Smart Monte Carlo: Various tricks using Malliavin calculus

Eric Benhamou[1]

Goldman Sachs International

Fixed Income Strategy

1st Floor River Court
120 Fleet Street
London EC4A 2BB- UK

January, 2001

[1] E-mail: Eric.Benhamou@gs.com. The views herein are the authors' ones and do not necessarily reflect those of Goldman Sachs.

Smart Monte Carlo: Various tricks using Malliavin calculus

Key words: Monte-Carlo, Greeks, Conditional expectation, Malliavin Calculus, Likehood Ratio, Homogeneity, Heston.

JEL classification: G13

Abstract: Current Monte Carlo pricing engines may face computational challenge for the Greeks, because of not only their time consumption but also their poor convergence when using a finite difference estimate with a brute force perturbation. The same story may apply to conditional expectation. In this short paper, following Fournié et al. (1999), we explain how to tackle this issue using Malliavin calculus to smoothen the payoff to estimate. We discuss the relationship with the likelihood ration method of Broadie and Glasserman (1996). We show on numerical results the efficiency of this method and discuss when it is appropriate or not to use it. We see how to apply this method to the Heston model.

1. Introduction

The growing emphasis on risk management issues as well as the development of more and more complicated financial products have urged to develop efficient techniques for the computation of price sensitivities with respect to model parameters. Moreover, the computation is not only done as the trader or book(s) level but also at the firm level, especially for the global computation of VAR and credit charge valuation, leading to raising concern about computational time.

In practice, generic Monte Carlo pricing engines may face computational challenge for the Greeks of discontinuous payoffs options, because of not only their time consumption but also their poor convergence when using a finite difference estimate with a brute force perturbation. In addition to the standard error on the numerical computation of the expectation, the finite difference Monte Carlo method contains another error on the approximation of the derivative function by means of its finite difference. This may give some hard time to the generic engine. The same story applies to conditional expectations where many paths might not be relevant.

In this short note, we discuss various methods to get fast convergence and show how these methods can apply to a generic Monte Carlo pricing engines as opposed to particular methods that would only spice up certain types of payoff but may not apply in a general framework. Mainly, we present a method based on Malliavin calculus that enables to smoothen the function to simulate, following ideas first established by Fournié et al. (1999). We also explain how to use various homogeneity properties in order to get the different Greeks. This allows us to only compute a few of the Greeks. Last but not least, we show that the integration by parts can also be applied to conditional expectation. We give in the appendix section an introduction to the Malliavin calculus.

2. Fast Greeks Computation

Introduction to Malliavin weights

We will fist see how to have smart Monte Carlo that compute fast Greeks. We will always assume that the functions are smooth enough to be able to perform the different computation referring to Benhamou (2000b) (2000c) and Fournié et al. (2001) for the technical assumptions required (mainly uniform ellipticity of the volatility operator). When using finite difference approximation for the Greeks, bumping the price and taking the sensitivity, one makes two errors: one on the numerical computation of the expectation via the Monte Carlo as for any simulations, and another one on the approximation of the derivative function by means of its finite difference. As of the gamma, this leads for example to $\frac{F(x+\varepsilon)-2F(x)+F(x-\varepsilon)}{2\varepsilon}$, which means that one approximate the second order derivative of the payoff function by $f''(x) \approx \frac{f(x+\varepsilon)-2f(x)+f(x-\varepsilon)}{2\varepsilon}$. This is obviously very inefficient for very discontinuous payoff like for binary, range accrual, barrier and other type of digital options for example. To overcome this inefficiency, Broadie and Glasserman (96) suggested using the likelihood ratio method. If we are interested in the sensitivity of the option price with respect to some parameter θ, and if we know explicitly the density function of the underlying variable and can expressed it in terms of the parameter θ by $p(x,\theta)$, we can compute the Greek by:

$$\frac{\partial}{\partial\theta}E\left[f\left(X_T\right)\right]=\frac{\partial}{\partial\theta}\int f\left(x\right)dp(x,\theta)dx=E\left[f\left(x\right)\frac{\partial}{\partial\theta}\ln p(x,\theta)\right] \quad (2.1)$$

The interest of this approach was to come up with an efficient way of avoiding the differentiation of the payoff function. In fact, rewriting it more formally, the Greeks can be computed as the expectation of the original payoff times a weight:

$$Greek=E\left[f\left(X_T\right)weight\right] \quad (2.2)$$

However, this method was quite restrictive since one needs to know explicitly the density function. This is precisely where M. calculus could provide a solution. In an inspiring article, Fournié et al. (1999) proved that any Greek could be expressed as an expectation of the payoff time a weight. They show that this weight could be expressed in terms of the Malliavin derivative (in the following M. derivative), without knowing explicitly the density function. Benhamou (2000b) (2000c) and Fournié et al. (2001) examined the different possible weights, mentioning that they exist an infinity of weighting function and proved that the weight of minimal total variance is precisely the one given by the likelihood ratio method. Benhamou (2000b) (2000c) also introduced the weighting function generator and showed that any weight could also be expressed as the Skorohod integral of the weighting function generator. We will now follow the presentation of Benhamou (2000c) for the delta (extensions to the other Greeks are easy and can be found in Benhamou (2000c))

We will denote in the following by X_T the underlying, $Y_T=\frac{\partial}{\partial x}X_T$ its first variation process (derivatives of X_T with respect to its initial condition x and $\partial(.)$ the Skorohod integral[2]. We want to take the derivative of the price with respect to the underlying initial condition:

$$\frac{\partial}{\partial x}E\left[f\left(X_T\right)\right]=E\left[\frac{\partial}{\partial x}f\left(X_T\right)\frac{\partial}{\partial x}X_T\right]=E\left[\frac{\partial}{\partial x}f\left(X_T\right)Y_t\right] \quad (2.3)$$

If this can be written as the expectation of the payoff function times a weight (expressed as a Skorohod integral $\partial(u)$). We have on the other hand

$$E\left[f\left(X_T\right)\partial(u)\right]=E\left[\int D_s f\left(X_T\right)uds\right]=E\left[\frac{\partial}{\partial x}f\left(X_T\right)\int D_s X_T uds\right]$$

$$=E\left[\frac{\partial}{\partial x}f\left(X_T\right)\int\sigma(s,X_s)Y_tY_s^{-1}1_{\{s<T\}}uds\right] \quad (2.4)$$

where we have successfully used the integration by part formula (A.3), the chain rule (A.2) and the expression of the M. derivative with respect to its first variation process (A.6). Expressions (2.3) and (2.4) are equal if and only if $E\left[\frac{\partial}{\partial x}f\left(X_T\right)Y_t\right]=E\left[\frac{\partial}{\partial x}f\left(X_T\right)\int\sigma(s,X_s)Y_tY_s^{-1}1_{\{s<T\}}uds\right]$ for any function f.

A trivial solution is given by $1=\int\sigma(s,X_s)Y_s^{-1}uds$ or $u=\frac{Y_s}{\sigma(s,X_s)T}$. (2.5)

Similarly, we can apply this for the gamma. We provide in table 1 a summary of the conditions required for the delta and gamma for European option. Extension to other payoff type can be found in Fournié et al (1999) and Benhamou (2000a) (Asian options) and in Gobet and Kohatsu Higa (2001) (barrier and lookback options). The definition of other Greeks, in particular the vega for stochastic volatility models, is very specific to the model. This is why we limit our study in this short note to the delta and gamma (more relationships to specific models can be found in Benhamou (2000b)).

[2] see the appendix for an introduction to Malliavin calculus.

Greek	Weight
Delta	$\frac{1}{T}\partial\left(\int\frac{Y_s}{\sigma(s,X_s)}ds\right)$ (2.6)
Gamma	$\partial\left(\frac{1}{T}\frac{\partial}{\partial x}\int\frac{Y_s}{\sigma(s,X_s)}ds\right)+\partial\left(\frac{\partial}{\partial x}\int\frac{Y_s}{\sigma(s,X_s)}ds\right)\partial\left(\frac{\partial}{\partial x}\int\frac{Y_s}{\sigma(s,X_s)}ds\right)$ (2.7)

Table 1: European Weight for a general diffusion.

Let us say in passing that these relationships are very general and assume that the underlying is modeled by a jump diffusion model with the jump component independent from the Brownian motion.

$$X : dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t + \lambda(t)dJ_t \tag{2.8}$$

The jump part can be of course null, leading to standard SDE, and the volatility can be either deterministic or stochastic. It is then easy to apply this to specific model (table 2), using the fact that:

$$\iint_{u,s} f(u)dW_u f(v)dW_v = \left(\int_u f(u)dW_u\right)^2 - \int_u f^2(u)du \tag{2.9}$$

Let us remind that Heston model is described by a stochastic volatility model given by

$$dX_t = rX_t dt + \sigma_t X_t dW_t^1, \qquad d\sigma_t^2 = \lambda(\theta - \sigma_t^2)dt + v\sigma_t dW_t^2 \tag{2.10}$$

$$\text{with } \rho dt = E\left[dW_t^1 dW_t^2\right] \tag{2.11}$$

The conditions of table 1 can be applied to many models. We have given in table 2 the explicit form of the weights for the Black Scholes and Heston model.

Greek	Black Scholes	Heston
Delta	$\frac{W_T}{\sigma x T}$ (2.12)	$\frac{1}{xT}\int_0^T\frac{dW_s^1}{\sigma_s}$ (2.13)
Gamma	$-\frac{W_T}{\sigma T x^2}+\frac{W_T^2-T}{(\sigma T x)^2}$ (2.14)	$-\frac{1}{x^2T}\int_0^T\frac{dW_s^1}{\sigma_s}+\frac{1}{(xT)^2}\left(\left(\int_0^T\frac{dW^1 s}{\sigma_s}\right)^2-\int_0^T\frac{ds}{\sigma_s^2}\right)$ (2.15)

Table 2: European Weight for the Black Scholes and Heston model.

Characteristics of Malliavin weights

Let us now summarise some important results about Malliavin weights (keeping in mind for the design of a general Monte Carlo engine):

- All Greeks can be written as the expected value of the payoff times a weight function. The weight functions are independent from the payoff function. This has two implications.
  - First, the Malliavin method will comparatively (to finite difference) increased its efficiency for discontinuous payoff options. As a rule of thumb, the Malliavin method is appropriate for option for which the mean-square convergence of a shifted option $P(x+\varepsilon)$ to the normal one $P(x)$ is linear in ε This is the case of any option with a payoff expressed as a probability that a certain event occurs conditionally to the underlying level at a certain time (case of any binary and corridor option). For a general pricing engine, using certain (numerical) criteria of smoothness, we shall be able to branch on the appropriate method. Because it is in a sense independent from the payoff function, the general implementation is simpler that the one of variance reduction technique that only apply to very specific payoff (like the use of control variate).
  - Second, no extra computation is required for other payoff function as long as the payoff is a function of the same points of the Brownian trajectory. This can be cached in memory to make it efficient.
- There is an infinity of solutions for the generator function. However, the optimal weighting function is the one which is measurable with respect to the payoff variables. This means in practice that the weight functions will be expressed with the same points of the Brownian motion trajectory as the option payoff, therefore requiring no extra points computation.

♦ The weighting function smoothens the function to simulate (as the payoff function does not require to be numerically differentiated) but introduces some extra noise. It smoothens twice the payoff function in the case of the gamma as it reduces a second order differentiation to no differentiation, leading to high efficiency for the simulation of the gamma (see figure1 for the comparative efficiency of the Malliavin method in the case of the gamma of a corridor option). It introduces a lot of noise in the simulation as the weighting function explodes for small maturities, imposing some criteria for critical maturities.

♦ For homogeneous model, like Black Scholes or Heston, we can derive some proportionality rules (see for example Reiss and Wystupe (2001)). In particular, there exists some relationship between the vega and the gamma in the Black Scholes model. This has two implications: the simulation of gamma and vega can be done at once and the performance of the vega computations is very similar to the one of the gamma. This can also be understood from the meaning of the vega. The vega in the case of Black Scholes is a compound differentiation. The smoothing introduced by Malliavin method is therefore twice for the vega. Using the homogeneity property of the Greeks makes sure that their computation is consistent and non arbitrageable.

♦ The Malliavin method leads to weighting functions which are roughly (polynomial) functions of the Brownian motion. The variance of the weighting function increases for high values of the Brownian motion. This implies that if the payoff function is very small for high value of the Brownian motion, the variance is going to be low. This indicates that Malliavin formulae are more efficient for put than call options. Two remarks should be made. First, it is more appropriate to use the put-call parity and therefore to calculate Greeks only for a put, second, one should use a localization of the Malliavin weight only at the discontinuity of the payoff and elsewhere avoid introducing extra noise with the Malliavin weight.

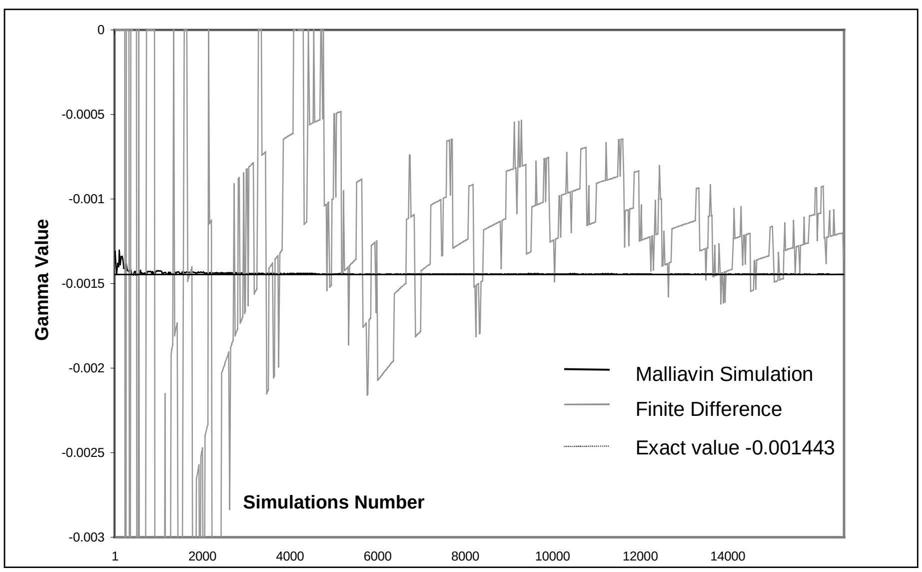

Figure 1: Efficiency of the Malliavin weighted scheme for the computation of the gamma of a Corridor option. The parameters of this option are: S_0=100, r=5%, σ =15%, T=1year, $S_{\min}$=95, $S_{\max}$ =105

In order to illustrate these remarks, we show two simulations done for the gamma of a European corridor and call option in a Black Scholes model. Figure 1 is a school case of an appropriate use of Malliavin method. It shows the gamma of a corridor option defined as an option to pay 1 if the underlying at maturity is between $S_{\min}$ and $S_{\max}$. The payoff of a corridor option has two discontinuities, the mean square convergence of the bumped price is only linear in ε and the Malliavin method smoothens twice the Greek to simulate in the case of the gamma. Figure 2 is an example of inappropriate use of Malliavin method. The mean square convergence of the bumped price is quadratic, the payoff is not discontinuous, it is only its derivative function that has only one discontinuity at the strike. The Malliavin method introduces extra noise in the simulation with the weight to simulate. The call put parity has not be used, therefore creating high variance for high values of the Brownian motion.

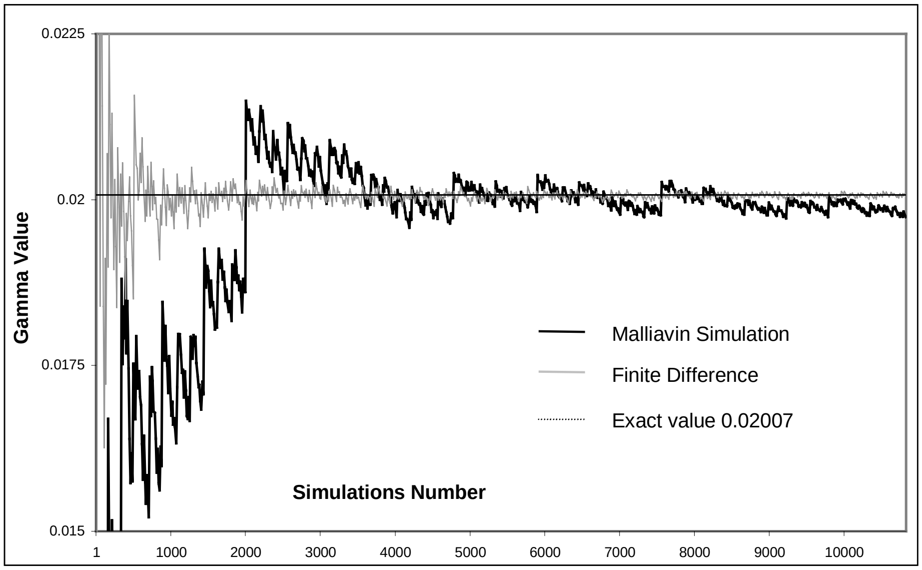

Figure 2: Efficiency of the Malliavin weighted scheme for the computation of the delta of a call option. The parameters are similar to the corridor option with a strike of 100.

Localisation of Malliavin weights

What we have shown so-far is that any Greeks could be written as $Greek = E\left[f\left(X_T\right)weight\right]$. This formula holds for any payoff. This formula will help to smoothen the function to simulate (as the payoff function does not require to be numerically differentiated) but the weight will introduce some extra noise. A good way of limiting this extra noise is to "localise" the integration by parts. Let us explain on a simple example. If the payoff function has some discontinuity at a strike K, we can rewrite it as

$$f\left(X_T, K\right) = f\left(X_T, K\right)\phi\left(X_T\right) + \left(1-\phi\left(X_T\right)\right)f\left(X_T, K\right) \tag{2.16}$$

where $\phi\left(X_t\right)$ is a smooth localisation function (say Lipschitz) that has its support in $\left[K-\alpha, K+\alpha\right]$. We can now process to the integration by parts and come up with a formula of the type

$$Greek = E\left[f\left(X_T\right)\phi\left(X_T\right)weight\right] + \frac{\partial}{\partial x}E\left[f\left(X_T, K\right)\left(1-\phi\left(X_T\right)\right)\right] \tag{2.17}$$

where the second part can be computed via an appropriate finite difference scheme that introduces no extra noise or even better, via an explicit differentiation of the payoff. Let us also mention that the localisation formula can be also done by taking a smoother function that "approaches" in a sense the payoff function. If the payoff is very discontinuous, we can always find a function that is smoother and is a good approximation of the payoff. In the case of a digital option, a smooth approximation $\phi_\alpha\left(X_T\right)$ could be a function that is piecewise linear, equal to 0 for. $X_T \le K-\alpha$ and 1 for $X_T \ge K+\alpha$ and linear in between. The payoff of an up digital, $1_{\{X_T \ge K\}}$ can be rewritten in terms of the smooth function $f\left(X_T\right) = \phi_\alpha\left(X_T\right) + f\left(X_T\right) - \phi_a\left(X_T\right)$. In this expression, only the second terms $f\left(X_T\right) - \phi_a\left(X_T\right)$ is now discontinuous and will require a smoothen expression expressed in terms of the Malliavin weight. Obviously, this can be repeated many times and we can for instance expressed our discontinuous function in terms of smooth polynomial approximation functions. We shall not pursue here in that direction even if we believe that the efficient approximation of the discontinuous function will be an interesting area of research in the coming years.

3. Conditional expectations and anticipative Monte Carlo

Malliavin weights

This section tackles what is believed to be one of the most promising application of Malliavin calculus to finance, namely the transformation of conditional expectation in non conditional ones and the use of Malliavin calculus for anticipative Monte Carlo. It is well known that conditional expectations offer the computational challenge to require a very high number of paths since "almost all" paths may miss the target event involved in the conditional expectations. In fact, at least when written formally, this problem is very similar to the one above (the computation of the Greeks). A conditional expectation can be formally represented as the ratio of two conditional expectation. We will here follow the presentation of Fournié et al. (2001). Let us assume that the condition is expressed in terms of a constraint of the type $G(X_T)=0$, of probability $E[\partial_0(G(X_T)=0)]$ where ∂_0 represents the Dirac function in zero. We have the symbolic calculation

$$E[F(X_T)\,|\,G(X_T)=0]=\frac{E[F(X_T)\partial_0(G(X_T))]}{E[\partial_0(G(X_T))]} \tag{3.1}$$

but of course the Dirac function is the derivative function of the Heavyside function $H(x)=1_{\{x\geq 0\}}+k$, and using similar computation as in section 2, we can immediately see that we can integrate this by parts. Let us assume that there exists a weight expressed as a Skorohod integral $\partial(u)$ so that we have

$$E[F(X_T)\partial_0(G(X_T))]=E[F(X_T)H(G(X_T))\partial(u)] \tag{3.2}$$

Using successively the integration by parts formula (A.3) and the rule for the M. derivatives of a product (A.8) and the chain rule (A.2), we get

$$E[F(X_T)H(G(X_T))\partial(u)]=E\left[\int D_t(F(X_T)H(G(X_T)))u_t\,dt\right] \tag{3.3}$$

$$E[F(X_T)H(G(X_T))\partial(u)]=E\left[\int F'(X_T)D_tX_TH(G(X_T))u_t\,dt\right]+E\left[\int F(X_T)\partial_0(G(X_T))D_tG(X_T)u_t\,dt\right] \tag{3.4}$$

If we want this to be true for any payoff function, we see that in fact the equation (3.2) cannot hold directly. In fact, we can rather remove the first term of the integration by part and impose the second term to be equal to $E[F(X_T)\partial_0(G(X_T))]$. A sufficient condition is

$$\int D_tG(X_T)u_t\,dt=1 \tag{3.5}$$

We then have the following very important way of computing conditional expectation. If we can find a weighting function generator u that satisfies the condition (3.5), we have immediately the obvious result

$$E[F(X_T)\,|\,G(X_T)=0]=\frac{E\left\{F(X_T)H[G(X_T)]\partial(u)-F'(X_T)H[G(X_T)]\int D_tX_Tu_t\,dt\right\}}{E[H(G(X_T)\partial(u))]} \tag{3.6}$$

Moreover, if we can find an orthogonal weight satisfying both (3.5) and the following orthogonality condition

$$E\left[\int D_tF(X_T)u_t\,dt\right]=0 \tag{3.7}$$

We then have that the conditional expectation is even simpler and equal to

$$E[F(X_T)\,|\,G(X_T)=0]=\frac{E\{F(X_T)H[G(X_T)]\partial(u)\}}{E[H(G(X_T)\partial(u))]} \tag{3.8}$$

Obviously, imposing the two conditions (3.5) and (3.7) may impose some restrictions on the two stochastic variables F,G and may not hold for any function F,G. This part is the subject of a small digression in the section below called "Functional dependence and Malliavin calculus", just after the numerical example. However, before embarking into a numerical example, we will see the explicit expression of the weight when we know the density function. This follows the same line as the comparison of the likelihood ration method and the M. weights for the Greeks.

Weights for explicit densities

Interestingly, when we know the density function, we can express explicitly the weight with respect to the density. In fact, there is two ways of doing it:

- If the function $F(X_T)$ is smooth, we may want to use it and shift the derivation operator on this function to inherit a formula with some smoothness. This integration by part is formally equal to

$$E[F(X_T)\partial_0(G(X_T))]=\int F(x)\partial_0(G(x))p(x)dx=-\int\frac{\partial}{\partial x}(F(x)p(x))H(G(x))dx \tag{3.9}$$

$$E[F(X_T)\partial_0(G(X_T))]=-\int\left(\frac{\partial}{\partial x}\ln[p(x)F(x)]\right)H(G(x))F(x)p(x)dx=E[F(X_T)H(G(X_T))\pi] \tag{3.10}$$

with the weight

$$\pi=-\frac{\partial}{\partial x}\ln[p(X_T)F(X_T)] \tag{3.11}$$

- If the function $F(X_T)$ is not smooth at all but independent from the function $G(X_T)$ we may want to split the expression in independent terms.

$$E[F(X_T)\partial_0(G(X_T))] = \int F(x)\partial_0(y)p(x,y)dx = -\int F(x)H(G(x))\frac{\partial}{\partial y}p(x,y)dx \qquad (3.12)$$

$$E[F(X_T)\partial_0(G(X_T))] = -\int F(x)H(G(x))q(x,y)p(x,y)dx = E[F(X_T)H(G(X_T))\pi] \qquad (3.13)$$

with the weight

$$\pi = -\frac{\partial}{\partial y}\ln p(X_{Tt}, Y_T) \qquad (3.14)$$

Numerical experiments and implementation rules

Conditional expectations are of great importance for calibration. For example, we may need to compute the overall volatility knowing the final value. Conditional expectations shall also change the understanding of Monte Carlo method. Usually, Monte Carlo methods are thought to be forward looking[3]. One gives an initial point and diffuses the underlying. Standing on the other extreme, PDEs methods are thought to be backward looking. One gives a final point and back propagates. This allows computing American and Bermudean option with the second method while path dependent products for the first one. But if one knows how to express any conditional expectation where the condition is that the underlying price is equal to a given value at a given time, one can also do some backward looking computation with Monte Carlo. This shows that the overall accepted separation between Monte Carlo and PDEs methods is too simplistic and misses some recent development (see Fournié et al (2001), Lions and Régnier (2001) for a deeper discussion on this).

Let us take again the Heston model described by (2.10) and (2.11). We are interested in computing the conditional volatility

$$E\left[\sigma_T^{\ 2} \mid S_T = S\right] \qquad (3.15)$$

We have in this case that the underlying is a two dimensional process with $X_t = (S_t, \sigma_t)$, $F(X_t) = \sigma_T^{\ 2}$, while $G(X_T) = S_T - S$

In order to simplify the computation of the Skorohod integral, we will assume zero correlation between the underlying and its stochastic volatility. We can easily apply the calculation of the previous paragraph. First, because of the independence of W_t^1 and W_t^2, we have that $D_s\sigma_T = 0$ and therefore $D_sF(X_T) = 0$, so that the orthogonality condition (2.7) holds for any weight. Moreover, using the relationship between the M. derivatives and its first variation process (A.6), we get that the condition (3.5) is equal to

$E\left[\int_0^T \sigma_s S_s \frac{Y_T}{Y_s} u_t dt\right] = 1$. An easy solution is given for $u_t = \frac{1}{T\sigma_t S_T}$, leading to the following weight

$$\partial\left(\frac{1}{T\sigma_t S_T}\right) = \frac{1}{TS_T}\int_0^T \frac{dW_t^{\ 1}}{\sigma_t} + \frac{1}{T}\int_0^T \frac{D_t S_T}{\sigma_t S_T^{\ 2}} = \frac{1}{TS_T}\int_0^T \frac{dW_t^{\ 1}}{\sigma_t} + \frac{1}{S_T} \qquad (3.16)$$

where we have used the rule for the Skorohod integral of a product (A.5), with $u = \frac{1}{\mathrm{T}\sigma_t}$ and $F = \frac{1}{S_T}$.

We have finally the following formula

$$E\left[\sigma_T^{\ 2} \mid S_T = S\right] = \frac{E\left\{\frac{\sigma_T^{\ 2}}{S_T}.H(S_T - S)\left[1 + \frac{1}{T}\int_0^T \frac{dW_t^{\ 1}}{\sigma_t}\right]\right\}}{E\left\{\frac{1}{S_T}.H(S_T - S)\left[1 + \frac{1}{T}\int_0^T \frac{dW_t^{\ 1}}{\sigma_t}\right]\right\}} \qquad (3.17)$$

The numerical experiment has been to compute the conditional expectation given by formula (3.17) for $S_0 = 100$, $r = 5\%$, $\lambda = 1\%$ $\sigma_0 = 30\%$, $\theta = 2.25\%$, $v = 5\%$. We have displayed for T =0 to 6

[3] Even if there has been some recent development for American Monte Carlo

months and for value of the spot between $S_{\min}$=70. Obviously, we get the characteristic U shape of Heston model with zero correlation.

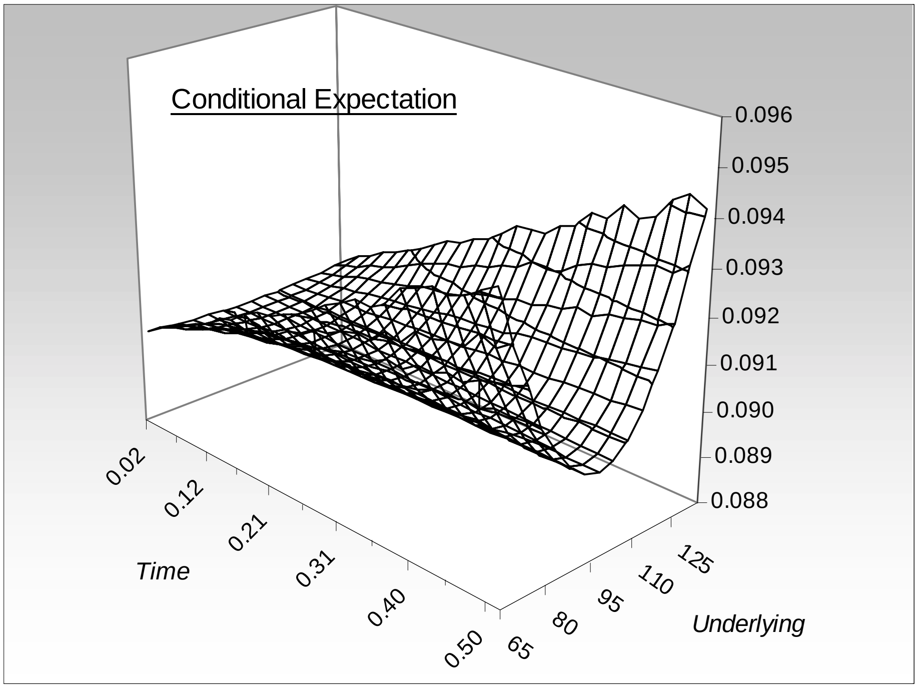

Figure 3: Example of Conditional expectation computed via Malliavin calculus. In this case, We computed in a Heston model the conditional volatility $E\left[\sigma_T^{\ 2} \mid S_T = S\right]$

Functional dependence and Malliavin calculus

We have seen in the previous section that the colinearity between the two functions, the one to estimate $F(X_T)$ and the one of the condition $G(X_T)$ plays a role in the weight. This was obvious when knowing explicitly the density, as we had to split the integration by parts into two cases. In fact, when looking at the two relationships (3.5) and (3.7), we could already realize that they could not hold both if the two functions D_tF and D_tG were proportional. This leads us to look at the notion of dependence between the two M. derivatives of $F(X_T)$and $G(X_T)$. We believe that the concept of functional dependence using M. derivative is very general and could be useful in finance to extend notion of non-linear correlation for any stochastic variables F,G. This could have certainly some influence over non-linear VAR approach. We will just briefly introduce the subject as this is slightly out of the scope of this short note.

We call functional dependence the function

$$C(F,G) = \sup_{\omega} ess \left\{ \frac{\left| \int D_t F D_t G dt \right|^2}{\int |D_t F|^2 dt \int |D_t G|^2 dt} \right\} \tag{3.18}$$

Obviously, we have $0 \le C(F,G) \le 1$ (the notion of sign for functional dependence is meaningless since the orientation of the vectorial functional space has no sense). $C(F,G)=0$ means that the two M. derivatives are orthogonal, which can be shown to be equivalent to the independence of F,G. Intuitively, the orthogonality of the stochastic gradient of the two functions means that their evolution is unrelated, or equivalently that these two variables evolve independently. $C(F,G)=1$ is equivalent to the two M. derivatives being co-linear. One can show that if F is G-measurable, the two M. derivatives will be co-linear. In a sense the M. derivatives functional dependence measure how a given variable F is G-measurable and vice versa.

4. Conclusion

In this paper, we have shown that there exist various tricks to enhance the performance of general pricing Monte Carlo. Using appropriate expression of the expectation to simulate is crucial for fast and accurate result. We have explained how to use Malliavin calculus to explicitly do some integration by parts when not knowing the density function of the underlying diffusion. We have applied this to two main applications: computation of the Greeks and of conditional expectations.
We believe that this is a very promising area of research and will progressively change the understanding of Monte Carlo methods as it paves the path for very generic forward/backward Monte Carlo, following the recent trend of improvement of American Monte Carlo.

Appendix: a primer on Malliavin calculus

The objective of this short primer is to give an intuitive presentation of Malliavin calculus. For a more rigorous and detailed explanation, we refer the reader to the exhaustive book of Nualart (1995).
Malliavin calculus is a synonym of calculus of variation of stochastic processes. Even if its original motivation was to provide a probabilistic proof of the existence and smoothness of solutions of particular PDEs (the of Hormander's sum of squares theorem), M. calculus has turned out to be a very powerful tool for giving other representation of stochastic processes, allowing to prove certain properties of stochastic processes (especially smoothness conditions). Because the Brownian motion is not differentiable in the traditional sense, M. calculus defines a derivative, using a local perturbation on the Brownian motion and more generally on a martingale process. It measures in a sense the impact of bumping locally the Brownian path. Let us take a function of the Brownian motion $(W_t)_{t\geq 0}$, $F: t \to F(W_t)$. Let us bump the Brownian motion only locally at a time s. In mathematical terms, the perturbed Brownian motion is the superposition of the original Brownian motion and a Kronecker function of total measure ε: $W_t + \varepsilon\delta_s$, where $\delta_s(u) = 1_{\{s=u\}}$. The M. derivative is defined intuitively as

$$D_s F : t \to \lim_{\varepsilon \to 0} \frac{F(W_t + \varepsilon\delta_s) - F(W_t)}{\varepsilon} \tag{A.1}$$

where the limit can usually be interpreted as a.s. This trivially leads to the M. derivative of a Brownian motion given by the indicative function: $D_s W_t = 1_{\{s \leq t\}}$

The interest of the M. calculus is to satisfy usual derivation rules:

- Chain rule for compound function, $\Phi : t \to G(F_1(W_t), F_2(W_t))$

$$D_s \Phi = \sum_i \frac{\partial}{\partial x_i} G.D_s F_i \tag{A.2}$$

- Integration by parts, (or duality between the M. derivative and the Skorohod integral).

$$E\left[\int D_s F u ds\right] = E[F\partial(u)] \tag{A.3}$$

where $\partial(u)$ is called the Skorohod integral. This relation is the cornerstone formula as it enables to smoothen the function inside the expectation. Intuitively, the Skorohod integral could be compared to the divergence operator[4] (up to the minus sign) as for deterministic function on (R^n, λ^n), we have $\int_{R^n} \langle \nabla f, u \rangle_{R^n} d\lambda^n = \int_{R^n} f(-divu) d\lambda^n$.

- Skorohod integration: for adapted processes, the Skorohod integral coincides with the Ito integral

$$\partial(u) = \int u dW_t \tag{A.4}$$

Moreover, the Skorohod integral satisfies some interesting properties

$$\partial(Fu) = F\partial(u) - \int D_t F u dt \tag{A.5}$$

- M. derivatives of a jump-diffusion: Let $(X_t)_{t\geq 0}$ defined by its jump-diffusion equation: $X : dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t + \lambda(t)dJ_t$ with initial condition $X_0 = x$

[4] Some authors refer to the Skorohod integral as the stochastic divergence operator.

And let us define its first variation process (also called the tangential process) $(Y_t)_{t\geq 0}$ defined as $Y_t = \frac{\partial}{\partial x} X_t$, obviously $Y_0 = 1$ and

$$Y : dY_t = \frac{\partial}{\partial x} b(t, X_t) Y_t dt + \frac{\partial}{\partial x} \sigma(t, X_t) Y_t dW_t$$

The M. Derivatives of $(X_t)_{t\geq 0}$ is then given by

$$D_s X_t = \sigma(s, X_s) Y_t Y_s^{-1} 1_{\{s\leq t\}} \quad \text{(A.6)}$$

Let us conclude by saying in passing that the M. derivative satisfies standard rule of derivation, namely for a product, we have

$$D_t(FG) = D_t F.G + F.D_t G \quad \text{(A.7)}$$

References

Benhamou E.: 2000a, An Application of Malliavin Calculus to Continuous Time Asian Options, London School of Economics Working paper.

Benhamou E.: June 2000b, Application of Malliavin Calculus and Wiener Chaos to Option Pricing Theory, Ph.D. Thesis, London School of Economics

Benhamou E.: December 200c, Optimal Malliavin Weighting Function for the Computation of the Greeks, Proceedings of the Monte Carlo Congress, Monte Carlo June 200.

Broadie, M., Glasserman, P.: Estimating security price derivatives using simulation, Manag. Sci. 42, 269–285 (1996)

Fournié E., Lasry J.M., Lebuchoux J., Lions P.L. and Touzi N.: 1999, Applications of Malliavin Calculus to Monte Carlo methods in Finance, Finance and Stochastics 3, 391-412.

Fournié E., Lasry J.M., Lebuchoux J. and Lions P.L.: 2001, Applications of Malliavin Calculus to Monte Carlo Methods in Finance. II., Finance and Stochastics .

E. Gobet, A. Kohatsu-Higa: 2001, Computation of Greeks for barrier and lookback options using Malliavin calculus, Ecole Polytechnique, CMAP, R.I. N464 .

Nualart, D.: Malliavin Calculus and Related Topics. (Probability and its Applications) Berlin Heidelberg New York: Springer 1995.

Lions P.L., Régnier H.: 2000, Monte-Carlo computations of American options via Malliavin calculus, Monte Carlo 2000 conference mimeo.

Reiss O., U. Wystup: 2001, Computing Option Price Sensitivities Using Homogeneity and Other Tricks, Journal of Derivatives, Winter 2001, 41-53