ the Malliavin derivative in the transformed Wiener space $(\Omega, \widehat{P}, \widehat{W})$. This with a little extra work extends to a general transformation of the form $(\omega \mapsto \widehat{\omega})$ where

$$\widehat{\omega}(t) - t\varphi(\widehat{\omega}) = \omega(t) \quad \text{for all } t \in [0, T]$$

assuming that φ is smooth, $(\omega \mapsto \widehat{\omega})$ is bijective and $(1 - \int_0^T \widehat{D}_t\varphi dt) \neq 0$ a.s.

If we compare the above considerations with the literature we are aware of (see [6], [5] and the references therein), we have changed notations and more importantly we have isolated a special class of transformation for which the Carleman-Fredholm determinant is computable (here, it is, up to various irrelevant changes of signs, $(1 - \int_0^T \widehat{D}_t\varphi dt)\exp(\widehat{W_T}\varphi - \delta(\varphi))$ where, as usual, δ denotes the Skorohod integral of φ).

Remark 5.1 We have deliberately defined the paths transformations in an implicit way in order to simplify the expressions of some jacobians. If we insist on a more traditional formulation namely

$$d\widehat{W_t} = dW_t + \varphi(W_T)dt$$

or more generally

$$d\widehat{W_t} = dW_t + \varphi dt$$

where φ is a smooth function of paths (restricted to $[0, T]$) then we may translate the above results in the following way

$$\mathbf{E}[F] = \mathbf{E}\left[\widehat{F}\exp\left(-\varphi W_T - \frac{T}{2}\varphi^2\right)\left(1 + \int_0^T D_t\varphi dt\right)\right] \tag{5.42}$$

where $\widehat{F} = F(\widehat{W})$.

Finally, the multidimensional analogue of the preceding formula consists in replacing φW_T by $\varphi.W_T$, φ^2 by $|\varphi|^2$ and $1 + \int_0^T D_t\varphi$ by $\det\left(I + \int_0^T D_t\varphi dt\right)$.

5.3 Applications

The first application we present concerns European options involving a pay-off which vanishes if the contingent asset at maturity S_T vanishes on the complement of an interval $[K_1, K_2]$ where $0 < K_1 < K_2 < +\infty$. More precisely, we consider $E\,[\phi(S_T)]$ where ϕ is Borel measurable, bounded and supported in $[K_1, K_2]$. For instance, we may take

$$\phi(S) \;=\; (S - K_1)_+ \; 1_{(S<K_2)} \tag{5.43}$$

i.e. a call with a knock-out, or

$$\phi(S) \;=\; 1_{(K_1<S<K_2)} \tag{5.44}$$

i.e. a digital with a knock-out. And we consider, in order to simplify the presentation, the simple situation where $S_t = S_0 + \sigma W_t + bt$, $\sigma > 0$, $b \in \mathbb{R}$ (Brownian case) or where $S_t = S_0 \exp\,(\sigma W_t + (b - \frac{\sigma^2}{2})t)$, $\sigma > 0$, $b \in \mathbb{R}$ (lognormal case). We begin with the Brownian case and we note first that $\phi(S_T) = \psi(W_T)$ where $\psi(x) = \phi(S_0 + bT + \sigma x)$ is now supported in $[L_1, L_2]$ where $L_i = \frac{K_i - S_0 - bT}{\sigma}$ $(i = 1, 2)$. We then choose $L = \frac{L_1 + L_2}{2\theta}$, $\varphi(x) = -\theta R$ if $x \le -R$, $= \theta x$ if $x \in [-R, +R]$, $= \theta R$ if $x \ge R$ where $\theta \in (0, 1)$, $R > 0$. We shall use the following transformation

$$d\widehat{W_t} \;=\; dW_t - \frac{1}{T}\varphi(W_T - L)\; dt \;,$$

so that $\widehat{W_T} - L \;=\; W_T - L - \varphi(W_T - L)$.

Then, applying the results of the previous section , we have

$$\begin{cases} \mathbf{E}(\phi(S_T)) = \mathbf{E}(\psi(W_T)) \;= \\ = \mathbf{E}\left[\psi(\widehat{W_T})\exp\left(-\frac{1}{T}\varphi(W_T)W_T - \frac{1}{2T}\varphi^2(W_T)\right).(1 - \varphi'(W_T))\right]. \end{cases}$$

In order to estimate what we gained with this transformation, we observe that $\phi(S_T)$ vanishes if $S_T \notin [K_1, K_2]$ and thus the probability of hitting that interval at time T is given by

$$P(W_T \in [L_1, L_2]) \;=\; \int_{L_1}^{L_2} \frac{e^{-|y|^2/2T}}{\sqrt{2\pi T}}dy, \tag{5.45}$$

while $\psi(\widehat{W_T})$ vanishes if $\widehat{W_T} \notin [L_1, L_2]$ and thus the probability of reaching that interval at time T is given by

$$P\left(W_T - \varphi(W_T - L) \in [L_1, L_2]\right) = P\left(L + R \leq W_T \leq L_2 + \theta R\right) +$$

$$+P(L_1 - \theta R \leq W_T \leq L - R) + P\left(W_T \in \left[\frac{L_1}{1-\theta} - \frac{\theta L}{1-\theta}, \frac{L_2}{1-\theta} - \frac{\theta L}{1-\theta}\right]\right).$$

In particular, if $R \geq \frac{1}{1-\theta} \max(L - L_1, L_2 - L)$, this probability reduces to

$$P\left(W_T \in \left[-\frac{L_2 - L_1}{1(1-\theta)}, \frac{L_2 - L_1}{2(1-\theta)}\right]\right)$$

which is obviously greater than (5.45) and in fact goes to 1 as θ goes to 1_-.

In the lognormal case, the analysis is the same defining now ψ, L_1, L_2, by

$$\psi(x) = \phi\left(S_0 \exp(\sigma x + (b - \frac{\sigma^2}{2}T)\right),$$

$$L_i = \frac{1}{\sigma}\left(\log(\frac{K_i}{S_0}) + (\frac{\sigma^2}{2} - b)T\right), \quad \text{for } i = 1, 2.$$

Appendix A
A PDE interpretation for the representation of the delta

Let us consider an expectation of the form

$$\mathbf{E}\left[\phi(X_T)\right] = V(x, T) \tag{A.1}$$

where ϕ is a Borel measurable function on $\mathbb{R}$ with at most polynomial growth at infinity, $T > 0$ and X_t is a non degenerate process given by

$$dX_t = \sigma(X_t)dW_t + b(X_t)dt, \quad X_0 = x \tag{A.2}$$

where σ, b are smooth, $\sigma^2 > 0$ on $\mathbb{R}$, σ', b' are bounded on $\mathbb{R}$.
As is well known, V is the unique solution of

$$\frac{\partial V}{\partial t} - LV = 0 \quad \text{on } \mathbb{R} \times (0, +\infty), \quad V\Big|_{t=0} = \phi \quad \text{on } \mathbb{R} \tag{A.3}$$

where $L = \frac{1}{2}\sigma^2 \frac{\partial^2}{\partial x^2} + b\frac{\partial}{\partial x}$.

Let us immediately emphasize the fact that we consider here only a one-dimensional setting although everything we discuss below adapts trivially to higher dimensions.

Our goal here is to explain the partial differential equation interpretation of various representation formula for the delta i.e. for $\frac{\partial V}{\partial x}(x, T)$. The first one is the following classical formula obtained by differentiating (A.1) with respect to x.

This, of course, requires ϕ to have some smoothness like, for instance, ϕ locally Lipschitz and ϕ' grows at in a polynomial way at infinity. We then have

$$\frac{\partial V}{\partial x}(x,t) \; = \; \mathbf{E}\left[\phi'(X_t)\xi_t\right] \tag{A.4}$$

where ξ_t solves

$$d\xi_t \; = \; \sigma'(X_t)\xi_t dW_t + b'(X_t)\xi_t dt, \quad \xi_0 = 1 \tag{A.5}$$

The PDE interpretation of this formula is clear. We introduce a diffusion process (X_t, Y_t) where Y_t solves

$$dY_t \; = \; \sigma'(X_t)Y_t dW_t + b'(X_t)Y_t dt, \quad Y_0 = y \tag{A.6}$$

so that $Y_t = y\xi_t$. And, obviously $\frac{\partial V}{\partial x}(x,t)y = W(x,y,t)$ should be the solution of

$$\begin{cases} \frac{\partial W}{\partial t} - \frac{1}{2}\sigma^2\frac{\partial^2 W}{\partial x^2} - b\frac{\partial W}{\partial x} - \frac{1}{2}(\sigma')^2 y^2\frac{\partial^2 W}{\partial y^2} - \sigma\sigma' y\frac{\partial^2 W}{\partial x\partial y} + \\ \qquad\qquad - b'y\frac{\partial W}{\partial y} \; = \; 0 \quad \text{on} \quad \mathbb{R}^2\times(0,+\infty) \\ W\Big|_{t=0} = \phi'(x)y \quad \text{on} \quad \mathbb{R}^2. \end{cases} \tag{A.7}$$

This can be checked directly by differentiating (A.3) with respect to x and multiplying by y. We then find indeed

$$\begin{aligned} \frac{\partial W}{\partial t} &= y\frac{\partial}{\partial x}LV = LW + y\sigma\sigma'\frac{\partial^2 V}{\partial x^2} + yb'\frac{\partial V}{\partial x} \\ &= LW + y\sigma\sigma'\frac{\partial^2 W}{\partial x\partial y} + yb'\frac{\partial W}{\partial y} + \alpha\frac{\partial^2 W}{\partial y^2} \end{aligned}$$

for any function α since W is linear in y. In particular, we observe that W solves any parabolic equation of the form (A.7) where we replace $\frac{1}{2}(\sigma')^2 y^2$ by any function $\alpha \geq \frac{1}{2}(\sigma')^2 y^2$.

Therefore, we can even make such an equation uniformly parabolic by choosing $\alpha > \frac{1}{2}(\sigma')^2 y2$.

We now turn to representation formula deduced from Malliavin calculus (see [1] for more details). And we begin with the simplest case namely the Brownian case ($\sigma = 1,\ b = 0$). Then, (A.3) reduces to the standard heat equation

$$\frac{\partial V}{\partial t} - \frac{1}{2}\Delta V \; = \; 0 \; . \tag{A.8}$$

And, we have : $t\nabla V(x,t) = \mathbf{E}\left[\phi(x, W_t)W_t\right]$.

Once more, the analytical interpretation of that formula is clear : we solve

$$\frac{\partial U}{\partial t} - \frac{1}{2}\Delta V \; = \; 0, \quad U\Big|_{t=0} = x\phi \tag{A.9}$$

and the preceding formula simply means that we have

$$t\nabla V(x,t) \equiv U(x,t) - xV(x,t) \ . \tag{A.10}$$

Indeed, we have

$$\left(\frac{\partial}{\partial t} - \frac{1}{2}\Delta\right)(U - xV) = \nabla V = \left(\frac{\partial}{\partial t} - \frac{1}{2}\Delta\right)(t\nabla V)$$

while $U - xV\Big|_{t=0} \equiv t\nabla V\Big|_{t=0} \equiv 0$. In other words, the crucial fact here is the following property of the commutation of Δ and the multiplication operator by x namely

$$[\Delta, x] = \nabla \ .$$

For a general diffusion process X_t, the interpretation is more elaborate. Indeed, let us recall first the formula

$$t\frac{\partial V}{\partial t}(x,t) = \mathbf{E}\left[\phi(X_t)\int_0^t \frac{\xi_s}{\sigma(X_s)}dW_s\right] \ . \tag{A.11}$$

We then introduce a diffusion process (X_t, Y_t, Z_t) where Y_t is defined as above (i.e. $Y_t = y\xi_t$) and Z_t satisfies

$$dZ_t = \frac{Y_t}{\sigma(X_t)}dW_t, \quad Z_0 = z \tag{A.12}$$

Then, the above lationship (A.11) can be interpreted as

$$t\frac{\partial V}{\partial x}(x,t)y = F(x,y,z,t) - zV(x,t) \tag{A.13}$$

where $F(x,y,z,t)$ is given by

$$F(x,y,z,t) = \mathbf{E}\left[\phi(X_t)Z_t\right] \tag{A.14}$$

and thus F solves the following degenerate parabolic equation

$$\begin{cases} \frac{\partial F}{\partial t} = \frac{1}{2}\sigma^2(x)\frac{\partial F}{\partial x^2} + \frac{1}{2}\sigma'(x)^2y^2\frac{\partial^2 F}{\partial y^2} + \frac{1}{2}\frac{y^2}{\sigma^2(x)}\frac{\partial^2 F}{\partial z^2} + \\ \qquad + \sigma(x)\sigma'(x)y\frac{\partial^2 F}{\partial x\partial y} + y\frac{\partial^2 F}{\partial x\partial z} + \\ \qquad + \frac{\sigma'(x)}{\sigma(x)}y^2\frac{\partial^2 F}{\partial y\partial z} + b(x)\frac{\partial F}{\partial x} + b'(x)y\frac{\partial F}{\partial y} \\ F\Big|_{t=0} = \phi(x)z \ . \end{cases} \tag{A.15}$$

The relationship (A.13) can be checked directly by writing down the equation satisfied by $G(x,y,z,t) = t\frac{\partial V}{\partial x}(x,t)y + zV(x,t)$. Indeed, differentiating (A.3) with respect to x and multiplying by ty, we obtain easily

$$\begin{aligned}\frac{\partial G}{\partial t} &= \frac{\partial V}{\partial x}y + \frac{1}{2}\sigma^2\frac{\partial^2 G}{\partial x^2} + b\frac{\partial G}{\partial x} + ty\sigma\sigma'\frac{\partial^2 V}{\partial x^2} + tyb'\frac{\partial V}{\partial x}\\ &= \frac{1}{2}\sigma^2\frac{\partial^2 G}{\partial x^2} + b\frac{\partial G}{\partial x} + y\sigma\sigma'\frac{\partial^2 G}{\partial x\partial y} + yb'\frac{\partial G}{\partial y} + y\frac{\partial V}{\partial x}\\ &= \frac{1}{2}\sigma^2\frac{\partial^2 G}{\partial x^2} + b\frac{\partial G}{\partial x} + y\sigma\sigma'\frac{\partial^2 G}{\partial x\partial y} + yb'\frac{\partial G}{\partial y} + y\frac{\partial^2 G}{\partial x\partial z}\end{aligned}$$

and thus

$$\left\{\begin{aligned} &\tfrac{\partial G}{\partial t} = \tfrac{1}{2}\sigma^2\tfrac{\partial^2 G}{\partial x^2} + b\tfrac{\partial G}{\partial x} + y\sigma\sigma'\tfrac{\partial^2 G}{\partial x\partial y} + yb'\tfrac{\partial G}{\partial y} + \\ &\qquad +y\tfrac{\partial^2 G}{\partial x\partial z} + \alpha\tfrac{\partial^2 G}{\partial y^2} + \beta\tfrac{\partial^2 G}{\partial z^2} + \gamma\tfrac{\partial^2 G}{\partial y\partial z}\end{aligned}\right. \tag{A.16}$$

for any functions α, β, γ since $\frac{\partial^2 G}{\partial y^2} = \frac{\partial^2 G}{\partial z^2} = \frac{\partial^2 G}{\partial y\partial z} = 0$. In particular, we may choose $\alpha = \frac{1}{2}(\sigma')^2y^2$, $\beta = \frac{1}{2}\frac{y^2}{\sigma^2}$ and $\gamma = \frac{\sigma'}{\sigma}y^2$ in which case (A.16) reduces to (A.15) and we conclude that $F \equiv G$ since the equation (A.1) is degenerate parabolic if and only if $\alpha \geq \frac{1}{2}(\sigma')^2y^2$, $\beta \geq \frac{1}{2}\frac{y^2}{\sigma^2}$, $|\gamma - y^2\frac{\sigma'}{\sigma}|^2 \leq \left(\alpha - \frac{1}{2}(\sigma')^2y^2\right)\left(\beta - \frac{1}{2}\frac{y^2}{\sigma^2}\right)$.

Let us observe that (A.13) also holds where F solves (A.16) instead of (A.15) with $\alpha \geq \frac{1}{2}(\sigma')^2y^2$, $\beta \geq \frac{1}{2}\frac{y^2}{\sigma^2}$, $\left(\gamma - y^2\frac{\sigma'}{\sigma}\right)^2 \leq (\alpha - \frac{1}{2}(\sigma')^2y^2)(\beta - \frac{1}{2}\frac{y^2}{\sigma^2})$. In particular, (A.16) can be made uniformly parabolic provided we take $\alpha > \frac{1}{2}(\sigma')^2y^2$, $\beta > \frac{1}{2}\frac{y^2}{\sigma^2}$, $(\gamma - y^2\frac{\sigma'}{\sigma})^2 < (\alpha - \frac{1}{2}(\sigma')^2y^2)(\beta - \frac{1}{2}\frac{y^2}{\sigma^2})$.

We conclude this appendix by giving the PDE interpretation of the classical Girsanov transform for diffusion processes. We mention in passing this fact since we are not aware of any previous reference on that subject and the interpretation follows the same lines as above.

We thus introduce another diffusion process $\overline{X}_t$ solution of

$$d\overline{X}_t = \sigma(\overline{X}_t)DW_t + \overline{b}(\overline{X}_t)dt \tag{A.17}$$

and we set $M_t^1 = \exp\left\{\int_0^t \sigma^{-1}(b-\overline{b})(\overline{X}_s)dW_s - \frac{1}{2}\int_0^t |\sigma^{-1}(b-\overline{b})|^2(\overline{X}_t)ds\right\}$. Then we have

$$V(x,t) = \mathbf{E}\left[\phi(\overline{X}_t)M_t^1\right]. \tag{A.18}$$

Once more, we introduce the diffusion process $(\overline{X}_t, M_t)$ where M_t is defined by

$$dM_t = \sigma^{-1}(\overline{X}_t)(b-\overline{b})(X_t)\, M_t dW_t\ ,\ M_0 = m \tag{A.19}$$

so that $M_t = mM_t^1$. Then, $W(x,m,t) = \mathbf{E}\left[\phi(\overline{X}_t)M_t\right]$ solves

$$\left\{\begin{aligned} &\tfrac{\partial W}{\partial t} = \tfrac{1}{2}\sigma^2\tfrac{\partial^2 W}{\partial x^2} + \tfrac{1}{2}\sigma^{-2}(b-\overline{b})^2m^2\tfrac{\partial^2 W}{\partial m^2} + \\ &+ (b-\overline{b})m\tfrac{\partial W}{\partial x\partial m} + \overline{b}\tfrac{\partial W}{\partial x},\ W\Big|_{t=0} = \phi(x)m\end{aligned}\right. \tag{A.20}$$

while $\widetilde{W} = V(x,t)m$ solves

$$\begin{cases} \frac{\partial \widetilde{W}}{\partial t} = \frac{1}{2}\sigma^2 \frac{\partial^2 \widetilde{W}}{\partial x^2} + \alpha \frac{\partial^2 \widetilde{W}}{\partial m^2} + (b - \overline{b})m \frac{\partial^2 \widetilde{W}}{\partial x \partial m} + \overline{b} \frac{\partial \widetilde{W}}{\partial x} \ , \\ \widetilde{W}\Big|_{t=0} = \phi(x)m \end{cases} \tag{A.21}$$

for any function α and thus in particular for $\alpha = \frac{1}{2}\sigma^{-2}(b - \overline{b})^2 m^2$. Since (A.20) is degenerate parabolic, we conclude that $W = \widetilde{W}$.

Let us once more remark that $W \equiv Vm$ solves (A.21) for any α and that this equation is degenerate parabolic if and only if $\alpha \geq \frac{1}{2}\sigma^{-2}(b - \overline{b})m^2$ (and uniformly parabolic if $\alpha > \frac{1}{2}\sigma^{-2}(b - \overline{b})^2 m^2$) ...

Appendix B
Functional dependence and Malliavin derivative

As is well known, the covariance provides a tool to analyze the linear correlation between two random variables F and G. We briefly propose here a tool based upon Malliavin derivative to analyze the "nonlinear" correlation between two random variables F and G. We assume for instance that F and G are $\mathscr{F}_T$-measurable and are smooth, say $I\!H^1$ for instance, Without loss of generality, we may take $T = 1$ by a simple time rescaling. Then, if $F = \varphi(G)$ for some, say, Lipschitz φ, we have obviously $D_t F = \varphi'(G) D_t G$ a.s. and thus $D_t F$ and $D_t G$ are a.s. proportional as functions of t. This leads us to consider

$$C(F,G) = \operatorname{supess}_{\omega} \left\{ \left| \int_0^1 D_t F D_t G dt \right|^2 \Big/ \left(\int_0^1 |D_t F|^2 dt)(\int_0^1 |D_t G|^2 dt \right) \right\} \tag{B.1}$$

where we agree that this ratio is 1 if $D_t F$ or $D_t G$ vanishes identically on $[0,1]$. Obviously, $C(F,G) = 1$ if F is a function of G (and conversely), $C(F,G) = C(\varphi(F), \psi(G))$ for any φ, ψ which are Lipschitz (for instance) and thus, by extension and density, $C(F,G)$ is constant on $\sigma(F) \times \sigma(G)$.

Let us also mention that $C(F,G) = 0$ means that $D_t F$ and $D_t G$ are orthogonal in $L^2(0,1)$ a.s. and the relationship between that property and the independence of F and G has been studied in A.S. Üstünel and M. Zakai [9], [10].

In order to form an idea of what C does, it is worth looking at the case when $F = \varphi\left(W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_1 - W_{t_{m-1}}\right)$, $G = \psi(W_{t_1}, \ldots, W_1 - W_{t_m-1})$ where $m \geq 2$, $t_0 = 0 < t_1 < t_2 < \ldots < t_m = 1$ and φ, ψ are smooth. Then, $D_t \varphi = \sum_{i=1}^m \frac{\partial \varphi}{\partial x_i}(\ldots) 1_{(t_{i-1}, t_i)}(t)$ and $D_t \psi = \sum_{i=1}^m \frac{\partial \psi}{\partial x_i}(\ldots) 1_{(t_{i-1}, t_i)}(t)$, so that

$$C(F,G) = \sup_{x \in \mathbb{R}^m} \left\{ \left| \sum_{i=1}^m h_i \frac{\partial \varphi}{\partial x_i} \frac{\partial \psi}{\partial x_i} \right|^2 \Big/ \left(\sum_{i=1}^m h_i (\frac{\partial \varphi}{\partial x_i})^2 \right) \left(\sum_{i=1}^m h_i \frac{\partial \psi}{\partial x_i} \right)^2 \right\}$$

with $h_i = t_i - t_{i-1}$. In particular, $C(F,G) = 1$ means that $\nabla\varphi$ and $\nabla\psi$ are colinear at each point $x \in \mathbb{R}^n$, which is the usual criterion for a functional dependence between φ and ψ. If $\nabla\psi$ vanishes at some point, then we cannot deduce that φ is a function of ψ as is well-known : consider for instance $\varphi = \frac{1}{4}x_1^4 - \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2$, $\varphi = \psi$ on $\{\psi > 0\} \cup \{\psi < 0, x_1 > 0\}$, $\varphi = 0$ on $\{\psi < 0,\ x_1 < 0\}$.

For simple G, it is possible to conclude that F if G-measurable if $C(F,G) = 1$. First, if $D_tF \equiv 0$, and thus $C(F,G) = 1$ for any G by convention, then by Clark-Ocone formula, F is constant and thus F is indeed G measurable for any G.

Another example consists in taking $G = W_1$, then $C(F,G) = 1$ means that D_tF is independent of t a.s. and we claim that this implies the W_1-measurability of F. A simple proof consists in using the following one-parameter family of paths transformations $T_\theta(\theta \in [0,1))$ defined by

$$(T_\theta\omega)(t) = (1-\theta)\omega(t) + \theta t\omega(1)$$

which is clearly bijective and falls within the class of anticipating Girsanov transforms studied in Sect. 5. Then, one shows easily by density (using for instance the anticipating Girsanov theorems) that, for each $\theta \in [0,1]$, $FoT_\theta \in \mathbb{H}^1$ and

$$D_t(FoT_\theta) = (1-\theta)(D_tF)oT_\theta + \theta\int_0^1 (D_sF)oT_\theta ds = (D_tF)oT_\theta$$

since D_tF is independent of t. And, thus $D_t(FoT_\theta)$ is independent of t.

Clearly, there just remains to show that FoT_θ is independent of θ. In order to do so, we fix $\theta_0 \in [0,1)$ and we consider a regularization of the paths : for instance, let $N \geq 1$, let $W_t^N = W_{\frac{k}{N}} + \left(t - \frac{k}{N}\right)\left(W_{\frac{k+1}{N}} - W_{\frac{k}{N}}\right)$ if $\frac{k}{N} \leq t < \frac{k+1}{N}$, $0 \leq k \leq N-1$. Then, arguing once more by density, we obtain for h small enough

$$\begin{aligned}\frac{d}{dh}&\left(FoT_{\theta_0}\right)\left(\omega + h(\omega^N - t\omega(1))\right)\\ &= \int_0^1 \left(D_tFoT_{\theta_0}\right) dW_t^N - \left(\int_0^1 D_tFoT_{\theta_0}\, dt\right) W_1 \;=\; 0\end{aligned}$$

Hence, letting N go to $+\infty$, $FoT_{\theta_0}(\omega + h(\omega - t\omega(1))) = FoT_{\theta_0}$ for h small enough i.e. $FoT_\theta = FoT_{\theta_0}$ for θ close enough to θ_0, and we conclude.

Let us mention that this argument can be adapted to more general situations when $D_tF = a(\omega)b(t)$ for some deterministic function b, but we shall not pursue in that direction here.

References

1. Fournié, E., Lasry, J.-M, Lebuchoux, J. Lions, P.-L, Touzi, T.: Applications of Malliavin calculus to Monte-Carlo methods in finance. Fin. and Stoch. **3**, 391–412 (1999)
2. Fournié, E., Lasry, J.-M, Lions, P.-L.: Some non linear methods to study far-from-the-money contingent claims. Numerical Methods in Finance. Cambridge: University Cambridge Press 1997
3. Fournié, E., Lasry, J.-M, Touzi, N.: Monte Carlo methods in stochastic volatility models. Numerical Methods in Finance. Cambridge: University Cambridge Press 1997
4. Kusuoka, S.: The non-linear transformation of Gaussian *mesure* on Banach space and its *absolule* continuity (I). J. Fac. Sci. Univ. Tokyo IA, 567–597 (1982)
5. Malliavin, P.: Stochastic analysis. Berlin: Springer 1997
6. Nualart, D.: Malliavin calculus and related topics. Berlin: Springer 1995
7. Üstünel, A.S.: An introduction to analysis on Wiener's space. Berlin: Springer 1992
8. Üstünel, A.S., Zakai, M.: On independence and conditioning on the Wiener space. Ann. Proba., 1441–1453 (1989)
9. Üstünel, A.S., Zakai, M.: On the structure of independence on the Wiener space. J. Funct. Anal., 113–137 (1990)
10. Watanabe, S.: (1984) Stochastic differential equations and Malliavin calculus. Tata Inst. Fund. Research. Bombay: Springer 1984