Application of Malliavin Calculus and Wiener Chaos to Option Pricing Theory

by

ERIC BEN-HAMOU

The London School of Economics and Political Science

Thesis submitted in partial fulfilment of the requirements of the degree of

Doctor of Philosophy

at the University of London

Abstract

This dissertation provides a contribution to the option pricing literature by means of some recent developments in probability theory, namely the Malliavin Calculus and the Wiener chaos theory. It concentrates on the issue of faster convergence of Monte Carlo and Quasi-Monte Carlo simulations for the Greeks, on the topic of the Asian option as well as on the approximation for convexity adjustment for fixed income derivatives.

The first part presents a new method to speed up the convergence of Monte-Carlo and Quasi-Monte Carlo simulations of the Greeks by means of Malliavin weighted schemes. We extend the pioneering works of Fournié et al. (1999), (2000) by deriving necessary and sufficient conditions for a function to serve as a weight function and by providing the weight function with minimum variance. To do so, we introduce its generator defined as its Skorohod integrand. On a numerical example, we find evidence of spectacular efficiency of this method for corridor options, especially for the gamma calculation.

The second part brings new insights on the Asian option. We first show how to price discrete Asian options consistent with different types of underlying densities, especially non-normal returns, by means of the Fast Fourier Transform algorithm. We then extends Malliavin weighted schemes to continuous time Asian options.

In the last part, we first prove that the Black Scholes convexity adjustment (Brotherton-Ratcliffe and Iben (1993)) can be consistently derived in a martingale framework. As an application, we examine the convexity bias between CMS and forward swap rates. However, for more complicated term structures assumptions, this approach does not hold any more. We offer a solution to this, thanks to an approximation formula, in the case of multi-factor lognormal zero coupon models, using Wiener chaos theory.

Acknowledgements

I enjoyed the help and support of numerous friends and colleagues while working on my dissertation. My greatest debt, however, goes to Doctor Pierre Mella-Barral, my supervisor at the London School of Economics and to Professor Nicole El Karoui, who initiated me to financial mathematics and nurtured my passion for this fantastic research field. Doctor Pierre Mella-Barral, whose fantastic breadth of knowledge and intuition, I have been privileged enough to have access to, guided me in countless meetings. He helped me to structure and clarify my thought and suggested valuable insightful comments. Professor Nicole El Karoui provided invaluable support in reading nearly all versions of all chapters of this Thesis. I am very much indebted to her for being thorough, where I was sloppy, and ingenious, where I was lost. I also thank her for numerous occasions on which her encouragement helped me have the stamina to get to the end of the task. I am very thankful to Professor David Webb, Head of the Financial Markets Group for his friendship and for his trust. I would like to thank Margaret Bray and Bob Nobay for their support and useful remarks.

I am particularly indebted to my girlfriend, Beatrice. She assisted me at each stage of the dissertation process, starting with early discussions of ideas through reading every draft of every chapter. She deserves more love and gratitude than could fit possibly on this page. I would like to express my gratitude to Alexandre Duguet for fruitful conversations and Grigorios Mamalis, whose kindly spent a couple of dozen of hours reading the chapters of this Thesis and giving me helpful and relevant feedback. Of course, all remaining errors are mine.

I appreciated the remarks of the participants of the Ph.D. seminar of the Financial Markets Group, the ones of the Ph.D. seminar of the Centre de Mathématiques Appliquées at Ecole Polytechnique as well as the ones of the Ph.D. seminar in Economics.

I am also indebted to several institutions for their supports: the Financial Markets Group and the Centre de Mathématiques Appliquées de l'Ecole

Polytechnique. I thank all of the F.M.G. members including the administrative staff, Sarah Vink, Tarah Lehane and Pippa Maddicott, for creating and maintaining such a wonderfully friendly place, which puts pressure off rather than on students and provides such an excellent environment for research and discussions. I am very appreciative of my fellow students Simon Emrich and Alexander Mürmann. At the Ecole Polytechnique, I was among the privileged ones to become a member of the research group in Financial Mathematics at the Centre de Mathématiques Appliquées at Ecole Polytechnique. I would like to mention Rama Comt, Jeanne Bailleul for their nice support.

Last but not least, I am especially grateful to all my family and especially to my parents, Gilbert and Sylvie Ben-Hamou, for their support and for the confidence they have put in me. I am very much indebted to my sister Agnes Ben-Hamou who kindly spent some hours correcting my english. I am thankful to my brother Joel Ben-Hamou for providing me a computer and an internet access every time I came back to France. And finally I would like to thank other relatives, friends, and teachers who have positively influenced my life.

London, June 07, 2000

Eric Ben-Hamou

Contents

List of Figures

List of Tables

Introduction

This dissertation provides a contribution to the option pricing literature by means of some recent developments in probability theory, namely the Malliavin Calculus and the Wiener chaos theory. It sheds new light on some old and complicated problems by means of these new techniques. It concentrates on the issue of faster convergence of Monte Carlo and Quasi-Monte Carlo simulations for the Greeks, on the topic of the Asian option as well as on the approximation for convexity adjustment for fixed income derivatives. This Thesis consists of six different chapters. Among these, chapters 3 and 5 try to contrast the aforementioned new tools of probability theory by presenting other approaches like the Fast Fourier Transform technique and the martingale framework.

The first part presents a new method to speed up the convergence of Monte-Carlo and Quasi-Monte Carlo simulations of the Greeks by means of the Malliavin weighted scheme. The pioneering works were the ones of Fournié et al. (1999), (2000). However, two important questions remained unsolved. Can we derive necesary and sufficient conditions for a function to serve as a weight function? Which weight function has the minimum variance?

The first chapter tries to answer these two questions. To be able to provide necessary and sufficient conditions for the weighting function, we introduce its generator defined as its Skorohod integrand. This new definition turns out to be very powerful since it provides a description of all weight functions. An integration by parts, by means of Malliavin calculus, leads to these conditions. These conditions, expressed through conditional expectations, provide the whole set of weight functions for an option pricing kernel in a continuous-time model. We show how to find the ones with minimum variance. This minimum-variance solution is the projection of any weight function on the filtration spanned by the payoff functional. We give some key examples of the weight function generator. It turns out that in some cases, the optimal solution is not easy to calculate explicitly. We discuss the question of the most appropriate weight function in this complicated case. We finally conclude that this method is very efficient for discontinuous payoff options, like binary and corridor options. This is a consequence of the fact that this method avoids differentiating the payoff function.

The second chapter is a numerical application of this general theory in the case of the Black pricing model. We quantify the gain in the variance reduction when using the Malliavin weighted scheme. We find evidence of spectacular efficiency of this method for corridor options, especially for the gamma calculation. Indeed, the Malliavin weighted scheme variance reduction should be more efficient for second order derivatives compared to first order ones, ceteris paribus. We examine, furthermore, a mixed strategy based on the Malliavin weighted scheme and finite difference approximation. The Malliavin weighted scheme is used only locally, at the kink of discontinuity. This leads to so called "local Malliavin" formulae. This method appears to be a very efficient way to simulate the Greeks, either for very standard payoffs like call options or more discontinuous ones. A subtle point of this method concerns the choice of the location of the discontinuity. We conjecture that this is depending on the form of the payoff functional.

The second part brings new insights on the Asian option by means of the Malliavin calculus. It first studies an alternative to this probabilistic method by means of the Fast Fourier Transform algorithm. It shows how to compute discrete Asian options consistent with different types of underlying densities, especially non-normal returns as suggested by the empirical literature (see Mandelbrot (1963) and Fama (1965) for the early ones). The interest of this method is its flexibility compared to standard option pricing ones. Based on Fast Fourier Transform, the algorithm is an enhanced version of the algorithm of Caverhill and Clewlow (1992). The contribution of this chapter is to improve their algorithm by a systematic recentering at each stage and to adapt it to non-lognormal densities. This enables us to examine the impact of fat-tailed distributions on price as well as on the delta. We find evidence that fat-tailed densities lead to wider jumps in the delta. We then examine the case of the Greeks for continuous Asian options and show how to extend the work of the first part to this case. The Malliavin weighted scheme turns out to be adaptable to this particular case. We conjecture indeed that these results should be adaptable to the case of the continuous lookback options.

The last part concentrates on the old but still very interesting problem of the convexity adjustment. We first introduce the notion of convexity. We show that old results of Brotherton-Ratcliffe and Iben (1993) and later by Hull (1997) and Hart (1997) can be consistently derived in a martingale framework. The motivation of this chapter lies in two directions. First, we set up a proper no-arbitrage framework illustrated by a relationship between yield rate drift and bond price. Second, making an approximation, we come to a closed formula with a specification of the error term. Earlier works (Brotherton et al. (1993) and Hull (1997)) assumed constant volatility and could not specify the approximation error. As an application, we examine the convexity bias between CMS and forward swap rates. However, for more complicated term structures assumptions, this approach does not hold any more. The contribution of the last chapter is precisely to provide a solution to this problem and to give good approximation formulae of the convexity adjustment for multi-factor lognormal zero coupon models, which are more general term structure yield curve models. We show how Wiener chaos theory enables us to derive a closed form solution. We apply results to various well-known one-factor models (Ho and Lee (1986), Amin and Jarrow(1992), Hull and White (1990), Mercurio and Moraleda(1996)).

Quasi Monte-Carlo simulations confirm the efficiency of the approximation. Its precision relies on the importance of second and higher order terms.

Brief Review of the literature about Malliavin calculus and Wiener Chaos

Traditionally, option pricing literature is divided into different fields depending on the background of the authors. Option pricing theory can be developed either from a probabilistic referred to as the martingale point of view (theory initiated by Bachelier (1900) for the former, Harrison, Kreps (1979) and Harrison, Pliska (1981) and El-Karoui et al. (1995) for the most famous articles), a partial differential equation one (with a stress on finite difference methods introduced in finance by Schwartz (1977) for the explicit scheme, Brennan and Schwartz (1978) for the implicit one and by Courtadon (1982) for the Crank Nicholson one), a lattice-based concern (Sharpe (1978), Cox Ross and Rubinstein (1979)) or a Monte Carlo simulation emphasis (Boyle (1977) and later Broadie and Glasserman (1996)). Indeed, over the last few years, it has turned out that these fields are not that different. Many "bridges" like the Feynman Kac formula tie the aforementioned fields. The seminal Black Scholes (1973) option pricing formula has been derived by various techniques. Moreover, lattice methods can be seen as a particular case of finite difference methods (Hull (1997)). A stochastic differential equation can be translated into a partial differential equation and the Black Scholes equation can be shown to be a modified version of the heat equation (see Wilmott et al. (1993)). Our aim in this Thesis has been to further integrate and link the various approach. We have tried to see how new developments of the probabilistic theory, that is to say Malliavin calculus and Wiener chaos could help solve option pricing problems.

Indeed, the Malliavin calculus and the Wiener chaos theory have turned out to be very powerful tools for various problems modelled by continuous-time stochastic processes. Moreover, these two theories are nowadays taught together since the Malliavin derivative can be expressed in terms of its Wiener chaos expansion.

The starting point of both Wiener chaos and Malliavin calculus lies in some mathematical considerations. The Wiener chaos theory was initially used to get a Hilbert basis of square integrable function expressed as an Ito integral. It intuitively relates the Cameron Martin sub-space with the wider Hilbert space of square integrable functions expressed as an Ito integral. In particular situations, this expansion could be intuitively thought of the generalization of Taylor's expansion to stochastic processes with some martingale considerations. This representation of stochastic processes initially proved for the Brownian motion by Wiener (1938) and later for Levy process (see Ito 1956) has been recently refocussed, motivated by the contemporary development of the Malliavin calculus theory and its application not only to probability theory but also to mechanics, economics and finance (1995).

The Malliavin calculus was initiated by Malliavin and further developed by Stroock, Bismut, Watanabe and others. The original motivation was to provide a probabilistic proof of Hormander's sum of squares theorem. One of the important conclusions is the existence of an adjoint operator of the Malliavin derivative called the "Skorohod integral". It has the elegant property to be an extension of the Ito integral for non-adapted process. The great advantage of this theory is also to allow the formulation of regular solutions of stochastic differential equations, in case where the solution is not adapted to the Brownian filtration. One can roughly say that the Malliavin calculus is the calculus of variation in a stochastic framework. Or, comparing with the deterministic framework, Ito calculus would correspond to the ordinary derivative in infinitesimal calculus and the Malliavin derivative on Wiener Space to the Frechet derivative on a function space. As an introduction to Malliavin calculus, we suggest the reader to refer to the appendix section A.

Interestingly, these two techniques have turned out to be very useful for many fields. There has been a growing literature on the use of Malliavin calculus theory as well as Wiener Chaos over the last ten years. Uwe et al. (1998) have applied Malliavin calculus to quantum mechanics. They showed that certain types of quantum stochastic processes could be defined by means of their Wigner densities on the Eisenberg-Weyl algebra and that they had to satisfy a diffusion equation. Using the integration by parts formula of the Malliavin calculus, they proved the existence and regularity of these solutions. They applied this theory to the phenomenon of creation and annihilation on Fock space.

Furthermore, going over to the modelling of perturbation for high-frequency telecommunication queuing networks, Decreusefond (1994) applied successfully Malliavin calculus to get high order derivatives as an input to the likelihood Ratio Method. In his model, fluxes are modelled by stochastic processes. Therefore, techniques about stochastic processes are valid. This enabled him to find estimators with faster convergence and to specify a criterion for the absolute continuity for the law of a reflected process.

Bally and Talay (1995) and later Kohatsu and Antonelli (1999) have used Malliavin calculus to study the convergence rate of an appropriate discretisation scheme on the solution to the McKean Vlasov equation, describing the behaviour of a high-density gas. Millet and Solé (1997) used Malliavin calculus to prove the regularity and the smoothness of the law of the solution of a stochastic wave equation in two dimensions.

In economics and finance, Malliavin calculus has been introduced in many works. Serrat (1996) has used the Malliavin derivatives in his model of dynamic equilibrium for two-country exchange economy with non-traded goods and complete financial markets. Øksendal (1997) was among the first ones to suggest the use of Malliavin calculus in economics and especially in the option pricing literature. Bermin (1998) (1999) has suggested an alternative approach to delta hedging by means of the Malliavin calculus. He has examined the complicated case of hedging strategies of barrier options. Recently, Gobet (2000) has used Malliavin calculus theory to study the convergence rate of killed diffusions using Euler schemes and has applied it to barrier options. Last but not least, Fournié et al. (1999) (2000) have suggested the use of Malliavin calculus for the faster computation of the Greeks.

The same is true for Wiener chaos theory. The study of chaos expansions and multiple Wiener-Ito integrals has become a field of considerable interest in applied and theoretical areas of probability, stochastic processes, mathematical physics, and statistics. It has been used in filtering theory (Rozovskii (1997)) stochastic physics, biological cybernetics (Johanessma and Victor (1986)) and pattern recognition. In Finance, it has been used by Lacoste (1996) to provide a probabilistic framework for transaction costs, by Brace and Musiela (1995) to find an approximation formula for interest rates derivatives and by Barucci and Mancino (1997) for a model similar to the one of Lacoste focussing on transaction costs and hedging problems.

Notation

General Notation

$\mathbb{R}$	Set of real numbers
$\mathbb{R}^+$	Set of real non-negative numbers
A^{-1}	Inverse of a number, a matrix, according to the definition of A
$log(x)$	Natural logarithm (Neper basis)
$f'(x)$	First order derivative function with respect to x
$f''(x)$	Second order derivative function with respect to x
$O(\varepsilon)$	Notation of Landau defined by there exists η so that it is bounded in absolute value by $\eta\varepsilon$
$L^2([0,T])^d$	Real separable Hilbert space of d -dimensional real functions squared integrable, defined on $[0,T]$
$\langle .,. \rangle$	Canonical scalar product of $L^2([0,T])^d$
$\Vert\Vert$	Canonical norm of $L^2([0,T])^d$
C_K^∞	Set of infinitely differentiable functions with compact support
$C_p^\infty(.)$	Set of infinitely differentiable functions with all their partial derivatives with polynomial growth
$C_b^\infty(.)$	Set of infinitely differentiable functions with their all partial derivatives bounded
$\delta_{n,m}$	Kronecker delta defined by $\delta_{n,m} = 1$ if $n = m$, $\delta_{n,m} = 0$ otherwise
1_A	$1_A(x) = 1$ if $x \in A$ and 0 otherwise
$\equiv$	Equals by definition

Particular sets

C_n	Set of strictly increasingly-ordered n-uplets $\{(s_1, ..., s_n) \in \mathbb{R}^n, 0 < s_1 < ... < s_n < t\}$
T_m	Set of $L^2[0,T]$ normalized functions with respect to the Lebesge measure under $[0,t_i]$

	defined by $\left\{a \in L^2[0,T] \mid \int_0^{t_i} a(t)\,dt = 1 \;\forall i = 1...m\right\}$
$\widetilde{T}_m$	Set of $L^2[0,T]$ normalized functions
	with respect to the Lebesge measure under $[t_{i-1}, t_i]$
	defined by $\left\{\widetilde{a} \in L^2[0,T] \mid \int_{t_{i-1}}^{t_i} \widetilde{a}(t)\,dt = 1 \;\forall i = 1...m\right\}$

Probability

(Ω, F, Q)	Complete probability space
$\{F_t, t \in [0,T]\}$	Augmented filtration generated by
	a standard Wiener Process $(W_t)_{t\in\mathbb{R}}$
$a.s.$	almost surely
$Var(X)$	Standard variance defined as $\mathbb{E}(X^2) - (\mathbb{E}(X))^2$
$Cov(X,Y)$	Standard covariance defined as $\mathbb{E}(XY)) - \mathbb{E}(X)\mathbb{E}(Y)$
P	Historical probability measure
Q	Risk neutral probability measure
Q^T	Forward neutral probability measure
$\mathbb{E}^Q[.]$	Expectation under the probability measure Q
$\mathbb{E}^Q_x[.]$	$\mathbb{E}^Q_x[.] = \mathbb{E}^Q[.\vert X_0 = x]$
$\mathbb{E}^Q_{x,X_{t_1},\ldots,X_{t_m}}[.]$	$\mathbb{E}^Q_{x,X_{t_1},\ldots,X_{t_m}}[.] = \mathbb{E}^Q_x[.\vert X_{t_1},\ldots,X_{t_m}]$

Stochastic Processes

$(W_t)_{t\in\mathbb{R}}$	Wiener process, either one dimensional
	or multi dimensional
$(M_t)_{t\in\mathbb{R}}$	Square-integrable martingale
	with respect to an appropriate filtration $\{F_t\}_{t\in\mathbb{R}}$
$\langle M\rangle_t$	Doob Meyer brackets defined through the
	requirement that $(M_t^2 - \langle M\rangle_t)$ be a martingale
$(\Phi_n)_{n\in\mathcal{N}}$	Morphism from $\mathcal{L}^2(C_n)$ to $\mathcal{L}^2(\mathcal{F}_\infty)$ defined by
	$\Phi_n(f) : \mathcal{L}^2(C_n) \to \mathcal{L}^2(\mathcal{F}_\infty)$
	$\Phi_n(f) = \int_{0\le s_1\le..s_i\le...s_n\le T} f(s_1,\ldots,s_n)\,dM_{sn}\,dM_{s_1}$

$\mathcal{S}$	Set of stochastic functions F of the form $f(\int_0^T h_1(t)dW_t, ..., \int_0^T h_n(t)dW_t)$
$\mathcal{D}^{1,2}$	Banach space, completion of $\mathcal{S}$ with respect to the norm $\| F \|_{1,2} = (\mathbb{E}(F^2))^{1/2} + (\mathbb{E}(\int_0^T (D_t F)^2 dt)^{1/2}$
$I_{i,i>0}(V,T,T_k)$	Wiener Chaos of order i spanned by the function $V(.,T,T_k)$ at time T defined by $\frac{\int_{0\leq s_1\leq ..s_i\leq ...s_n\leq T}\langle V(s_1,T,T_k),dW_{s_1}\rangle ... \langle V(s_i,T,T_k),dW_{s_i}\rangle}{i!}$

Malliavin Calculus

D_tF	Malliavin derivative
$\delta(.)$	Skorohod integral
$weight$	Malliavin weight
w^{delta}	Malliavin weight for the delta
w^{gamma}	Malliavin weight for the gamma
w^{rho}	Malliavin weight for the rho
w^{vega}	Malliavin weight for the vega
π_0	Malliavin weight with minimal variance

Underlying

X_t	Price of the underlying security at time t
x	Initial value of the underlying security at time $t=0$
R_t	Log return defined as $R_{t_i} = log(X_{t_i}/X_{t_{i-1}})$
r or r_s	Risk-free rate
σ	Constant Black Scholes volatility
σ_s	Black volatility (deterministic, time dependent)
$b(t,X_t)$	Drift term
$\sigma(t,X_t)$	Diffusion term

Perturbation on the Underlying

Y_t	First variation process
$b'(t, X_t)$	Partial derivative of the drift term with respect to the second variable
$\sigma'(t, X_t)$	Partial derivative of the diffusion term with respect to the second variable
$\widetilde{b}(t, X_t)$	Perturbation of the drift term
$\widetilde{\sigma}(t, X_t)$	Perturbation of the diffusion term
$X_t^{\varepsilon, rho}$	Perturbation of the underlying along the drift term
$X_t^{\varepsilon, vega}$	Perturbation of the underlying along the diffusion term
Z_t^{rho}	Sensitivity of the underlying along the perturbed drift term
Z_t^{vega}	Sensitivity of the underlying along the perturbed diffusion term

Payoff and Price

$f(X_T)$	Payoff of an option
$f(X_{t_1}, .., X_{t_m})$	Payoff of an option depending on a finite set of dates $t_1...t_m$
$f\left(\int_0^T X_s ds\right)$	Payoff of an option depending on the continuous arithmetic average (simple Asian option)
$f\left(X_T, \int_0^T X_s ds\right)$	Payoff of an option depending on the terminal value of the underlying security as well as the continuous arithmetic average (complex Asian option)
$F = e^{-\int_0^T r_s ds} f(X_t)$	Discounted payoff
$P(x)$	Price of the option as a function of the underlying level

Interest Rate

$e^{-\int_0^T r_s ds}$	Discount factor
$B(t,T)_{t<\tau, T<\tau}$	Price at time t of a default-free forward zero coupon maturing at time T
Y^f	Forward swap rate
B_{T_i}	Value at time $t=0$ of the T-forward zero coupon maturing at time T_i defined as $\frac{B(0,T_i)}{B(0,T)}$
CA	Convexity adjustment between different products depending on the context
K	Sensitivity of the forward swap defined as the sum of the forward zero coupon bonds $K=\sum_{i=1}^n B_{T_i}$.
$V(s,T_i)$	Instantaneous volatility at time s of a forward zero coupon maturing at time T_i
$V_s^{(T,T_i)}$	Forward volatility of a T-forward zero coupon maturing at time T_i defined as $V(s,T_i)-V(s,T)$
$C(T_i,T_j)$	Correlation term between the returns of T-forward zero coupon bond defined as $\int_0^T \left\langle V_s^{(T,T_i)}, V_s^{(T,T_j)} \right\rangle$

Chapter 1

Malliavin Weighted Scheme for Fast Computation of the Greeks

Summary of the chapter

This chapter presents a new method to speed up the convergence of Monte-Carlo and Quasi-Monte Carlo simulations of the Greeks by means of the Malliavin weighted scheme. The contribution to the pioneering works is to derive necessary and sufficient conditions for a function to serve as a weight function and to find the weight function with the minimum variance. To do so, we introduce the generator of the weighting function defined as its Skorohod integrand. This new definition turns out to be very powerful since it provides a description of all weight functions. An integration by parts, by means of Malliavin calculus, leads to these conditions. These conditions are expressed through conditional expectations. We show that the minimum-variance solution is the projection of any weight function on the filtration spanned by the payoff functional. We give some key examples of the weight function generator. For complicated diffusion, the optimal solution is not easy to calculate explicitly. We discuss the question of the most appropriate weight function. We finally conclude that this method is very efficient for discontinuous payoff options, like binary and corridor options.

1.1 Introduction

Since price sensitivities are an important measure of risk, growing emphasis on risk management issues has suggested a greater need for their efficient computation. Collectively referred to as "the Greeks", theses sensitivities are mathematically defined as the derivatives of a derivative security's price with respect to various model parameters.

The traditional way to compute a Greek is through its finite difference approximation. If we denote by $P(x)$ the price of the option for an underlying's initial value equal to x, one calculates the delta by means of $(P(x+\varepsilon) - P(x))/\varepsilon$. This can produce a significant error since one takes the difference of terms $P(x)$ and $P(x+\varepsilon)$ which are already calculated by approximations. When looking at Monte Carlo and Quasi Monte Carlo methods, Glynn (1989) showed that the quality of this approximation was depending on the way of approximating the derivative: forward difference $(P(x+\varepsilon) - P(x))/\varepsilon$, central difference $(P(x+\varepsilon) - P(x-\varepsilon))/2\varepsilon$, or even backward difference scheme $(P(x) - P(x-\varepsilon))/\varepsilon$. In the case of the forward and backward difference scheme, if the simulation of the two estimators of $P(x+\epsilon)$ and $P(x)$ or $P(x)$ and $P(x-\varepsilon)$ is drawn independently, he proved that the best theoretical convergence rate is $n^{-1/4}$. As of the central difference scheme, the optimal rate is $n^{-1/3}$. When taking common random numbers, this optimal rate becomes $n^{-1/2}$. This is the best to be expected by standard Monte Carlo simulation as described by Boyle, Broadie and Glasserman (1997) Glasserman and Yao (1992), Glynn (1989), and L'Ecuyer and Perron (1994). However, the finite difference method has a slow convergence rate when dealing with discontinuous payoffs. This restriction applies to many of the exotic options such as digital, corridor, Asian and lookback options.

To overcome this poor convergence rate, Curran (1994), (1998) and Broadie and Glasserman (1996) suggested to take the differential of the payoff function inside the expectation required to compute a price. This leads to a convergence rate of $n^{-1/2}$. However, this can be applied only to simple payoff functions.

Fournié et al. (1999) (2000) extended their method to payoffs depending on a finite set of dates, in very general conditions. The original idea comes from a result by Elworthy (1992) which suggests, in a probabilistic framework, to shift the differential operator from the payoff functional to the diffusion kernel, introducing a weighting function. They came to the central result that the common Greeks could be written as an expected value of the payoff times a weight function.

$$Greek = \mathbb{E}^{Q}\left[e^{-\int_0^T r_s ds} f(X_T).weight\right] \tag{1.1}$$

The theoretical tool used was the stochastic calculus of variations, traditionally called Malliavin calculus. Their results were given for particular examples of weight functions. However, a natural question, starting point of this research was to examine all the weight functions and to determine which conditions a weight function should satisfy. Another important question is to find the minimal-variance solution.

The contribution of this chapter is precisely to answer these questions. We show how to characterize by necessary and sufficient conditions the weight functions in the Malliavin weighted scheme. Expressing weight functions as Skorohod integral, we introduce the weight function generator defined as the Skorohod integrand. We show that these functions can be characterized by necessary and sufficient conditions on their generator. We then examine the different weight functions and show how to find the one with minimal variance. We then give some key examples of the weighting function generator. We finally discuss the issue of the most appropriate weight function.

The remainder of this chapter is organized as follows. In section 2, we explicit the intuition of the methodology with the Black Scholes model as well as some preliminary definitions and results. In section 3, we derive the necessary and sufficient conditions for the weight function generator. In section 4, we show various examples for the weight function generator. We conclude in section 5. For clarity reasons, we put all the proofs which turned to be quite involved, in the appendix section.

1.2 Mathematical framework and preliminary results

1.2.1 Intuition

In this subsection, we show by means of the Black Scholes (1973) model, how to derive a formula that reduces the variance of the Greeks when computed by simulation methods. The core of our methodology lies in an integration by parts formula. This allows us to avoid taking the derivative of the payoff functional and instead shift the differential operator on the diffusion kernel.

Following Harrison and Kreps (1979), Harrison and Pliska (1981), the price of a contingent claim is traditionally calculated as the expected value of the discounted payoff value in the risk neutral probability measure Q uniquely defined in complete markets with no-arbitrage. We consider a continuous-time trading economy with a finite horizon $t \in [0,T]$. The uncertainty in this economy is classically modelized by a complete probability space (Ω, F, Q). The information evolves according to the augmented filtration $\{F_t, t \in [0,T]\}$ generated by a standard one dimensional standard Wiener process $(W_t)_{t\in[0,T]}$. The price $P(x)$ of our contingent claims at time $t = 0$ with expiry date T is defined by the expected value of the discounted payoff function at expiry $f(X_T)$ (for a call $f(X_T) = (X_T - K)_+$) conditional to the present information, described by σ-algebra $F_{t=0}$

$$P(x) = \mathbb{E}^Q\left[f(X_T)\, e^{-\int_0^T r_s ds} | F_0\right] \tag{1.2}$$

$\mathbb{E}^Q[.]$ is the expectation under the risk neutral measure Q, X_t is the underlying price, and r_s is the risk-free rate. Following Black Scholes assumptions, the underlying, either an equity, a commodity, an interest rate or an index price, follows a geometric Brownian motion characterized by the following diffusion equation:

$$\frac{dX_t}{X_t} = rdt + \sigma dW_t \tag{1.3}$$

Let us denote by X_T the unique continuous strong solution of (1.3) with initial condition x ($X_0 = x$). Replacing in (1.2) X_T by its probability density function gives us that the price $P(x)$ can be written as an explicit integral:

$$P(x) = \int_{-\infty}^{+\infty} e^{-rT} f(xe^{rT+\sigma\sqrt{T}y-\frac{1}{2}\sigma^2 T}) \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy$$

To calculate a Greek, traditional methods compute numerically the finite difference between two shifted prices. For the clarity of the proof, we chose the delta. This leads in the case of a centered scheme to:

$$delta \simeq \frac{P\left(x+\frac{\varepsilon}{2}\right) - P\left(x-\frac{\varepsilon}{2}\right)}{\varepsilon}$$

Its continuous limit leads then to take the derivative of the payoff function since the expression $\frac{f\left(X_T^{x+\frac{\varepsilon}{2}}\right) - f\left(X_T^{x-\frac{\varepsilon}{2}}\right)}{\varepsilon}$ inside the expectation operator in (1.4) tends to the derivative of the function f as ε tends to zero.

$$\frac{P\left(x+\frac{\varepsilon}{2}\right) - P\left(x-\frac{\varepsilon}{2}\right)}{\varepsilon} = \mathbb{E}^Q\left[e^{-\int_0^T r_s ds} \frac{f\left(X_T^{x+\frac{\varepsilon}{2}}\right) - f\left(X_T^{x-\frac{\varepsilon}{2}}\right)}{\varepsilon} | F_0\right] \tag{1.4}$$

The driving idea of this chapter is to avoid taking the derivative of the function, by doing an integration by parts. Assuming that $f(.)$ is a.s. differentiable with derivatives with polynomial growth[1], we can show that the derivative with respect to x is proportional to the derivative with respect to y:

$$\frac{\partial}{\partial x} f(xe^{rT+\sigma\sqrt{T}y-\frac{1}{2}\sigma^2 T}) = \frac{1}{x\sigma\sqrt{T}} \frac{\partial}{\partial y} f(xe^{rT+\sigma\sqrt{T}y-\frac{1}{2}\sigma^2 T})$$

[1]These are assumptions that justify the interchange of the integration and the differential operator by dominated convergence.

leading to the following integration by parts:

$$\begin{aligned}\frac{\partial}{\partial x}P &= \frac{\partial}{\partial x}\left(\int_{-\infty}^{+\infty} e^{-rT} f(xe^{rT+\sigma\sqrt{T}y-\frac{1}{2}\sigma^2 T})\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy\right)\\ &= e^{-rT}\left[\frac{1}{x\sigma\sqrt{T}} f(xe^{rT+\sigma\sqrt{T}y-\frac{1}{2}\sigma^2 T})\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}\right]_{-\infty}^{+\infty}\\ &\quad + e^{-rT}\int_{-\infty}^{+\infty}\frac{1}{x\sigma\sqrt{T}} f(xe^{rT+\sigma\sqrt{T}y-\frac{1}{2}\sigma^2 T})\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}y\,dy\end{aligned}$$

This enables us to write the delta as the expectation of the discounted payoff times a weight function:

$$\frac{\partial P}{\partial x} = \mathbb{E}^Q\left(\frac{e^{-rT}}{x\sigma T}W_T f(X_T)\right)$$

In the above formula, the differential operator has disappeared. Instead, this methodology has introduced a weight function $\frac{e^{-rT}}{x\sigma T}W_T$. The weight is not depending on the payoff function and is easy to simulate. This indicates that the efficiency of this method does not depend on the payoff type. On the contrary, the standard way to compute the Greeks relies on the payoff function since it takes the finite difference approximation of the derivative of the payoff function (1.4). Since this integration by parts method smoothens the payoff function with a weight independent of the payoff function, it is all the more efficient that the payoff function is discontinuous. This is the case of digital, simple, double barrier and many other exotic options. Furthermore, we can conjecture that this method should be more efficient for second order Greeks, like gamma, than first order ones, like delta. Moreover, this methodology should provide us with similar rates of convergence for the Greeks as for the price. The only difference between the price simulation and the Greek simulation comes from the weight function to simulate.

1.2.2 Notations and hypotheses

To avoid heavy notations, and for clarity reason, we present our results in one dimension. However, the results can be easily extended to the multi-dimensional case. Following the traditional literature on continuous-time option pricing (see Duffie (1995), Musiela and Rutkowski (1997) or Lamberton and Lapeyre (1991)) the evolution of the underlying price, Ito process $(X_t)_{t\in[0,T]}$, is described by a very general stochastic differential equation (SDE):

$$dX_t = b(t, X_t)\, dt + \sigma(t, X_t) dW_t \tag{1.5}$$

with the initial condition $X_0 = x$, $x \in \mathbb{R}$. The function $b : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ represents the determinist drift of our process and the function $\sigma : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ its volatility. The risk-free interest rate is denoted by $r(t, X_t)$. We assume that:

- the functions b and σ are continuously differentiable with bounded derivatives and verify Lipschitz conditions, i.e., there exists a constant $K < +\infty$ such that

$$\begin{aligned} |b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| &\leq K|x-y| && (1.6)\\ |b(t,x)| + |\sigma(t,y)| &\leq K(1+|x|) && (1.7) \end{aligned}$$

Inequalities (1.6) and (1.7) are classical conditions to ensure the existence and uniqueness of a continuous, strong solution of the SDE (1.5) with its initial condition. We denote by X_t^x the continuous, strong solution X_t starting at x.

- the diffusion function $\sigma(t,x)$ is uniformly elliptic[2]:

$$\exists \epsilon > 0, \quad \forall t \in [0,T], \forall x \in \mathbb{R} \quad |\sigma(t,x)| \geq \epsilon |x| \tag{1.8}$$

We denote by $(Y_t)_{t\in[0,T]}$ the first variation process of $(X_t)_{t\in[0,T]}$, which is characterized as the unique strong continuous solution of the linear stochastic differential equation (1.9) with initial condition $(Y_{t=0} = 1)$:

$$\frac{dY_t}{Y_t} = b'(t, X_t)dt + \sigma'(t, X_t)dW_t \tag{1.9}$$

[2]This is to ensure that we can find some solutions for the weighting functions, since it often requires to take the inverse of the volatility function.

where the prime stands for the derivatives with respect to the second variable. We can show that the first variation process is the derivative of $(X_t)_{t\in[0,T]}$ with respect to x, $\left(Y_t = \frac{\partial}{\partial x}X_t\right)$. Malliavin calculus theory proves (see Nualart (1995) page Theorem 2.3.1 page 110 on the absolute continuity) that the Malliavin derivative can be written as an expression of the first variation process as well as the volatility function:

$$D_s X_t = Y_t Y_s^{-1}\sigma(s, X_s)1_{\{s\leq t\}}a.s. \tag{1.10}$$

To be as general as possible, we assume that our payoff is depending on a finite set of payment dates: $t_1, t_2, ..., t_m$ with the convention that $t_0 = 0$ and $t_m = T$. The price $P(x)$ of the contingent claim given an initial value of the underlying price x is traditionally computed as the expectation under the risk neutral probability measure of discounted future cash flow:

$$P(x) = \mathbb{E}_x^Q\left[e^{-\int_0^T r(s,X_s)ds} f\left(X_{t_1}, X_{t_2}, ..., X_{t_m}\right)\right]$$

with the traditional shortcut notation $\mathbb{E}_x^Q[.] = \mathbb{E}^Q[.|X_0 = x]$. The function $f : \mathbb{R}\times\mathbb{R}\times...\times\mathbb{R} \rightarrow \mathbb{R}$ denotes the payoff, and is supposed to be first order differentiable with derivatives with at most polynomial growth. We denote by F the discounted payoff $F = e^{-\int_0^T r(s,X_s)ds} f\left(X_{t_1}, X_{t_2}, ..., X_{t_m}\right)$. If we need to specify that the underlying is a function of the initial value x, we denote the discounted payoff by F^x.

1.2.3 Generalizing Greeks

We take the common definition of the delta and gamma as the first (respectively the second order derivative) of the price with respect to the underlying initial level. However, for the rho and vega, the definitions need to be extended. Since by assumptions, the drift and volatility terms are functions of the underlying and time but not constant coefficient, we need to develop a more robust framework than the common sensitivity with respect to a fixed parameter. The meaning of the rho and vega is to quantify the impact of small perturbation, in a specified direction, on either the drift term or the volatility term. We therefore define an "extended" rho as well as an "extended" vega defined as the derivative function of the price along a specified perturbation direction either on the drift term or the volatility term.

Let denote by $\widetilde{b} : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ a direction function for the drift term and $\widetilde{\sigma} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ for the stochastic term. We assume that, for every $\epsilon \in [-1,1]$, $\widetilde{b}(.,.)$, $\left(b+\varepsilon\widetilde{b}\right)(.,.)$, $\widetilde{\sigma}(.,.)$ and $(\sigma+\varepsilon\widetilde{\sigma})(.,.)$ are continuously differentiable with bounded derivatives and verify Lipschitz conditions and moreover that $\widetilde{\sigma}(.,.)$ and $(\sigma+\varepsilon\widetilde{\sigma})(.,.)$ satisfy the uniform ellipticity condition (1.11). $\forall\varepsilon \in [-1,1]$

$$\exists\eta > 0, \forall\, x \in \mathbb{R}, \forall t \in [0,T], \;\; |(\sigma+\varepsilon\widetilde{\sigma})(t,x)| \geq \eta\,|x| \tag{1.11}$$

We then define two different perturbed underlying processes, with their respective prices. The "drift-perturbed" process is the stochastic process $\left\{X_t^{\varepsilon,rho}, t \in [0,T]\right\}$[3] solution of the perturbed diffusion equation, in the direction $\widetilde{b}$, defined by (1.12) and the unmodified initial condition ($X_0^{\varepsilon,rho} = x$)

$$dX_t^{\varepsilon,rho} = \left[b\left(t, X_t^{\varepsilon,rho}\right) + \varepsilon\widetilde{b}\left(t, X_t^{\varepsilon,rho}\right)\right]dt + \sigma\left(t, X_t^{\varepsilon,rho}\right)dW_t \tag{1.12}$$

Similarly, the volatility-perturbed underlying process is the stochastic process $\{X_t^{\varepsilon,vega}, t \in [0,T]\}$ solution of the perturbed diffusion equation in the direction $\widetilde{\sigma}$ defined by (1.13) and the unmodified initial condition ($X_0^{\varepsilon,vega} = x$)

$$dX_t^{\varepsilon,vega} = b\left(t, X_t^{\varepsilon,vega}\right)dt + \left[\sigma\left(t, X_t^{\varepsilon,vega}\right) + \varepsilon\widetilde{\sigma}\left(t, X_t^{\varepsilon,vega}\right)\right]dW_t \tag{1.13}$$

We can relate two perturbed prices to these two perturbed processes: $P_{rho}^{\varepsilon}(x)$ and $P_{vega}^{\varepsilon}(x)$ defined by

$$\begin{aligned} P_i^{\varepsilon}(x) &= \mathbb{E}_x^Q\left[e^{-\int_0^T r\left(s, X_s^{\varepsilon,i}\right)ds} f\left(X_{t_1}^{\varepsilon,i}, X_{t_2}^{\varepsilon,i}, ..., X_{t_n}^{\varepsilon,i}\right)\right] \\ \text{with } i &= \text{rho or vega} \end{aligned}$$

[3] we put either *rho* or *vega* in upperscript so as to be able to distinguish the two perturbed process $X_t^{\varepsilon,rho}$ and $X_t^{\varepsilon,vega}$. One is corresponding to a perturbation on the drift term whereas the other one on the stochastic term..

The physical meaning of the above definitions is to set an appropriate framework so as to see the impact on the underlying process as well as on the price function of a structural change of either the drift or the volatility term. The extended rho and vega quantify this effect. They are given by the following definitions:

Definition 1 *The extended rho is the Gateau derivative of the perturbed price function $P^{\varepsilon}_{rho}(x)$ in the direction given by the function $\widetilde{b}(.)$:*

$$rho \equiv \frac{\partial}{\partial \varepsilon} P^{\varepsilon}_{rho}(x)\bigg|_{\varepsilon=0,\widetilde{b}\ given} \tag{1.14}$$

where the sign $\equiv$ stands for a definition Similarly, the extended vega is the Gateau derivative of the perturbed price function $P^{\varepsilon}_{vega}(x)$ in the direction given by the function $\widetilde{\sigma}(.,.)$:

$$vega \equiv \frac{\partial}{\partial \varepsilon} P^{\varepsilon}(x)\bigg|_{\varepsilon=0,\widetilde{\sigma}\ given} \tag{1.15}$$

1.2.4 Results on the first variation process

This section shows that the first variation process $(Y_t)_{t\in[0,T]}$ is at the core of the extended Greeks theory. In this section, we introduce Gateau derivatives implied by the extended Greeks. We show that these two Gateau derivatives can be expressed as a simple function of the first variation process Y_t. We denote by $\left(Z^{rho}_t\right)_{t\in[0,T]}$ (respectively $(Z^{vega}_t)_{t\in[0,T]}$) the Gateau derivative of the drift-perturbed underlying process $\left\{X^{\varepsilon,rho}_t, t\in[0,T]\right\}$, respectively the volatility-perturbed underlying process $\{X^{\varepsilon,vega}_t, t\in[0,T]\}$ along the direction $\widetilde{b}$, respectively $\widetilde{\sigma}$. These two quantities are defined as the limit in L^2, uniformly with respect to time t :

$$Z^{rho}_t = \lim_{L^2,\epsilon\to 0} \frac{X^{\varepsilon,rho}_t - X_t}{\varepsilon} \tag{1.16}$$

respectively

$$Z^{vega}_t = \lim_{L^2,\epsilon\to 0} \frac{X^{\varepsilon,vega}_t - X_t}{\varepsilon} \tag{1.17}$$

Interestingly, these two processes can be expressed in terms of the first variation process $(Y_t)_{t\in[0,T]}$ as the following proposition states:

Proposition 1 *The process* $\left(Z_t^{rho}\right)_{t\in[0,T]}$ *can be expressed in terms of the first variation process by*

$$Z_t^{rho} = \int_{s=0}^{t} \frac{Y_t \widetilde{b}(s, X_s)}{Y_s} ds \tag{1.18}$$

Similarly, the process $(Z_t^{vega})_{t\in[0,T]}$ *can be expressed in terms of the first variation process* $(Y_t)_{t\in[0,T]}$ *by*

$$Z_t^{vega} = \int_0^t Y_t \frac{\widetilde{\sigma}(s, X_s)}{Y_s} dW_s - \int_0^t Y_t \sigma'(s, X_s) \frac{\widetilde{\sigma}(s, X_s)}{Y_s} ds \tag{1.19}$$

Proof: see the appendix section, section B.1, page 158.□

The proposition above explains intuitively why the Malliavin weights for the rho and vega can be expressed in terms of the first variation process. The difference between the volatility-perturbed framework and the drift-perturbed one comes from an additional term in the case of the volatility-perturbed one.

1.3 Malliavin weighted scheme: a new method for computing the Greeks

This section shows the necessary and sufficient conditions for a function to serve as a weight function. We first give the state of the art, then give the necessary and sufficient conditions and finally show how to extend these conditions to models where the risk-free interest rate is a function of the underlying as in interest rates models for the spot rate (model of Vasicek (1977), Cox Ingersoll Ross (1981), Black Derman Toy (1990), Black Karinski (1990) and so on).

1.3.1 State of art

Fournié et al. (1999) and (2000) were the first to suggest that the three Greeks delta, vega and rho could be computed as an expected value of the discounted payoff times a suitable weight function (1.20)

$$Greek = \mathbb{E}^Q \left[e^{-\int_0^T r_s ds} f(X_T) weight \right] \tag{1.20}$$

The article of Fournié et al. (1999)leaves many questions unanswered, like which condition(s) a function should satisfy to serve as a weight function and which weight function is the one with minimal variance.

It is worth noticing that all weight functions could be expressed as a Skorohod integral. This is because the Skorohod integral is the adjoint operator of the Malliavin derivative. It means that the only way to have an integration by parts by means of the Malliavin derivatives is that there exists a weight function that could be written as a Skorohod integral. The following subsections shows that the weight function generator (defined below) should satisfy necessary and sufficient conditions. Interestingly, these conditions are different for each Greek but independent of the payoff function. Therefore, the Malliavin weight is independent from the payoff function.

1.3.2 Generalization of the methods: Exact determination of the Malliavin Weights

Writing the weight function *weight* as a Skorohod integral, we call weight function generator w the Skorohod integrand

$$weight = \delta(w) \tag{1.21}$$

We will assume as well that the weight is L^2 integrable that is

$$\mathbb{E}\left[weight^2\right]^{1/2} < \infty \tag{1.22}$$

This equation is the condition to ensure the existence of the Skorohod integral. Since the Skorohod integral is at the core of the Malliavin integration by parts formula, the weight function is better characterized by its weight function generator. We first examine the most common case where we assume that the instantaneous risk-free interest rate does not depend on the underlying process $r'(s, X_s) = 0$ where the prime stands for the derivative function with respect to the second variable. Denoting by $\mathbb{E}^Q_{x,X_{t_1},\dots,X_{t_m}}$ the conditional expectation with respect to $X_{t_1}, \dots, X_{t_m}$, i.e. $\mathbb{E}^Q_{x,X_{t_1},\dots,X_{t_m}}[.] = \mathbb{E}^Q_x[|X_{t_1}, \dots, X_{t_m}]$, we show that:

Greeks	Necessary and Sufficient conditions on the Malliavin Weights
delta	(1.M1) : $\mathbb{E}^{Q}_{x,X_{t_1},\ldots,X_{t_m}}\left[Y_{t_i}\int_0^{t_i}\frac{\sigma(t,X_t)}{Y_t}w^{delta}(t)\,dt\right] = \mathbb{E}^{Q}_{x,X_{t_1},\ldots,X_{t_m}}\left[Y_{t_i}\right]$
gamma	(1.M2) : $\mathbb{E}^{Q}_{x,X_{t_1},\ldots,X_{t_m}}\left[\delta\left(w^{gamma}\right)\right] = \mathbb{E}^{Q}_{x,X_{t_1},\ldots,X_{t_m}}\left[\delta\left(w^{delta}\delta\left(w^{delta}\right)+\frac{\partial}{\partial x}w^{delta}\right)\right]$
"extended" rho	(1.M3) : $\mathbb{E}^{Q}_{x,X_{t_1},\ldots,X_{t_m}}\left[Y_{t_i}\int_0^{t_i}\frac{\sigma(t,X_t)}{Y_t}w^{rho}(t)\,dt\right] = \mathbb{E}^{Q}_{x,X_{t_1},\ldots,X_{t_m}}\left[Y_{t_i}\int_0^{t_i}\frac{\tilde{b}(t,X_t)}{Y_t}dt\right]$
"extended" vega	(1.M4) : $\mathbb{E}^{Q}_{x,X_{t_1},\ldots,X_{t_m}}\left[Y_{t_i}\int_0^{t_i}\frac{\sigma(t,X_t)}{Y_t}w^{vega}(t)\,dt\right] = \mathbb{E}^{Q}_{x,X_{t_1},\ldots,X_{t_m}}\left[\begin{array}{c}\int_0^{t_i}\frac{\tilde{\sigma}(t,X_t)Y_{t_i}}{Y_t}dW_t \\ -\int_0^{t_i}\sigma'(s,X_s)\frac{\tilde{\sigma}(s,X_s)Y_{t_i}}{Y_t}ds\end{array}\right]$

Table 1.1: Necessary and Sufficient conditions for the Weighting Function Generators in a model with interest rates independent of the underlying. The proofs for the equations (1.M1), (1.M2), (1.M3) and (1.M4) are given in the appendix section, respectively in section B.2, B.3, B.4, B.5

Theorem 1 *Malliavin formula for the Greeks*
There exist necessary and sufficient conditions for a function w to serve as a weighting function generator for the simulation of the Greeks. The first condition is the Skorohod integrability of this function. The second condition, different for each Greeks and summarized in table 1.1, is depending only on the underlying diffusion characteristics and is independent from the payoff function.

1.3.3 Extension to models with stochastic interest rates

When we assume that the risk-free interest rate is a function of the underlying, we need to take into account the dependency of the risk-free rate from the underlying process. The necessary and sufficient conditions given in table 1.1 are not sufficient and need to be completed by other conditions. We need to include in the expectation operator the discount factor $e^{-\int_0^T r(s,X_s^x)ds}$, term which is stochastic. This provides a second condition. The second condition is obtained the same way the first one was derived. However, since this expression does not bring any new intuition and is tedious rephrasing of the simpler results of table 1.1, we have put the set of these condition in the appendix section, section B.7, page 175 in table B.2.

1.3.4 The minimal variance weighting problem

If we want the weight function with the minimal variance, we have to understand the way the Greeks are calculated. We have found that the Greeks are expressed as the expectation of a weighting function times the discounted payoff. The only information we have about the payoff function is its measurability with respect to its spanned filtration F_T. It means that the product inside our expectation can be seen as the scalar product of the weighting function with any F_T-measurable function. The projection theorem proves us that the weight function with minimal variance is the conditional expectation of any weight function with respect to the filtration F_T by means of the theorem of projection. More precisely, we have the following proposition

Proposition 2 *The weight function with minimal variance denoted by π_0 is the conditional expectation of any weight function with respect to the filtration F_T*

$$\pi_0 = \mathbb{E}\left[weight|\boldsymbol{F}_T\right]$$

Proof: Let π be a weight function. The Greek ratio can be expressed as the expected value of the scalar product of the discounted payoff, F, with this weight function time $Greek = \mathbb{E}[F.\pi]$. The variance V of this estimator is given by the quadratic variation of our estimator of the Greek minus the true value of the Greek.

$$V = \mathbb{E}\left[(F.\pi - Greek)^2\right]$$

We can introduce the conditional expectation π_0, leading to

$$\begin{aligned} V &= \mathbb{E}\left[(F.(\pi - \pi_0) + F.\pi_0 - Greek)^2\right] \\ &= \mathbb{E}\left[(F.(\pi - \pi_0))^2\right] + \mathbb{E}\left[(F.\pi_0 - Greek)^2\right] \\ &\quad +2\mathbb{E}\left[(F.(\pi - \pi_0)).(F.\pi_0 - Greek)\right] \end{aligned}$$

But indeed the last term in the equation above is equal to zero since

$$\begin{aligned} \mathbb{E}\left[(F.(\pi - \pi_0)).(F.\pi_0 - Greek)\right] &= \mathbb{E}\left[\mathbb{E}\left[(F.(\pi - \pi_0)).(F.\pi_0 - Greek)\,|F_T\right]\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[(F.(\pi - \pi_0))\,|F_T\right].(F.\pi_0 - Greek)\right] \\ &= 0 \end{aligned}$$

where we have used the fact that $(F.\pi_0 - Greek)$ and F are F_T-measurable and therefore $\mathbb{E}\left[(F.(\pi - \pi_0))\,|F_T\right] = 0.\square$

This is a strong result. It indicates that the best weighting function should always be the one that is F_T-measurable. It indicates as well that without any more specification on the payoff function, the variance is lower-bounded by the variance of the particular weight function π_0. This indicates as well that with more information on the payoff function, we can have more efficient estimators. This is the case when for example, we have a payoff function which can be expressed in terms of some particular points of the Brownian motion trajectory. In this case, the best weight function would be the one expressed in terms of these particular points.

1.4 Examples of Malliavin weights

In this section, we give examples of weight functions generator. Instead of using the necessary and sufficient conditions derived above, expressed as an equality

Greeks	**Weighting Function Generators of Founie et al.**
delta	$a(t)\frac{Y_t}{\sigma(t,X_t)}$
"extended" rho	$\frac{1}{\sigma(t,X_t)}\widetilde{b}(t,X_t)$
"extended" vega	$\frac{1}{\sigma(t,X_t)}\widetilde{a}(t)\sum_{i=1}^{m}\left(Z_{t_i}^{vega}-Z_{t_{i-1}}^{vega}\right)1_{\{t_{i-1}\leq t<t_i\}}$

Table 1.2: Summary of Particular Malliavin Weights given by Fournié et al.

of conditional expectations, we look for solutions that satisfy the equality of the two terms inside the expectation. Of course, these conditions are stronger and are only sufficient but not necessary.

We show that the solutions given by Fournié et al. (1999) are particular solutions for generator functions. But we exhibit other solutions. This raises the interesting question of the choice of the weight function generator.

1.4.1 Fournié et al. solutions

Let us define $T_m=\left\{a\in L^2[0,T]\mid\int_0^{t_i}a(t)\,dt=1\;\forall i=1...m\right\}$ and $\widetilde{T}_m=\left\{\widetilde{a}\in L^2[0,T]\mid\int_{t_{i-1}}^{t_i}\widetilde{a}(t)\,dt=1\;\forall i=1...m\right\}$. Rewriting all the weight functions of Founié et al. (1999) as Skorohod integral, we can see immediately that of course these functions satisfies the necessary and sufficient conditions. Indeed, an easy way to check that the conditional expectations of the equations (1.M1), (1.M2), (1.M3) and (1.M4) are equal is to verify that the terms inside the expectations are equal. The table 1.2 summarizes the different weight function generators of Fournié et al.

1.4.2 Other examples

In fact, there many other judicious choices of weight function generators that can be used. We only need to find functions that satisfy the necessary and sufficient conditions and are elements of the Skorohod operator domain denoted by $D^{1,2}$.

Such functions written as piecewise stochastic constant are given below:

$$\begin{aligned} w^{delta}(t) &= \sum_{i=1}^{m} \alpha_i^{delta} 1_{t_{i-1} \leq t \leq t_i} && (1.23) \\ \text{with} \quad \sum_{j=1}^{i} \alpha_j^{delta} \int_{t_{j-1}}^{t_j} \frac{\sigma(t, X_t)}{Y_t} dt &= 1 \quad \forall i = 1...m && (1.24) \end{aligned}$$

It is interesting as well to examine the case of the gamma. However, for some simple diffusion assumption, we can find a proportionality relationship between gamma and vega. This is when the first variation process is proportional to the underlying. This implies that the underlying process follows a Geometric Brownian motion (see Benhamou (2000c)). For a more general model, the calculation of the formula for gamma cannot be avoided. And because of the second order differentiation, formulae become soon complicated. This might be the reason why gamma calculation is missing in previous works like Broadie and Glasserman (1996) and Fournié et al. (1999). We need to assume for this calculation that b and σ are continuously differentiable up to second order with bounded first and second order derivatives. These conditions are to justify the existence of the weight function. We can then show that one particular solution of the weighting function for the gamma is given by:

$$weight_{\Gamma} = \left[\begin{array}{c} \left(\int_0^T a(t) \frac{Y_t}{\sigma(t,X_t)} dW_t \right)^2 - \int_0^T \mathbb{E}\left[\left(a(t) \frac{Y_t}{\sigma(t,X_t)} \right)^2 \right] dt \\ - \int_0^T a(t) \frac{Y_t \sigma'(t,X_t)}{\sigma^2(t,X_t)} dW_t \\ \int_{0=s \leq s_2}^T \frac{a(s_2) Y_{s_2} Y_{s_1}}{\sigma(s_2, X_{s_2})} \left(b''(s_1, X_{s_1}) - \sigma'(s_1, X_{s_1}) \sigma''(s_1, X_{s_1}) \right) ds_1 dW_{s_2} \\ + \int_{0=s_1 \leq s_2}^T \frac{a(s_2) Y_{s_2} Y_{s_1}}{\sigma(s_2, X_{s_2})} \sigma''(s_1, X_{s_1}) dW_{s_1} dW_{s_2} \end{array} \right] \tag{1.25}$$

Proof: given in the appendix section, section B.3 page 166.□

We can as well define piecewise solutions for the other Greeks : rho and vega:

$$w^{rho}(t) = \sum_{i=1}^{m} \alpha_i^{rho} 1_{t_{i-1} \le t \le t_i}, \qquad w^{vega}(t) = \sum_{i=1}^{m} \alpha_i^{vega} 1_{t_{i-1} \le t \le t_i}$$

We have seen that the generator has to satisfy some necessary and sufficient conditions. Indeed, when rendering these conditions stronger, as when demanding the equality of the terms inside the conditional expectations, we get that the generator satisfy some technical conditions, which can be expressed in terms of the different elements α_i^{delta}, α_i^{vega}, α_i^{rho}. We have summarized these conditions in the table 1.3.

Greeks	**Strong Conditions for the generator in terms of the elements α_j^{greek}**
delta	$\sum_{j=1}^{i} \alpha_j^{delta} \int_{t_{j-1}}^{t_j} \frac{\sigma(t,X_t)}{Y_t} dt = 1$
"extended" rho	$\sum_{j=1}^{i} \alpha_j^{rho} \int_{t_{j-1}}^{t_j} \frac{\sigma(t,X_t)}{Y_t} dt = \int_0^{t_i} \frac{\tilde{b}(t,X_t)}{Y_t} dt$
"extended" vega	$\sum_{j=1}^{i} \alpha_j^{vega} \int_{t_{j-1}}^{t_j} \frac{\sigma(t,X_t)}{Y_t} dt = \frac{Z_{t_i}^{vega}}{Y_{t_i}}$

Table 1.3: Strong Conditions for piecewise constant generator

We can also define weights, which emphasize the role of the first variation process, as linear combination of first variation processes, with linear coefficients β_i^{greek} being stochastic:

$$w^{greek}(t) = \sum_{i=1}^{m} \beta_i^{greek} Y_t 1_{\{t_{i-1} \le t < t_i\}}$$

where the index *greek* stands for either delta, vega or rho. Like in the previous case, we can express the sufficient conditions of the generator in terms of these elements. Like in the previous case, we have summarized all these results in the table 1.4

Greeks	**Conditions for the generator in terms of the elements β_j^{greek}**
delta	$\sum_{j=1}^{i} \beta_j^{delta} \int_{t_{j-1}}^{t_j} \sigma(t, X_t)dt = 1$
"extended" rho	$\sum_{j=1}^{i} \beta_j^{rho} \int_{t_{j-1}}^{t_j} \sigma(t, X_t)dt = \int_0^{t_i} \frac{\tilde{b}(t,X_t)}{Y_t} dt$
"extended" vega	$\sum_{j=1}^{i} \beta_j^{vega} \int_{t_{j-1}}^{t_j} \sigma(t, X_t)dt = \frac{Z_{t_i}^{vega}}{Y_{t_i}}$

Table 1.4: Conditions for piecewise constant times the first variation process solution for the generator

1.4.3 Choice of the generator

When dealing with Malliavin weight, the true question is the choice of the best generator. Since the Skorohod integral coincides with the Ito integral for adapted processes, it is very interesting to find an adapted generator. A second feature is to base the choice of weight on a variance minimization criterion as well. However, this problem is extremely difficult to treat in its general framework. To tackle this issue, one needs to specify our diffusion parameters : drift and volatility term. The problem is then to determine the adapted generator with the lowest formula variance. However, this problem cannot be solved in this too general framework. We need stronger assumptions on the diffusion of the underlying for a fruitful discussion about the choice of the generator.

1.5 Conclusion

In this chapter, we have presented the theoretical framework for the simulation of the Greeks with no differentiation of the payoff function. Its innovation can be classified into two parts:

- We have found the exact condition for a function to serve as a weight function. The problem was solved by means of the Skorohod integrand, referred to as the generator of the weight function.

- We have given the weight function with minimal variance. It is the projection of any weight function on the filtration spanned by the payoff functional.

There are many possible extensions and applications of this theoretical chapter. One area of research is to extend the previous results to other option types: e.g. continuous time Asian options Benhamou as explained in chapter 4 of this Thesis. Another domain of interest is to find specific examples of weight function, according to a certain criterion. The question of the choice of generator needs to refer to stronger hypotheses on the diffusion of the underlying. Another question is a comparison study of the efficiency of Malliavin weights compared to traditional methods. In chapter 2, we examine numerical examples in the particular case of the Black diffusion. They concluded that Malliavin formulas are very efficient for non-linear payoffs but not for vanilla options. Their main conclusion is that one should be cautious when using the Malliavin formulae. As a suggestion, one should use locally the Malliavin formulae at regions of discontinuity and the finite difference method elsewhere. This point will be precisely the subjects of the following chapter.

Chapter 2

Faster Greeks for Discontinuous Payoff Options (A Malliavin Calculus Approach in Black World)

Summary of the chapter

This chapter is a numerical application of the general theory introduced in the first chapter. It examines the case of the Black pricing model. We quantify the gain in the variance reduction when using the Malliavin weighted scheme. We found evidence that this method is very efficient for corridor options, especially for the gamma calculation. Indeed, it can be shown that second order derivatives are the most efficient quantities for the Malliavin weighted scheme. We examine a mixed strategy based on Malliavin weighted scheme as well as a finite difference approximation. The Malliavin weighted scheme is used only locally at the kink of discontinuity. This leads to so called local Malliavin formulae. Local Malliavin formulae offset the drawback of slow convergence of Malliavin weighted scheme for very standard payoffs like call options.

2.1 Introduction

The traditional approach of option pricing relies on hedging. Since the seminal work of Black Scholes (1973), the fair price of an option is given by the portfolio that replicates exactly the option payoff at maturity. If we introduce incompleteness in our model, the hypothesis of perfect replication should be relaxed. One should use different types of criteria to find a price. There is an extensive literature on super-replication or risk minimizing (see for example El Karoui and Quenez (1995), Jouini et al. (1996), Frey (1999)). However, this is not always very realistic since this approach leads to too expensive prices. As a consequence, the derivatives industry still assumes a perfect replicating portfolio and is still very much concerned about the way of calculating it. This problem is commonly referred to as the computation of price sensitivities known as the Greeks.

In this chapter, we examine the particular case of the Black diffusion. We try to quantify the variance reduction induced by the Malliavin method and to define an empirical typology of option for which the Malliavin based formula is more efficient than the traditional finite difference method.

The remainder of this chapter is organized as follows. In section 2, we explain why the finite difference method fails to get fast Greeks for discontinuous payoffs. This suggests to use Malliavin based formulae. In section 3, we give explicit formulae of Malliavin weights for options depending on a finite set of dates. In section 4, we examine simulation results that confirm our theoretical predictions: Malliavin formulae are more efficient for strongly discontinuous payoff options. We define a typology of option types for which the Malliavin based formula should be efficient and quantify the variance reduction on our numerical simulations. We briefly conclude in section 5, giving possible extensions.

2.2 Why a new method for the estimation of the Greeks?

In this section, after summarizing our model hypothesis, we explain why the finite difference, advocated to be quite fast by the use of common random numbers, fails to get efficient estimates of the Greeks in the case of discontinuous payoff options. An exact knowledge of the Greeks is important for risk management issues. Indeed, traders usually delta hedge their portfolio no matter how important transaction costs are. (see the extensive literature on transaction costs: Leland (1985) for the early one, Bensaid et al. (1992), Hodges and Neuberger (1989), Boyle and Vort (1992), Davis, Pana and Zariphopoulou (1993), Hodges and Clewlow (1997), Jouini and Kallal (1991))

2.2.1 Description of the Black pricing model

We consider a continuous-time trading economy with a limited period of horizon $T \in [0, T_\infty]$ ($T_\infty < +\infty$). The uncertainty is characterized by a complete probability space (Ω, F, Q) where Ω is the state space, F is the σ-algebra representing the measurable events, and Q is the risk neutral probability measure[1]. The information evolves according to the augmented right continuous complete filtration $\{F_t, t \in [0, T_\infty]\}$ generated by a standard one dimensional Brownian Motion $\{W_t, t \in [0, T_\infty]\}$. We assume the underlying price process $(X_t)_{t\in[0,T]}$ follows a geometric Brownian motion with a time-dependent volatility,given by the Ito process solution of the following Stochastic Differential Equation:

$$dX_t = r_t X_t dt + \sigma_t X_t dW_t \tag{2.1}$$

with initial condition $X_0 = x$ and with r_t the deterministic risk-free interest rate and σ_t the deterministic Black (1976) volatility. The instantaneous variance

[1]since this basic model assumes markets completeness, this risk neutral probability measure exists and is uniquely defined

$\sigma(t, X_t)$ of the Ito process $(X_t)_{t\in[0,T]}$ is given by $\sigma_t X_t$ and its drift $b(X_t)$ by $r_t X_t$. The "first variation" process, a Ito process itself, defined as a derivative of the Ito Process $(X_t)_{t\in[0,T]}$ with respect to its initial condition x, is proportional to the underlying process

$$Y_t = \frac{X_t}{x}$$

The proportionality between the first variation process and the underlying implies a proportionality between vega and gamma (see Benhamou (2000c)). Therefore, the gamma can be obtained easily for a known vega and vice versa.

2.2.2 The failure of finite difference for discontinuous payoff options

As pointed out by Glynn (1989), by Glasserman and Yao (1992), Boyle, Broadie and Glasserman (1997) and by L'Ecuyer and Perron (1994), a finite difference scheme can be improved by taking common random numbers for the computation of the Greeks. If we denote by $P(x)$ the option price with an initial underlying's level of x, by $X_T(\varepsilon)$ the underlying's value at time T with an initial condition $x+\varepsilon$, and by K the strike of the option, a finite difference scheme for the particular case of the delta leads to approximate the delta by a finite difference approximation like $(P(x+\varepsilon) - P(x))/\varepsilon$. The decisive element for the variance of this estimator is the variance of the numerator, which turns out to be equal to

$$Var(P(x)) + Var(P(x+\varepsilon)) - 2Cov(P(x+\varepsilon), P(x))$$

Therefore, the more positively correlated the two prices $P(x+\varepsilon)$ and $P(x)$ are, the more efficient the above estimate of the Greek is. This is why using common random numbers for the simulation of the two options prices: $P(x+\varepsilon)$ and $P(x)$, is very efficient. Going even further, L'Ecuyer and Perron (1994) proved that the convergence rate is $n^{-1/2}$, which is the best that can be obtained from Monte Carlo simulations. The dramatic success of common numbers relies on the fast rate of the mean-square convergence of $P(x+\varepsilon)$ to $P(x)$. The rate of $n^{-1/2}$ unfortunately does not apply in all cases. For example, it fails to hold in the case of the digital call, an option paying 1 in the case of an underlying above the strike $X_T > K$ and zero elsewhere. This comes from the slow mean square convergence of $P(x+\varepsilon)$ to $P(x)$. The difference between the shifted digital call $P(x+\varepsilon)$ and the regular digital call $P(x)$ is given by a probability times a discount rate squared:

$$\mathbb{E}\left[|P(x+\varepsilon) - P(x)|^2\right] = e^{-2rT} P\left[X_T < K < X_T(\varepsilon)\right]$$

Assuming an homogeneous underlying process, $X_T(\varepsilon) = X_T * \left(1+\frac{\varepsilon}{x}\right)$, it leads to a convergence rate of ε for this probability. Writing with Landau notation, we get that the convergence of $P(x+\varepsilon)$ to $P(x)$ is only linear in ε :

$$\mathbb{E}\left[|P(X_0+\varepsilon) - P(X_0)|^2\right] = O(\varepsilon)$$

On the contrary, in the case of the plain vanilla call option, it can be shown (see for example Broadie and Glasserman (1996)) that for the geometric Brownian motion, the convergence rate is of ε^2

$$\begin{aligned}\mathbb{E}\left[|P(x+\varepsilon) - P(x)|^2\right] &\leq \mathbb{E}\left[|X_T(\varepsilon) - X_t|^2\right] \\ &\leq \varepsilon^2 \mathbb{E}\left[e^{(r-\mu)T+\sigma\sqrt{T}Z}\right]\end{aligned}$$

where Z is a normal variable $N(0,1)$, leading to

$$\mathbb{E}\left[|P(x+\varepsilon) - P(x)|^2\right] = O\left(\varepsilon^2\right)$$

This is why the methodology of finite difference under-performs for all discontinuous type options like simple digital, corridor (option which pays 1 if the underlying at time T is inside an interval $L < X_T < H$), barrier option and so on.

2.3 Determination of the Malliavin Weights

To overcome this problem, Fournié et al. (1999) and Benhamou (2000a) (reprinted as chapter 1 of this Thesis) advocated the use of an integration by parts formula so as to construct smooth estimators of the Greeks. This section show how to apply the results derived by in the chapter 1. We remind important characteristics of Malliavin weights:

- 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. The method efficiency is therefore increased for discontinuous payoff options.
- the weight functions are given as the Skorohod integral of some generator, characterized by necessary and sufficient conditions being expressed through conditional expectations. However, since it is easier to handle and still very robust, we use sufficient and stronger conditions that are the equality almost surely of the terms inside the conditional expectations.
- there is an infinity of solutions for the generator function. However, it is more efficient to choose weight functions expressed with the same points of the Brownian motion trajectory as the option payoff. For an option depending on a series of dates $t_1 < t_2 < ... < t_m$ it is appropriate to choose a weight function expressed in terms of $W_{t_1}, ..., W_{t_n}$. No extra simulation is required for the computation of the weight function and it can be shown that the variance of the weight function is minimum.

The latter stronger condition can in some cases not be fulfilled. In these cases, it becomes very difficult to determine the most efficient weight functions. In the rest of the chapter, we take the convention that $t_0 = 0$. We denote by F the discounted payoff. The option is depending on a series of increasing dates $t_1 < t_2 < ... < t_m$ with $t_m = T$. This dependance is very general. It can represent many options depending on a finite set of dates, as in discrete Asian, barrier and lookback options.

2.3.1 Delta

In chapter 1, we have seen that the delta is equal to the expected value of the discounted payoff F times a weight function expressed as a Skorohod integral (equation (2.2))

$$\text{delta} = E_x^Q \left[F\delta \left(w^{\text{delta}} \right) \right] \tag{2.2}$$

where the function w^{delta}, called the weighting function generator, has to satisfy a sufficient condition given by:

$$\int_0^T x\sigma_t w^{\text{delta}}(t) 1_{\{t \leq t_i\}} dt = 1 \quad \forall i = 1...m \tag{2.3}$$

as well as the L^2 integrability of its Skorohod integral, which is the condition for the existence of the Skorohod integral (Øksendal (1997) page 22).

Among the different weight function generators, it can be shown that the one with the lowest variance is the one expressed in terms of the same points of the Brownian motion as the option payoff, that is to say $W_{t_1}, ..., W_{t_n}$. This implies to use a piecewise constant generator. We denote by $(\lambda_i)_{i=1..n}$ the sequence initialized with $\lambda_1 = \frac{1}{x\int_0^{t_1} \sigma_t dt}$ and defined by the following recurrence: for $1 \leq i_0 < n$, λ_{i_0+1} is given by

$$\lambda_{i_0+1} = \frac{\frac{1}{x} - \sum_{i=1}^{i_0} \int_{t_{i-1}}^{t_i} \lambda_i \sigma_t dt}{\int_{t_{i_0}}^{t_{i_0+1}} \sigma_t dt}$$

With these definitions, the most appropriate solution for our generator (in terms of computation) is given by the following proposition:

Proposition 3 *The piecewise solution for our generator is given by*

$$w^{delta}(t) = \sum_{i=1}^{n} \lambda_i 1_{[t_{i-1}, t_i[}(t) \tag{2.4}$$

leading to the following expression for the delta

$$delta = \mathbb{E}_x^Q \left[F \sum_{i=1}^{n} \lambda_i \left[W_{t_i} - W_{t_{i-1}} \right] \right]$$

Proof: the solution (2.4) verifies condition (2.3).□

Remark 1 *In the particular case of an option depending only on a final date T (European option) with a Black Scholes diffusion ($\sigma_t = cte = \sigma$), we find the following particular solution:*

$$\delta = \mathbb{E}\left[e^{-\int_0^T r_s ds} f(X_T) \frac{W_T}{T\sigma x} \right] \tag{2.5}$$

The weight function is very simple in this special case. It is the Brownian motion divided by the maturity of the option times the volatility times the initial condition. This suggests that for an option close to maturity, the Malliavin weight of the delta should explode. Indeed, when the option is close to maturity, the condition (2.3) leads to increase the generator. The problem of a wider hedge close to the maturity is well-known, especially in the literature about barrier options. As far as the volatility is concerned, the intuition is that more volatility makes the option price more convex. It smoothens in a way the Greeks. This is why it is consistent with the decrease of the Malliavin weight with respect to the volatility parameter.

2.3.2 Gamma

The gamma (Γ) computation is harder than for the other Greeks since it is a second order derivative. However, since in the Black model, the first variation and underlying process are proportional, there is a proportionality between gamma and vega (see Benhamou (2000c)). The vega v is given by (in the case of a European option)

$$v = x^2 \sigma T * \gamma$$

Using this property enables us to compute easily the gamma. That is why we do not develop any further our analysis for the gamma and refer to the vega section for an elegant way of calculation. So as to be as extensive as possible, we can mention that a straightforward computation of the gamma can be done. Using the theoretical results on the Malliavin weight function (see Benhamou (2000a)), we find that one particular solution for the weight function of the gamma is given by:

$$weight_{\Gamma} = \left[\left(\int_0^T \frac{\lambda_t}{x\sigma_t} dW_t \right)^2 - \int_0^T \left(\frac{\lambda_t}{x\sigma_t} \right)^2 ds - \int_0^T \frac{\lambda_t}{x^2\sigma_t} dW_t \right] \tag{2.6}$$

Remark 2 *In the particular case of an option depending only on a final date T (European option) with a Black Scholes diffusion ($\sigma_t = cte = \sigma$), we find:*

$$\Gamma = E\left[e^{-\int_0^T r_s ds} f(X_T) \frac{1}{T\sigma x^2} \left(\frac{W_T^2}{\sigma T} - \frac{1}{\sigma} - W_T \right) \right]$$

2.3.3 Rho

The meaning of the classical rho is to examine the price sensitivity with respect to the risk-free rate. The results derived by Benhamou (2000a) are for a perturbation on the drift part of the diffusion of the underlying. However, a change in the risk-free rate impacts in two ways. It alters the drift of the underlying diffusion but it also modifies the discount factor:

$$\begin{aligned} rho &= \frac{\partial}{\partial \varepsilon} \mathbb{E}_x^Q \left[e^{-\int_0^{t_m} r_s ds} f\left(X_{t_1}^{\varepsilon,rho}, ..., X_{t_m}^{\varepsilon,rho} \right) \right] \\ &\quad + \frac{\partial}{\partial \varepsilon} \mathbb{E}_x^Q \left[e^{-\int_0^{t_m} r_s + \varepsilon \widetilde{r}_s ds} f(X_{t_1}, ..., X_{t_m}) \right] \end{aligned} \tag{2.7}$$

where $X_t^{\varepsilon,rho}$ stands for the underlying with a perturbed drift $\widetilde{b}(u, X_u) = X_u \widetilde{r}_u$, and where the limit is almost surely, taken for $\varepsilon = 0$. The second term can be calculated by interchanging the expectation and the derivative operator and differentiating the discount factor with respect to ε:

$$\frac{\partial}{\partial \varepsilon} \mathbb{E}_x^Q \left[e^{-\int_0^{t_m} r_s + \varepsilon \widetilde{r}_s ds} f(X_{t_1}, ..., X_{t_m}) \right] = \mathbb{E}_x^Q \left[-\int_0^{t_m} \widetilde{r}_s ds \, e^{-\int_0^{t_m} r_s} f(X_{t_1}, ..., X_{t_m}) \right]$$

Like in the case of the delta, the first term of the right hand side of equation (2.7) can be expressed as the expected value of the discounted payoff F times a weight function expressed as a Skorohod integral

$$\frac{\partial}{\partial\varepsilon}\mathbb{E}_x^Q\left[e^{-\int_0^{t_m} r_s ds} f\left(X_{t_1}^{\varepsilon},...,X_{t_m}^{\varepsilon}\right)\right] = \mathbb{E}_x^Q\left[F\delta\left(w^{\text{rho}}\right)\right] \tag{2.8}$$

where the function w^{rho}, called the weight function generator has to satisfy a sufficient condition given by the following equation

$$\int_0^{t_i} x\sigma_t w^{\text{rho}}(t)\,dt = \int_0^{t_i} x\widetilde{r}_t dt \quad \forall i = 1...m \tag{2.9}$$

as well as the L^2 integrability of its Skorohod integral. An obvious solution is $w^{\text{rho}} = \frac{\widetilde{r}_t}{\sigma_t}$. Using the fact that the classical rho is the price sensitivity with respect to the risk-free rate, we get the following proposition:

Proposition 4 *The rho is given by:*

$$rho = E^Q\left[F * \left(\int_0^{t_m} \frac{\widetilde{r}_t}{\sigma_t} dW_t - \int_0^{t_m} \widetilde{r}_t dt\right)\right]$$

2.3.4 Vega

The vega is the perturbation along the volatility term of the diffusion. We write the perturbation as $\widetilde{\sigma}(t, X_t) = \widetilde{\sigma}_t X_t$. Like in the case of the other Greeks, we can define a weight function characterized by its generator. The weight function generator $w^{vega}(.)$ should satisfy:

$$\int_{t=0}^{t_i} \sigma_t w^{vega}(t)\,dt = \int_{t=0}^{t_i} \widetilde{\sigma}_t dW_t - \int_{t=0}^{t_i} \sigma_t\widetilde{\sigma}_t dt \quad \forall i = 1...n \tag{2.10}$$

as well as the L^2 integrability of its Skorohod integral, which is the existence of the Skorohod integral. One possible solution is a piecewise constant solution. We denote by $(\lambda_i)_{i=1..n}$ the sequence initialized with

$$\lambda_1 = \frac{\int_{t=0}^{t_1} \widetilde{\sigma}_t dW_t - \int_{t=0}^{t_1} \sigma_t\widetilde{\sigma}_t dt}{\int_{t=0}^{t_1} \sigma_t dt}$$

and defined by the following recurrence: for $1 \leq i_0 < n$, λ_{i_0+1} is given by

$$\lambda_{i_0} = \frac{\int_{t=0}^{t_{i_0}} \widetilde{\sigma}_t dW_t - \int_{t=0}^{t_{i_0}} \sigma_t\widetilde{\sigma}_t dt - \sum_{i=1}^{i_0-1} \int_{t_{i-1}}^{t_i} \lambda_i\sigma_t dt}{\int_{t_{i_0-1}}^{t_{i_0}} \sigma_t dt} \tag{2.11}$$

We then get that there is a piecewise constant solution as the following proposition states it:

Proposition 5 *piecewise constant solution*

One particular solution for the weight function generator is defined by:

$$w^{vega}(t) = \sum_{i=1}^{n} \lambda_i 1_{[t_{i-1}, t_i[}(t) \tag{2.12}$$

Proof: The solution given by equation (2.12) verifies the necessary and sufficient condition to be a weighting function generator equation (2.10).□

Corollary 1 *In the case of an option depending only on a final date denoted by T, we get*

$$vega = \mathbb{E}\left[e^{-\int_0^T r_s ds} f(X_T)\, \delta\left(\frac{\int_0^T \widetilde{\sigma}_t dW_t - \int_0^T \sigma_t \widetilde{\sigma}_t dt}{\int_{t=0}^T \sigma_t dt}\right)\right]$$

Proof: immediate, since the weight function is defined as the Skorohod integral of the particular solution for the weighting function generator w^{vega} given by the equation (2.12).□

Corollary 2 *In the Black Scholes case, we get*

$$vega = \mathbb{E}\left[e^{-\int_0^T r_s ds} f(X_T) \frac{\widetilde{\sigma}}{\sigma T}\left(W_T^2 - T - \sigma T W_T\right)\right] \tag{2.13}$$

Proof: Using the fact that the Skorohod integral is a linear operator and that the Skorohod integral reduces to the Ito integral for adapted process, we get

$$\delta\left(\frac{\widetilde{\sigma} W_T}{\sigma T} - \widetilde{\sigma}\right) = \frac{\widetilde{\sigma}}{\sigma T}\int_{u=0}^{T}\int_{v=0}^{T} dW_u dW_v - \widetilde{\sigma} W_T$$

We need to calculate $\int_{u=0}^T \int_{v=0}^T dW_u dW_v$. This expression can be seen as a Wiener Chaos term of second order and is related to the Hermite polynomial of second order, so that (Øksendal (1997) page 19)

$$\int_{u=0}^{T}\int_{v=0}^{T} dW_u dW_v = W_T^2 - T$$

Putting all these terms together leads to the result. □

Corollary 3 *The classical vega is given by*

$$Classical\ vega = E\left[e^{-\int_0^T r_s ds} f(X_T)\left(\frac{W_T^2}{\sigma T} - \frac{1}{\sigma} - W_T\right)\right]$$

Proof: to obtain the classical vega, we must divide the above formula (2.13) by $\widetilde{\sigma}$.□

2.4 Numerical Result on the Efficiency of Malliavin weights

In this section, we compare the results of Malliavin weighted simulations with the ones obtained by a centred finite difference approximation $(P(x+\epsilon) - P(x-\epsilon))/2\epsilon$ for three different types of options in the Black-Scholes framework (so as to have closed formula):

- a corridor option: the payoff, given by $1_{\{S_{\max} > S_T > S_{\min}\}}$, displays two discontinuities. This is exactly the type of options we are targeting to since it has two discontinuities.
- a binary call : the payoff given by $1_{\{S_T > S_{\min}\}}$ displays only one discontinuity. The payoff is smoother than the one of a corridor option.
- a vanilla call. This last example is to examine the impact of the formula when there is a smooth payoff.

A point we had to resolve at first, was the type of simulations to use. Boyle Broadie and Glasserman (1997), Caflisch Mokoroff and Owen (1997), Galanti and Jung (1997), Boyle Joy and Tan (1997), Papageorgiou and Traub (1996), Paskov (1994), Paskov and Traub (1995) and Williard (1997) show that low-discrepeny sequences are more efficient than random sequences for low dimension problems. Bratley, Fox and Niederreiter (1992), Galanti and Jung (1997), Morokoff (1997) and Moskovitz and Caflish (1995) demonstrate that low discrepancy sequences become less efficient for high dimensions. Galanti and Jung (1997) demonstrate that the Sobol sequence exhibits better convergence properties than either the Halton and Fauré sequences. Therefore, we used the Sobol sequence.

Since the Sobol sequence fills the space with a pseudo periodicity, the simulations display pseudo-periodicity as well. We took the same parameters in the three option examples: X_0=100, r=5%, σ=15%, T=1 year, $S_{\min}$=95,

S_{max}=105, K=100. We display for each option the delta and the gamma. Rho and vega parameters lead to same results and are given for illustrative purpose in the particular case of the corridor option.

The results consist in two remarks:

- For discontinuous payoff function, as is the case of digital and corridor option, with a mean-square convergence of the shifted option $P(X_0+\varepsilon)$ to $P(X_0)$ linear in ε (see section 2.2.2, page 49), Malliavin formula outperforms finite difference method. This is because Malliavin simulation has lower simulation variance and converges faster. This comes from two self-reinforcing facts. First, Malliavin technique uses a smoothened payoff. Second, finite difference method is lengthy because of the slow mean-square convergence of the shifted option $P(X_0+\varepsilon)$ to the normal one $P(X_0)$.
- On the contrary, for vanilla options, smooth enough not to require the integration by parts technique and for which the mean-square convergence of a shifted option $P(X_0+\varepsilon)$ to the normal one $P(X_0)$ is quadratic in ε (see section 2.2.2, page 49), finite difference outperforms the Malliavin-based method.

2.4.1 Comparative analysis: Finite Difference versus Malliavin weighted scheme

Corridor Option

This important example illustrates the drastic efficiency of the Malliavin theory. The corridor option pays 1 if the underlying at maturity is inside the corridor: payoff equal to $1_{\{S_{max}>S_T>S_{min}\}}$. The outperformance of the Malliavin simulation is illustrated by the figures 2.1, 2.2 which display the delta and gamma of the corridor option. Results on the vega and rho are similar. They are given only for illustrative purpose as figures 2.3 and 2.4. A more quantitative analysis of the result is given in the section 2.4.2.

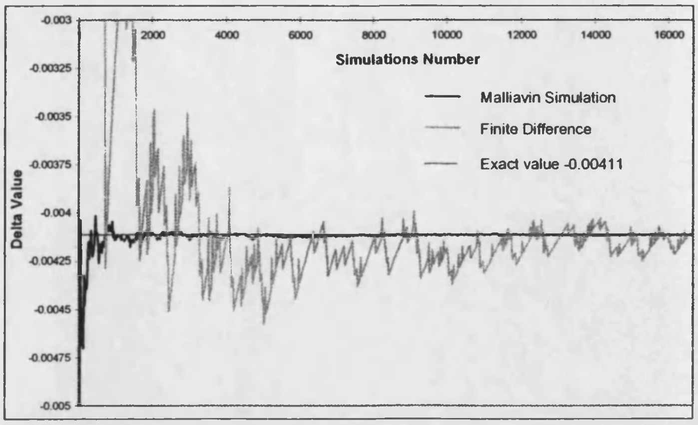

Figure 2.1: Efficiency of the Malliavin weighted scheme for the computation of the delta of a Corridor option

The figure 2.1 compares the two methods: Malliavin weighted scheme (black line) and the finite difference method (grey line). The Malliavin weighted scheme converges to the right answer fast with almost no oscillations, whereas the finite difference estimator fluctuates with a pseudo-periodicity around the correct value

The figure 2.2 examines the computation of the gamma. Like in the case of the delta, the Malliavin weighted scheme outperforms dramatically compared to the finite difference method. It is worth noticing that this outperformance is even more pronounced for the gamma than for the delta.

Vega and Rho for the Corridor option

We decided to study the delta and gamma to compare different type of options. However, the efficiency (or not) of the Malliavin weighted scheme for the rho and vega is similar to the case of the delta and gamma.

Binary Option

The binary option is a didactic example of a payoff function with small discontinuity ($1_{S_T > S_{\min}}$). Like for the corridor option, Malliavin weighted simulations

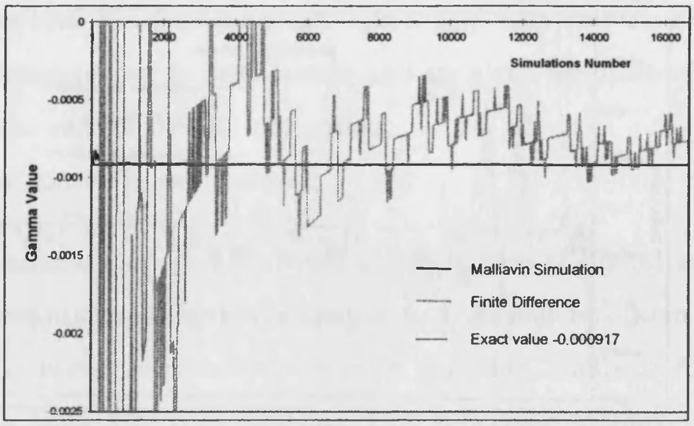

Figure 2.2: Efficiency of the Malliavin weighted for the computation of the gamma of a Corridor option

compute faster and more accurately the Greeks than the finite difference method. Finite difference simulation performs poorly since the mean-square convergence of the shifted option $P(X_0 + \varepsilon)$ to $P(X_0)$ is only linear in ε. Figures 2.5 and 2.6 are respectively example of delta and gamma computation. They illustrate the outperformance of the Malliavin weighted scheme.

Like for the corridor option, Malliavin outperformance is more pronounced for the gamma than the delta as a comparative study of figure 2.5 and 2.6 shows. Gamma is a second order Greek. This suggests an increased efficiency for higher order Greeks.

Call Option

Last but not least, the Call option is an instructive case of a smooth payoff function. Since the payoff function does not present any strong discontinuity, it is smooth enough not to require any integration by parts smoothing. Therefore, the Malliavin-weighted simulations do not provide any technical advantage. Indeed, since the mean-square convergence of the shifted option $P(X_0 + \varepsilon)$ to the normal one $P(X_0)$ is quadratic in ε, the finite difference method embodies a pseudo antithetic variance reduction. As a consequence, it converges faster than the Malliavin method as shown by figures 2.7 and 2.8, which represent respectively the delta and gamma.

Even for the case of the gamma, which is a second order derivative, the finite difference is more efficient than the Malliavin based formula as proved by figure 2.8. The second order smoothing is not enough to offset the quadratic convergence of the shifted option $C(X_0 + \varepsilon)$ to the normal one $C(X_0)$ as shown by figure 2.8. Still, the comparative study of the figures 2.7 and 2.8 indicates an increased efficiency of the Malliavin weighted scheme for second order Greeks like gamma.

2.4.2 Typology of options requiring Malliavin weighted scheme

As shown by simulation examples, a paradox of the Malliavin weighted scheme is the implication of the payoff function on the method. The weight function does not rely on the payoff function. However, its comparative advantage versus finite differences does depend on the form of the payoff. Indeed, the finite difference method are crucially related to the form of the option payoff. Moreover, the weight function is not depending on the payoff. However, the total variance of the Greeks simulated by a Malliavin weighted scheme is the variance of the product of the payoff function times the Malliavin weights. This does depend on the payoff type. An interesting and open problem is to classify the option types for which the Malliavin weighted scheme should be preferred. In this section, we precisely try to define a typology of options for which the Malliavin technology outperforms the traditional finite difference method. We can make many remarks:

- the Malliavin weight function is independent from the option payoff. This indicates that the disturbance caused by the weight function is not influenced by the payoff. This is not the case of the finite difference method for which the payoff function matters crucially.

- the weight function explodes for small maturities. This suggests that the Malliavin technology is inappropriate for small maturities options.

- the computation of the gamma is similar to the vega since in the case of the Black and Black Scholes model, there is a direct proportionality between the gamma and vega coefficient. The proportionality can be read on the weight function, whereas it is not obvious in the finite difference method. A standard finite difference method would lead to compute the gamma by the finite difference approximation

$$\Gamma \approx \frac{\text{Price}\,(S_0 + dS_0, \sigma) - 2\text{Price}\,(S_0, \sigma) + \text{Price}\,(S_0 - dS_0, \sigma)}{dS_0^2}$$

  as well as the vega $v \approx \frac{\text{Price}(S_0,\sigma+d\sigma)-\text{Price}(S_0-,\sigma-d\sigma)}{2d\sigma}$.

- the Malliavin technology in the case of the gamma reduces a second order differentiation to no differentiation. This implies that the efficiency of the Malliavin method is enhanced in the case of the gamma compared to the delta.

Before, giving an empirical typology of option, we quantify the variance reduction induced by the Malliavin method. And we can claim that Malliavin based formula is a variance reduction technique. This is well illustrated by table 2.1, where we have given the ratio between the estimated volatility of the Malliavin simulation and the finite difference simulation. We can see that the method is more efficient for gamma, then for vega.

The variance reduction is of comparable order for delta and rho. The number of simulation draws for the table 2.1 was N=20,000. We give the ratio of simulation variances between finite difference and Malliavin-based simulation. Since the variance decreases roughly linearly in n, a ratio of ten means that we

Option type	Variance ratio $\widehat{\sigma}^2_{\text{Finite Difference}}/\widehat{\sigma}^2_{\text{Malliavin}}$	delta	gamma	rho	vega
Call	N=20,000	0.1273	0.1272	0.401	0.0735
Binary	N=20,000	7.15	4916	6.56	81
Corridor	N=20,000	144.98	6864	33	5920

Table 2.1: Comparison of the Malliavin weighted scheme and the finite difference method

need to do 10 n draws in the finite difference method to get the same variance as the one obtained by the Malliavin method with only n draws.

We found that for the corridor option, the Malliavin weighted scheme, when compared to finite difference, improved the computation of the Greeks by a factor bigger than 100 for the case of the delta, 6000 for the case of the gamma, 33 for the case of the rho and 5900 for the case of the vega as stated by table 2.1. These are big numbers. It means for example that we need about 6 millions of draws to compute the gamma with the finite difference method to get the same accuracy as a simulation based on a Malliavin weighted scheme.

The faster convergence of Malliavin weighted scheme over the finite difference method with common random numbers comes from the fact that the Malliavin method avoids differentiating and smoothens considerably the payoff of the option to simulate.

Summarizing all the results given by the simulations, we draw the following conclusions:

- The Malliavin method is appropriate for option for which the mean-square convergence of a shifted option $P(X_0 + \varepsilon)$ to the normal one $P(X_0)$ 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. This is the case of any binary and corridor option.

- The maturity of the option is a crucial factor for the Malliavin method since it leads to an exploding weight function. However, the traditional method underperforms as well.

- The Malliavin method leads to weight functions which are roughly (polynomial) functions of the Brownian motion. The variance of the weight function increases for high values of the Brownian motion. To get the Greek, we multiply the weight function by the option payoff. 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. Therefore, it is more appropriate to use the put-call parity and calculate Greeks only for put options. Furthermore, we can use a mixed strategy referred to as the Malliavin local approach, which smoothens the discontinuity locally. This is the subject of the next subsection.

- The relative performance shows however that the Malliavin weighted underperforms only slightly for the case of call but outperforms greatly for the case of the corridor option as shown by the figures 2.9, 2.10, 2.11, 2.12. The relative performance is calculated as the ratio of the difference between the simulation result and the theoretical result over the absolute value of the theoretical one. Therefore, in the charts, a positive relative performance means that the simulation's estimate is greater than the theoretical one by the given relative percentage. The opposite holds. A negative relative performance of 10% means that the simulation underestimates the result by a relative 10%.

2.4.3 Local Malliavin formulae

The intuition behind the integration by parts is to smoothen the payoff at the discontinuity kink. However, there is no advantage in using the Malliavin formula when the payoff is smooth. This hints at using a mixed strategy. At the discontinuity, we use an integration by parts by means of the Malliavin formula. Elsewhere, we use the traditional finite difference. The finite difference method contains a variance reduction method of antithetic variate when implemented with common numbers (as explained in section 2.2.2, page 49). Let us describe the idea on the case of the delta of a call. We have seen that the delta can be written as the expected value of a payoff times a weighting function (section 2.3.1, page 52), which in the simple case of the Black Scholes framework leads for a European option to

$$\delta = \mathbb{E}\left[e^{-\int_0^T r_s ds}(X_T - K)^+ \frac{W_T}{x\sigma T}\right]$$

The weight function is multiplied by the term $(X_T - K)^+$ which is big for large values of X_T, corresponding to large values of the Brownian motion W_T. This generates some increased variance because of the weight function $W_T/x\sigma T$. When X_T is "large", W_T is "large" and therefore $(X_T - K)^+ * W_T/x\sigma T$ is even "larger" with a substantial variance. Writing the delta as the sum of two terms, we get

$$delta = \frac{\partial}{\partial x}\mathbb{E}\left[e^{-\int_0^T r_s ds}(X_T - K)^+ 1_{\{K<X_T<K+\varepsilon\}}\right] + \frac{\partial}{\partial x}\mathbb{E}\left[e^{-\int_0^T r_s ds}(X_T - K)^+ 1_{\{K+\varepsilon\le X_T\}}\right]$$

Using the Malliavin integration by parts only for the first term, and interchanging the expectation and the differentiation operator for the second, we come to

$$delta = \mathbb{E}\left[e^{-\int_0^T r_s ds}(X_T - K)^+ 1_{\{K<X_T<K+\delta\}}\frac{W_T}{x\sigma T}\right] + \mathbb{E}\left[e^{-\int_0^T r_s ds}1_{\{K+\delta<X_T\}}Y_T\right]$$

Indeed, it is very efficient to take a small localization parameter like $\varepsilon = 1$. In this case, it leads to a reduction of variance, that is to say $\widehat{\sigma}^2_{\text{Malliavin}}/\widehat{\sigma}^2_{\text{LocalMalliavin}}$ $=13.88$ and a variance reduction from finite difference to local Malliavin of 1.77, that is to say $\widehat{\sigma}^2_{\text{Finite Difference}}/\widehat{\sigma}^2_{\text{LocalMalliavin}}$ $=1.77$. Therefore, the Malliavin local formula is more efficient than the standard finite difference method. The factor of 1.77 means intuitively that we need a simulation of 17,000 draws with a finite difference approximation to get the accuracy of a 10,000 draws simulation with the local Malliavin based formula. Indeed, the crucial point in this formula is to find an interesting value of ε, taken here as 1% of the underlying initial level. It would be nice to examine the impact of this parameter on the variance reduction with some theoretical considerations. That is one of the promising future area of research for the Malliavin technique.

2.5 Conclusion

In this chapter we have seen that using Malliavin calculus and its integration by parts formula, we can smoothen the function to be estimated by the Monte Carlo or Quasi Monte Carlo procedure. This outperforms traditional finite difference method in the case of digital option as well as corridor, with a gain on the variance of the simulation of more than 4900 and 6800 for the gamma of respectively the digital and the corridor option.

However, we recommend a cautious use of the Malliavin formula. It turns out to be very efficient for discontinuous functions like a digital, corridor payoff function. However, for smooth functions, it can handle the computation of the Greeks more inefficiently than a finite difference method. This is because the finite difference method includes an antithetic variate variance reduction method. We suggest to use a local version of the Malliavin method, so as to smoothen the payoff at the kink and elsewhere to use finite difference method with common random numbers. Other relationships like put-call parity should be used as well.

There are many extensions to this chapter, especially to more complicated models than the Black one. An interesting enlargement is the advanced study of the local Malliavin method. As a conclusion, we conjecture that the Malliavin method is going to have an increasing influence over the next years since it is a powerful method to compute the Greeks.

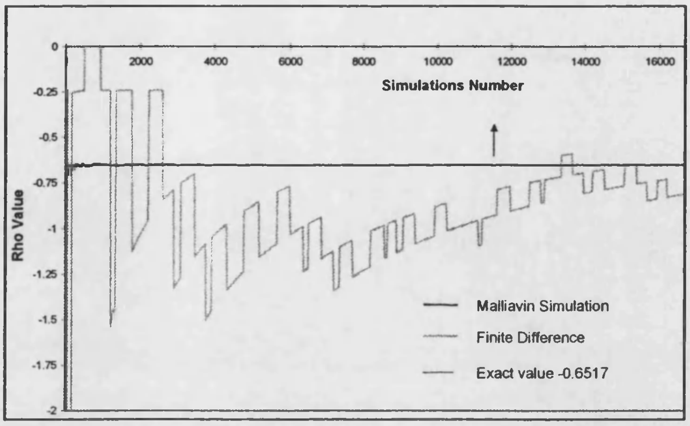

Figure 2.3: Efficiency of the Malliavin weighted for the computation of the rho of a Corridor option. The pseudo periodicity of the finite difference comes from the pseudo periodicity with which the Sobol sequence fills the space

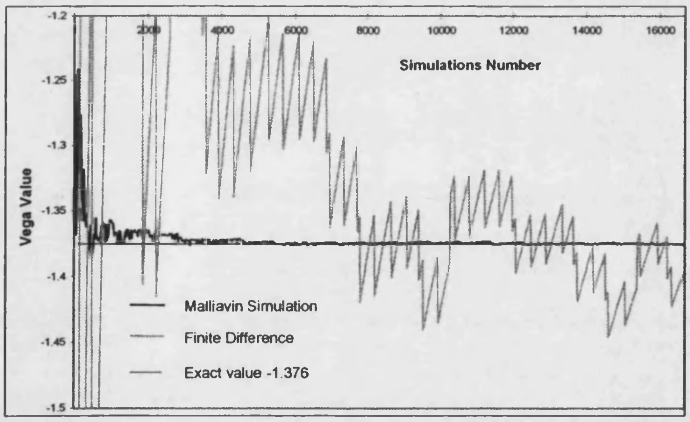

Figure 2.4: Efficiency of the Malliavin weighted for the computation of the vega of a Corridor option. The case of the vega is very similar to the one of gamma since these two sensitivities are proportional in the case of a single option in the Black Scholes framework

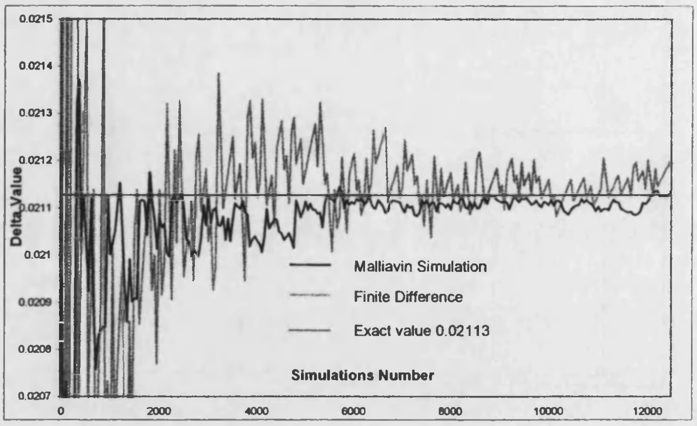

Figure 2.5: Comparison of the computation of the Delta of a Binary option by finite differences and by Malliavin weighted scheme

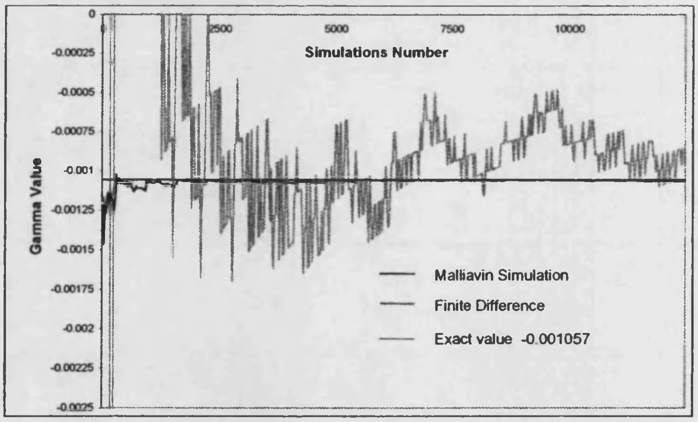

Figure 2.6: Comparison of the computation of the Gamma of a Binary option by finite differences and by Malliavin weighted scheme

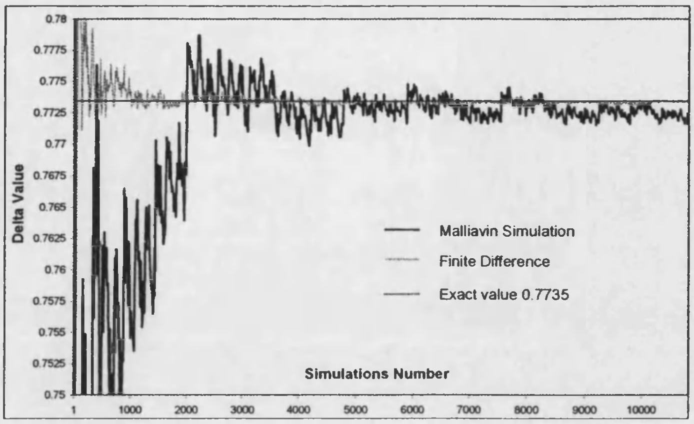

Figure 2.7: Comparison of the computation of the Delta of a Call option by finite differences and by Malliavin weighted scheme

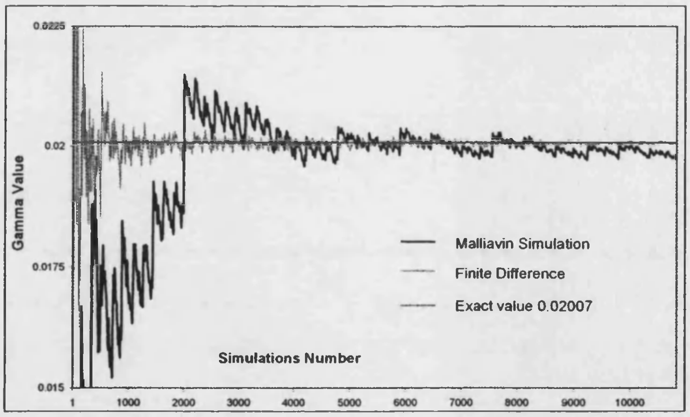

Figure 2.8: Comparison of the computation of the Gamma of a Call option by finite differences and by Malliavin weighted scheme

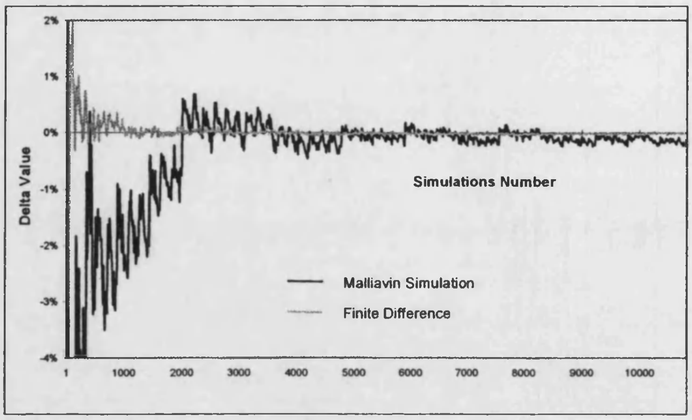

Figure 2.9: Relative performance of the Finite difference method over the Malliavin weighted scheme for the delta of a call option

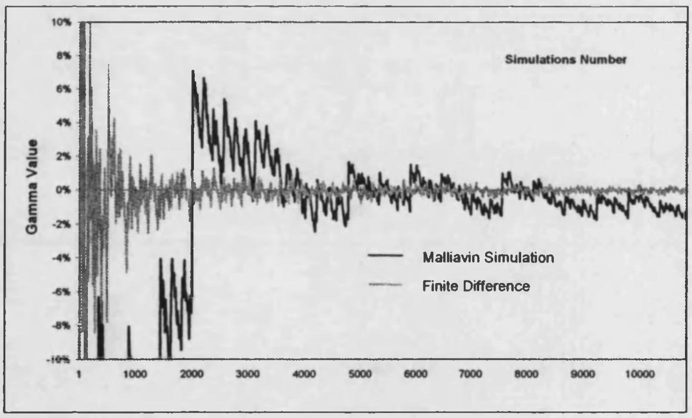

Figure 2.10: Relative performance of the Finite difference method over the Malliavin weighted scheme for the gamma of a call option

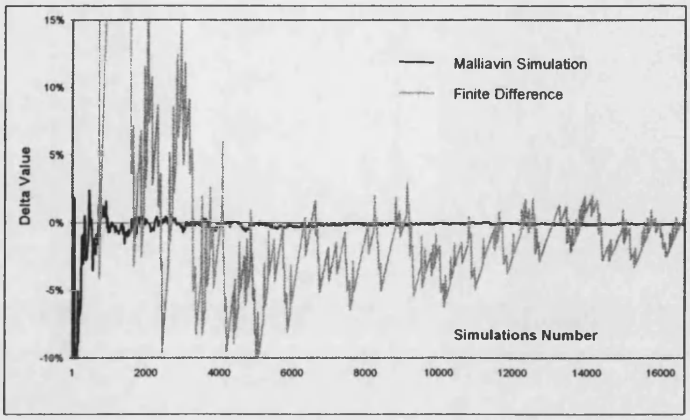

Figure 2.11: Relative performance of the Malliavin weighted scheme over the Finite difference method for the delta of a corridor option

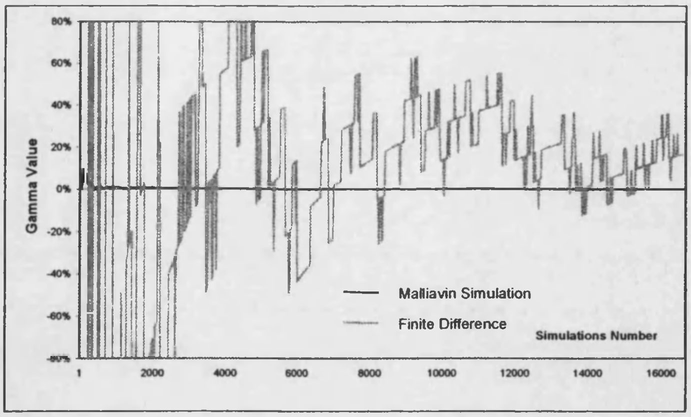

Figure 2.12: Relative performance of the Malliavin weighted scheme over the Finite difference method for the gamma of a corridor option

Chapter 3

Fast Fourier Transform for Discrete Asian Options

Summary of the chapter

This chapter presents a new methodology for pricing discrete Asian options consistent with different types of underlying densities, especially non-normal returns as suggested by the empirical literature (see Mandelbrot (1963) and Fama (1965) for the early ones). The interest of this method is its flexibility compared to the more standard ones. Based on Fast Fourier Transform, the algorithm is an enhanced version of the algorithm of Caverhill and Clewlow (1992). The contribution of this chapter is to improve their algorithm by a systematic recentering at each stage and to adapt it to non-lognormal densities. This enables us to examine the impact of fat-tailed distribution on price as well as on delta. We find evidence that fat-tailed densities lead to wider jumps in the delta.

3.1 Introduction

First introduced in Tokyo, Asian options are options based on any type of average of underlying equity prices, interest rates or indices. They are among the most popular path-dependent derivatives, since their characteristics capture partially the trajectory of the underlying, with often reduced exposure to volatility. In addition, Asian options are less sensitive to possible spot manipulations or extreme movements at settlement and offer flexibility in the way the average is settled. Consequently, they have become very attractive for investors since they provide a customized cheap way to hedge periodic cash-flows (see Longstaff (1995) for a discussion of the efficiency of Asian interest-rate options for corporations with reasonably predictable cash flows)

When pricing an option, one of the first questions that arises concerns the distributional assumptions for the underlying. Very often the distribution of the latter is taken to be lognormal as in the Black Scholes model. However, when it comes to arithmetic Asian options, one is confronted with the problem of the distributions. Indeed, the empirical literature has rejected normality of returns and hence the geometric Brownian motion. It has rather suggested fat-tailed distributions (see Mandelbrot (1963) and Fama (1965) for the early ones).

The motivation of this chapter is therefore to provide an efficient method for the pricing of Asian options consistent with various underlying densities, especially non log-normal ones. Because of the challenge of getting a correct price for Asian option with a widely used option pricing model, previous research has focussed on the Black Scholes model, adopting different strategies. It has first focussed on the geometric Asian option case (Vorst (1992), Turnbull and Wakeman (1991), Zhang (1995)). It has as well looked at the question of the continuous-time Asian options (Geman and Yor (1993), Rogers and Shi (1995), Alziary et al. (1997), He and Takahashi (1996), Forsyth et al. (1998), Nielsen and Sandmann (1998)). However, the type of average for traded Asian options is arithmetic and discrete: daily, weekly or monthly. Approximating these options by their continuous-time limit is inaccurate and misleading for options with a period of time between two fixing dates longer than a day.

To account for the discrete arithmetic averaging, it has been suggested to use different approximations of the density of the sum of lognormal variables leading to various closed-form solutions: approximation via the geometric average (Vorst (1992)), via a lognormal density (Turnbull and Wakeman (1991)), via an Edgeworth expansion (Levy (1992) and Jacques (1996)), via a Taylor expansion (Zhang (1998) and Bouaziz et al. (1998)) or via the reciprocal Gamma distribution (Milesvky and Posner (1997)).

It has also been advocated to use different numerical methods: Monte Carlo (Kemma and Vorst (1990)), tree methods (Hull and White (1997)) and Fast Fourier Transform techniques (Caverhill and Clewlow (1992)). However, none of these works has considered non-lognormal distributions.

When the underlying density is not lognormal, the approximation methods do not hold any more since they heavily rely on the lognormal assumption. Numerical methods like PDE or lattice methods are as well not easy to adapt to the non-lognormal case, since we need to restrict ourselves to certain types of diffusion like stochastic volatility or deterministic volatility models which implies strong assumptions on the underlying diffusion. It is not very straightforward to derive an empirical density from market data, requiring very often a calibration stage. The two methods adaptable to an ad-hoc empirical non lognormal distribution without too much difficulty, are indeed the Monte Carlo and the Fast Fourier Transform method. However, these two methods perform poorly for non-lognormal case as well as for lognormal one. The Monte Carlo has the drawback to be slow. The algorithm of Caverhill and Clewlow (1992) requires large discretization grid and has slow convergence.

In this chapter, we offer a solution to improve the method of Caverhill and Clewlow (1992) and to adapt it to the case of non-lognormal densities. To reduce the size of the grid and therefore the computational time, we recenter intermediate densities. We test this algorithm in the lognormal case since it is only in this particular situation that we have benchmarks in the literature. We then examine the impact of non-lognormal densities on the price as well as on the delta.

The remainder of this chapter is organized as follows. In section 2, we describe our algorithm in detail. In section 3, we examine numerical results for the lognormal case, using it as a benchmark for the efficiency of our method. Section 4 deals with non-lognormal densities. It examines the impact of various densities on the price of the option as well as on the delta. We conclude briefly in section 5 suggesting further developments.

3.2 Description of the method

3.2.1 Framework

We consider a continuous-time trading economy with infinite horizon. The uncertainty in the economy is classically modelled by a complete probability space (Ω, F, Q). The underlying is denoted by $(S_t)_{t\in\mathbb{R}^+}$. The information evolves according to the natural filtration $(F_t)_{t\in\mathbb{R}^+}$ implied by the underlying process. Following the traditional empirical literature, we assume that returns $(R_{t_i})_{i\in\mathrm{N}}$, defined by $R_{t_i} = \log\left(S_{t_i}/S_{t_{i-1}}\right)$ for a given sequence of time $(t_i)_{i\in\mathrm{N}}$, are independently distributed and have a well-known density $f_i(.)$, with a well-known mean denoted by μ_i. In the case of the Black Scholes model, each of these densities is a normal distribution with mean $\left(r - \frac{\sigma^2}{2}\right)(t_i - t_{i-1})$ and variance $\sigma^2 (t_i - t_{i-1})$. The underlying price is then calculated as the initial price S_{t_0} increased by the different returns $e^{R_{t_i}}$:

$$S_{t_i} = S_{t_0} e^{R_{t_1}+R_{t_2}+\ldots+R_{t_i}}$$

Assuming that we have n fixing dates for the average, denoted by t_1,t_2,...,t_n, the arithmetic average A is defined through:

$$A = \frac{1}{n}\sum_{i=1}^{n} S_{t_i} \tag{3.1}$$

In complete markets with no arbitrage opportunity, there is a unique risk neutral martingale measure denoted by Q. In this framework, the price P of an Asian call, with strike K, expiring at time T, is defined as the expected value of the time-T payoff discounted at the risk-free rate r :

$$P = \mathbb{E}^{Q}\left[e^{-rT}(A-K)^{+}\right] \tag{3.2}$$

where X^{+} stands for $\max(X,0)$. Since the discrete average process has no well-known density, there is no closed formula. However, we show in this chapter that we can compute numerically this density, giving a method which converges to the real densities as long as the size of the discretization grid tends to the infinity.

3.2.2 Why Fast Fourier Transform?

Well known in signal theory, Fast Fourier Transform (FFT) is efficient for the resolution of many numerical problems. More specifically, the FFT is an efficient algorithm for computing the sum:

$$FFT(f(k)) = \frac{1}{\sqrt{2\pi}}\sum_{j=1}^{N-1} e^{-i\frac{2\pi}{N}(j-1)(k-1)} f(j) \quad \text{for } k = 1...N$$

where N is typically a power of 2. This algorithm reduces the number of multiplications in the required N summations from $O(N^2)$ to that of $O(N\log_2(N))$. This suggests that for a grid with 2^p points, the complexity is $p2^p$, which is typically the complexity of a binomial tree.

Recently, this technique has gained popularity in option valuation (Baskhi and Chen (1998), Scott (1997), Chen and Scott (1992), Carr and Madan (1999)) in view of its numerical efficiency. The property of the Fourier transform used here is its efficiency to calculate convolution products. The Fourier transform of such a product is simply the product of the Fourier transforms. This is helpful in getting the density of the sum of two variables since this is just the convolution product of the individual densities as long as the variables are independent. In the case of the Asian option, the expression involved is not a straightforward sum of independent variables. In the algorithm section, we show how to use independent variables in a recursive scheme.

The interest of this method is its efficiency compared to a straightforward computation of the density. Instead of computing an $n-1$ dimensional integral with a complexity of $O\left(N^{n-1}\right)$, we reduce this complexity by means of Fast Fourier Transform to $O\left(N^2 \log\left(N\right)\right)$.

The use of FFT method for Asian option valuation was first suggested by Carverhill and Clewlow (1992). However, their work assumes lognormal densities and is not very efficient since it requires large grid and converges rather slowly. To speed up convergence, one needs to reduce the size of the grid required by the FFT algorithm. To cope with smaller grid, we introduce a proxy for the mean of intermediate densities. This enables us to recenter the different variables. We extend as well the FFT method to non-lognormal densities. We look particularly on the Student-density case since the latter is a well-known example of a fat-tailed distribution. We use the FFT algorithm as described in Press et al. (1992). Indeed, the method explained here is very general and can be applied to many other fat-tailed densities, like extreme value, Pareto and generalized Pareto distributions.

3.2.3 Algorithm

Inefficiency of the Carverhill and Clewlow method

A simple way to calculate the density of a sum of dependent variables is to transform then into independent variables. With our assumptions on the independence of returns, this comes naturally. Notice that when the underlying distribution is lognormal, returns are normal and their Fourier transform has a closed form solution equal to $f(w) = \frac{1}{\sqrt{2\pi}} e^{\left(iwm - \frac{w^2\sigma^2}{2}\right)}$, where m stands for the mean and σ^2 the variance. We introduce the sequence $(B_i)_{i=0..n-1}$ defined by its initial condition: $B_1 = R_{t_n}$ and for the recursion $i = 2...n$,

$$B_i = R_{t_{n+1-i}} + \log\left(1 + \exp B_{i-1}\right) \tag{3.3}$$

The Steward and Hodges factorization expresses the sum variable A defined by (3.1) in terms of the variable B_n as stated in the following proposition:

Proposition 6 *The sum variable A can be expressed in terms of the last term of the sequence $(B_i)_{i=0..n-1}$: B_{n-1}; through :*

$$A = \frac{S_{t_0}}{n} e^{B_n}$$

Proof : We decompose the underlying price as a function of the difference of returns: $S_{t_i} = S_{t_0} e^{R_{t_1} + R_{t_2} + ... + R_{t_i}}$. Factoring terms leads to a multiplicative expression of the sum variable:

$$A = \frac{S_{t_0}}{n}\left[e^{R_{t_1}} * \left(1 + e^{R_{t_2}} * \left(1 + e^{R_{t_3}}\left(1 + ... \left(1 + e^{R_{t_n}}\right)\right)\right)\right)\right]$$

When taking the logarithm of the above equation, we get an additive expression:

$$A = \frac{S_{t_0}}{n} * \exp\left(R_{t_1} + \log\left(1 + \exp\left(R_{t_2} + \log\left(... + \log\left(1 + R_{t_n}\right)\right)\right)\right)\right)$$

The term inside the outermost exponential can be calculated recursively using the sequence $(B_i)_{i=1..n}$.□

The Proposition 6 together with the recursive equation (3.3) was the starting point of the work of Carverhill and Clewlow (1992). At the i^{th} step of the recursive equation (3.3) the return $R_{t_{n+1-i}}$ is added. The latter is, however, not centred and has often a positive mean which for high volatilities can become negative (see the expression for the mean $\left(r - \frac{\sigma^2}{2}\right) * \frac{T}{n}$). For positive mean, the distribution of B_{i+1} is consequently shifted to the right of the distribution of B_i. If we discretize the distribution of B_{i+1} on the same grid as the one of B_i, this implies that the discretization grid must be large enough to contain the two

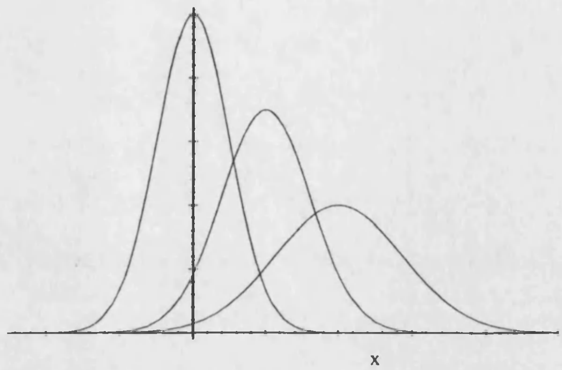

Figure 3.1: Evolution of the densities

distributions. When we have n dates in our arithmetic average, this tends to shift more and more in one direction as the order of the distribution increases as shown in figure 3.1. This is precisely why the algorithm of Carverhill and Clewlow requires a large grid.

Recentering intermediate densities

To cope with a smaller grid and therefore reduce computational time, we can recenter the densities at each step. The difficulty here is that we do not know the exact mean of the variable B_i. Denoting by μ_i the mean of the return R_{t_i} ($\mu_i = \mathbb{E}[R_{t_i}]$), which is supposed to be known, we can approximate the mean of the variable B_i with the following sequence: $(m_i)_{i=1..n}$ initialized with $m_1 = \mu_n$ and for $i = 2...n$

$$m_i = \mu_{n+1-i} + \log(1 + \exp m_{i-1}) \tag{3.4}$$

The term m_i acts as a proxy for the mean of the variable B_i. The approximation of the average is done by taking the lagged B_{i-1} equal to its mean m_{i-1} in the recursive equation (3.3). It is worth noticing that even if we do an approximation on the mean, it does not mean that we approximate the density of B_i. It just means that we do not perfectly center this variable. However, there is no new error implied by the recentering. Indeed, since the function $\log(1+e^x)$ is convex, we are underestimating some convexity adjustment term as stated by the Jensen inequality for convex functions $f(\mathbb{E}(X)) \leq \mathbb{E}(f(X))$.

The recentered sequence is defined as $(A_i)_{i=1..n}$ with $A_i = B_i - m_i$. Replacing B_{i-1} by its expression in terms of A_{i-1} and m_{i-1} leads to a recursive two dimensional sequence summarized by the following proposition:

Proposition 7 *The sum variable A can be expressed in terms of the last term of the recursive sequence A_n and m_n : as follows:*

$$A = \frac{S_{t_0}}{n} e^{A_n + m_n}$$

where the sequence $(m_i)_{i=1...n}$ is defined as above (3.4) and the sequence $(A_i)_{i=1..n}$ is given by the initial condition $A_1 = R_{t_n} - m_1$ and for $i = 2..n$

$$A_i = R_{t_{n+1-i}} + \log(1 + \exp A_{i-1} \exp m_{i-1}) - m_i$$

To get the density of A_i with respect to the one of A_{i-1}, we use the standard change of variable theorem, which relates the density of a variable $Y = g(X)$, denoted by dq_Y, with the one of the variable X, denoted by $dp_{f^{-1}(y)}$, through the Jacobian of the function $f^{-1}(Y) = X$

$$dq_Y = dp_{f^{-1}(x)} \left| J_{f^{-1}(y)} \right| dy$$

leading to the interpolation formula:

$$dp_{\log\left(1+e^{m_{i-1}+X}\right)-m_i}(y) = \frac{e^{y+m_i}}{e^{y+m_i}-1} p_X \left(\log\left(e^{y+m_i}-1\right) - m_{i-1}\right) 1_{\{y>-m_i\}} dy \tag{3.5}$$

We can now describe the different steps of the first algorithm. The algorithm is initialized with the value of the two dimensional sequence $m_1 = \mu_n$ and $A_1 = R_{t_n} - m_1$. It finishes when we get m_n and A_n.

The recursive sequence is calculated as follows. Assume that we know the value of the bi-dimensional sequence at step $i-1$, that is m_{i-1} and A_{i-1}.

- We then interpolate the variable A_{i-1} by means of the remark (3.5) to get the density of the variable $\log\left(1+e^{m_{i-1}+A_{i-1}}\right)-m_i$.

- We calculate the density of A_i as the sum of the two independent variables $R_{t_{n+1-i}}$ and $\log\left(1+e^{m_{i-1}+A_{i-1}}\right)-m_i$ by calculating the convolution product via FFT.

- Having obtained the density of the average, we calculate the payoff of the option, defined as an expectation, by a numerical integration, using the Simpson rule.

Discussion of the numerical techniques

The FFT algorithm requires the density function to be represented at a sufficient number of equally spaced points. The grid for the discretization of the different densities needs to be sufficiently dense as well as sufficiently large to avoid interference errors implied by the periodisation of the density function in the FFT algorithm. We use the FFT algorithm as described in Press et al. (1992)

Errors in the numerical integration by the Simpson rule (exact for the integration of polynom up to degree 3) are negligible compared to the ones produced by the discretization of the distribution. The error in the Simpson rule for the integral of a function f infinitely differentiable $\int_a^b f(x)\,dx$ can be shown to be an $O\left(\left(\frac{b-a}{2}\right)^5 f^{(4)}\right)$.

3.3 Efficiency of the algorithm for the lognormal case

3.3.1 Black Scholes assumption

The information evolves according to the augmented filtration $\{F_t, t\in[0,T]\}$ generated by a standard one-dimensional standard Brownian motion $(W_t)_{t\in R_+}$.

We assume the underlying price process is a geometric Brownian motion, solution of the Black Scholes (1973) diffusion defined by equation (3.6) with initial condition $S_{t=0} = S_0$

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t \tag{3.6}$$

In this case the returns R_{t_i} have a normal density with mean $\left(r - \frac{\sigma^2}{2}\right)(t_i - t_{i-1})$ and variance $\sigma^2 (t_i - t_{i-1})$.

3.3.2 Choice of the Grid

The choice of an efficient grid is not easy. The grid is determined by its range as well as its number of points. Choosing a range not correctly leads to interference errors. Taking a grid not dense enough leads as well to inaccurate Fourier Transform computation. We choose a centred grid with 4096 points, that is 2^{12} and with a width of $9n\sigma\sqrt{dt}$, where n stands for the number of fixing dates, σ the volatility, dt the time between two fixings. For a one year weekly Asian option, with fifty fixings, the number of fixing n is equal to 50 and the period of time between two successive prices dt is equal to one week or $1/52$ of a year.

Recentering the densities

The improvement of this chapter is to recenter densities at each step. Since we approximate the mean, the recentering is imperfect as figure 3.2 shows. For low volatilities up to 20%, densities are perfectly recentered for a one-year weekly Asian option. For volatilities higher than 20%, the approximation of the mean is not rigorously correct and leads to a shift of the different densities to the right. Indeed, since the function $\log(1 + e^x)$ is convex, we are underestimating a convexity adjustment term, as stated by the Jensen inequality for a convex function f, $f(\mathbb{E}(X)) \leq \mathbb{E}(f(X))$. However, the bias in our estimation is quite small, since for large values of x, the function $\log(1 + \exp(x))$ is very little convex, roughly equal to x, justifying our method.

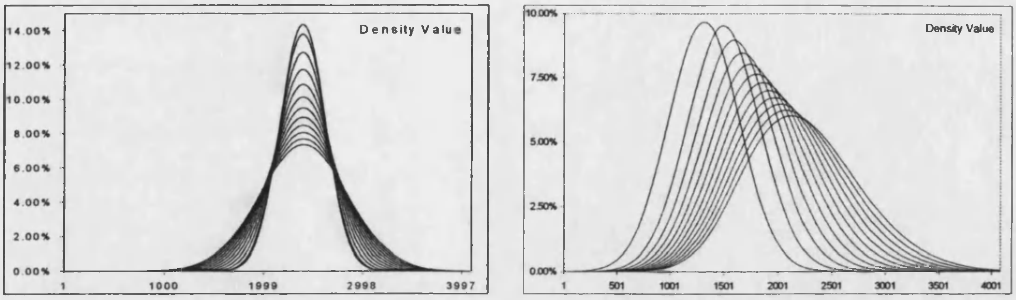

Figure 3.2: Evolution of the density with recentering at each step. The two graphics concern a one year weekly Asian option. The figure on the right is with 30% of volatility whereas the one on the left is for 20% volatility

In the figure 2, we can see that for small volatility level (20%, figure on the left), the recentering is perfect whereas for higher volatility (30%, figure on the right), we are missing the convexity adjustment term. In the original $(9n\sigma\sqrt{dt}$ithm of Caverhill and Clewlow, the grid size can be shown to be equal to $9n\sigma\sqrt{dt}+\mu_n$. The gain in our method can be measured by the grid width ratio $\left(9n\sigma\sqrt{dt}+m_n\right)/9n\sigma\sqrt{dt}$. For the case of a one year option with 10% volatility and a risk-free rate of 20%, this gain is equal to 1.317. This means that with the old algorithm, we need 1317 points to get the same precision as 1000 points with the new one. This means that the equivalent of 4096 points with the new algorithm is about 5400 points with the old algorithm.

Interference on the FFT algorithm

When the grid is not large enough, interference alters the results'quality as shown in figure 3.3, where we used a grid width of only $4n\sigma\sqrt{dt}$. This comes from the fact that the FFT algorithm assumes the periodicity of our function. It can cause interference terms when the grid size is too small.

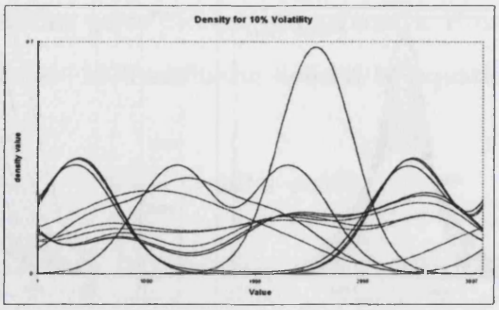

Figure 3.3: Interferences on the densities

3.3.3 Comparison of the different methods

Because of no well-known example, we arbitrary decided to use as a benchmark the same option example as in the work of Levy (1992) . We compute the price of a one year Asian option, with the underlying starting at 100 ($S = 100$), with a risk-free interest rate of 10% ($r = 10\%$), and 50 fixings per year (weekly average with two weeks of holidays).

The results, given in the table 3.1, compare different methods and show that the convolution method is efficient for the pricing of Asian option. Regarding the column titles, MC stands for Monte Carlo with its standard error given in the next column SE. WE means Wilkinson Estimates, E Edgeworth method, RG the reciprocal Gamma approximation, CV the Convolution method of Caverhill and Clewlow, CVR the convolution method with recentering. The reference price is the one of the Monte Carlo simulation. The efficiency of a formula is given by its comparison with this reference price.

We found that recentering the density improves significantly the efficiency of the Fast Fourier Transform method for high volatilities since the estimation of the density becomes more important. Among the traditional approximation methods, we tested Wilkinson estimates, Edgeworth expansion, and the reciprocal gamma approximation. We found that Wilkinson estimates was the most robust method. The Edgeworth expansion formula can blow up when the third

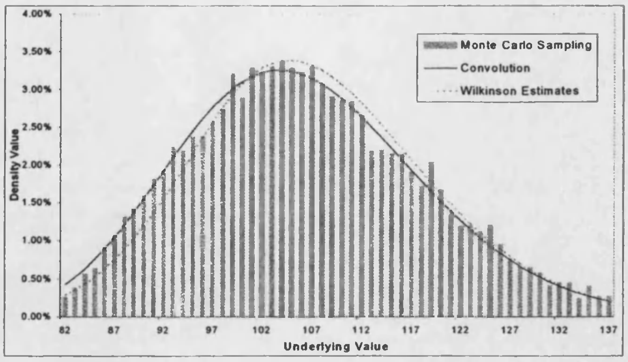

Figure 3.4: Comparison of the different densities for a one year Asian option for 30 % volatility

and fourth moments are too different from the ones implied by a lognormal. We also got poor results for the reciprocal gamma approximation. This comes from the small number of variables in our Asian options. The density of the average is therefore far from its asymptotic limit, which can be shown to be a reciprocal gamma density (see Milesvky and Posner (1997)).

3.3.4 Density Comparison

Our results confirm that the lognormal approximation slightly overprices Asian options (Levy and Turnbull (1992), Zhang (1998)). This is indicated by the skew to the right of the Wilkinson estimates (or lognormal approximation) density in figure 3.4. The efficiency of the FFT method is confirmed by the close fit with the Monte Carlo sampling in figure 4. The Monte Carlo sampling was based on a simulation of a Sobol sequence with 30,000 draws.

It is worth noting that the precision of the method is heavily depending on the type of the options: in, at or out-of-the-money. One should expect little difference in price for options depending on a wide part of the distribution like in or at-the-money options. However, for out-of-the-money options, that are depending mainly on the tails of the distribution, there is a real advantage

σ	K	MC	Std Err	WE	E	RG	CV	CVR
10%	80	22.78	0.00	22.78	22.78	21.64	22.78	22.78
	90	13.73	0.00	13.73	13.73	13.1	13.73	13.73
	100	5.24	0.00	5.25	5.25	4.98	5.25	5.25
	110	0.72	0.00	0.72	0.72	0.71	0.72	0.72
	120	0.03	0.00	0.02	0.02	0.02	0.02	0.02
30%	80	23.07	0.01	23.14	23.07	21.92	23.09	23.08
	90	15.22	0.01	15.30	15.16	14.46	15.29	15.26
	100	9.01	0.01	9.08	9.00	8.56	9.08	9.05
	110	4.83	0.01	4.84	4.85	4.58	4.86	4.84
	120	2.35	0.01	2.33	2.40	2.23	2.40	2.33
50%	80	24.83	0.03	25.06	24.10	23.58	25.01	24.88
	90	18.32	0.03	18.57	17.83	17.40	18.50	18.37
	100	13.18	0.03	13.34	13.02	12.52	13.47	13.20
	110	9.23	0.03	9.33	9.36	8.77	9.45	9.19
	120	6.36	0.03	6.37	6.63	6.04	6.68	6.40

Table 3.1: Comparison of different methods for the Asian option σ stands for the volatility, K for the strike

in terms of precision to use the Fast Fourier Transform method compared to Wilkinson estimates. Indeed, fat tails are the true motivation of this chapter. It is already interesting to realize that even in the case of a lognormal underlying, the Fast Fourier Transform method takes better account for fat tails than most standard approximation methods with closed form.

3.4 Using non-lognormal densities

3.4.1 Interest of the method

It is now widely accepted that markets differ from the seminal Black Scholes (1973) lognormal model. The empirical literature has extensively reported on these anomalities, especially on two of them, which indeed are closely linked. First, it is has been shown that unconditional returns show excess kurtosis and skewness, inconsistent with normality assumptions (see Mandelbrot (1963) and Fama (1965) for the early ones, Kon (1984), Jorion (1988) and Bates (1996)). Second, research has concentrated its attention on the implied volatility smile or skew (see Dumas et al. (1995) for a survey). Interestingly, the second fact is just another hint of the non-normality of returns. However, research has focussed at implied Black Scholes volatility since implied volatility has become a key concept in option pricing. Option prices are often quoted by their implied volatility. A more rigorous justification of the interest in modelling volatility is its less volatile character when compared with prices. Since, corresponding prices fluctuate more than implied volatilities, the trading environment is best captured by a model about the implied volatility.

How to cope with the smile in option pricing has become an extensive field of research. Classically, it is divided into two different approaches: parametric and non-parametric ones.

In the first method, the equation of the evolution of the underlying process is given. This description can consist in a continuous diffusion process with either a so called deterministic volatility (Rubinstein (1994), Dupire (1993) and Derman and Kani (1994)) or a stochastic volatility process (Hull and White (1987), Wiggings (1987), Melino and Turnbull (1990), Stein and Stein (1991), Amin and Ng (1993) and Heston (1992)) or a model with jumps (Aase [1993, Ahn and Thompson (1988), Amin (1993), Bates (1991), Jarrow (1984), Merton (1976)).

Other works close in spirit are assuming constant elasticity of volatility distribution often called power-law (Cox Ross (1976)) or a mapping principle between normal and lognormal distributions (Hagan (1998), Pradier and Lewicki (1999)).

The second type of methods involves infering the underlying distribution from market data. This has been called the expansion method where one induces the different terms of the expansion and can reconstitute the distribution (Jarrow and Rud (1982), Bouchaud et al.(1998), Abken et al. (1996)).

The interest of our methodology lies in its flexibility on the distributions of returns. We do not assume any specific distribution. The distribution is an input like all other parameters. Therefore, we can use distribution derived from market data, like option prices. In this chapter, we decided to illustrate the fat-tailed distribution with the specific case of a Student density. This is because this density is often used in the literature. It has the additional advantage to converge to the normal density when the number of degree of freedom tends to infinity. Indeed, there are many other densities which could have been used, like Pareto, generalized Pareto, power-laws distributions and many more.

3.4.2 Densities for leptokurtic effect

To account for leptokurtic returns, we assume that centred and pseudo-normalized (with a parameter $\lambda > 1$) returns $\frac{R_{t_i} - \left(r - \frac{\sigma^2}{2}\right)(t_i - t_{i-1})}{\sqrt{\frac{\sigma^2 (t_i - t_{i-1})}{\lambda}}}$ have a density given by a Student distribution with a degree of freedom $n = \frac{2\lambda}{\lambda - 1}$ given by $\frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{\pi n}} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}$.

The cumulative distribution is then given by

$$\Pr(X \leq t) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{\pi n}} \int_{-\infty}^{t} \left(1 + \frac{1}{n}u^2\right)^{-\frac{n+1}{2}} du$$

where $\Gamma(y) = \int_0^\infty e^{-x} x^{y-1} dx$ is the Gamma function at y. Since a Student density has always a variance bigger than one we need to specify this variance by the parameter λ.

3.4.3 Numerical results

Effect on the price

As expected, fat-tailed distributions hereby illustrated by the Student density lead to a more expensive price of the Asian option. The Fast Fourier method is efficient as confirmed by a comparison with Monte Carlo simulations with 20,000 draws. To simulate the Student density, we simulate uniform distribution and inverse the cumulative distribution by means of the approximation given in the appendix section D.

Without any surprise, the discrepancy between the lognormal distribution and a distribution with fat tails increases with the volatility. It also grows for distribution with fatter tails as shown by the increase of price between the Student density with 44 degrees of freedom and the one with only 22. We have chosen the Student density since its asymptotic distribution is precisely the normal distribution when the degrees of freedom tend to infinity.

Interestingly, practitioners have kept on using the lognormal approximation for the Asian option. We have seen that the approximation of a sum of lognormal by a lognormal distribution is not correct. It tends to overprice the Asian option. However, when assuming a fat-tailed distribution for the underlying, we also found that the price of the option was more expensive than the corresponding one with lognormal individual underlyings. This explains why practitioners have been very keen on using the lognormal approximation since this includes the rise of price due to fat-tailed distributions.

σ	K	Lognormal	Student 44 df	MC Student 44 df	Student 22 df	MC Student 22 df
10%	80	22.7838	22.7911	22.7914	22.8021	22.8028
	90	13.7347	13.7420	13.7425	13.7539	13.7550
	100	5.2438	5.2843	5.2850	5.3278	5.3294
	110	0.7211	0.7642	0.7649	0.8078	0.8094
	120	0.0336	0.0358	0.0362	0.0423	0.0430
30%	80	23.0733	23.2033	23.2050	23.3372	23.3406
	90	15.2231	15.4014	15.4036	15.5808	15.5855
	100	9.0110	9.2238	9.2277	9.4335	9.4416
	110	4.8338	5.0345	5.0379	5.2355	5.2426
	120	2.3545	2.5097	2.5117	2.6682	2.6728
50%	80	24.8324	25.2213	25.2243	25.6099	25.6168
	90	18.3207	18.7471	18.7510	19.1694	19.1779
	100	13.1811	13.6182	13.6239	14.0508	14.0628
	110	9.2300	9.6476	9.6530	10.0610	10.0721
	120	6.3615	6.7363	6.7412	7.1104	7.1209

Table 3.2: Price of Asian option with Fat-tailed distributions. σ stands for the volatility, K for the strike, df for degrees of freedom

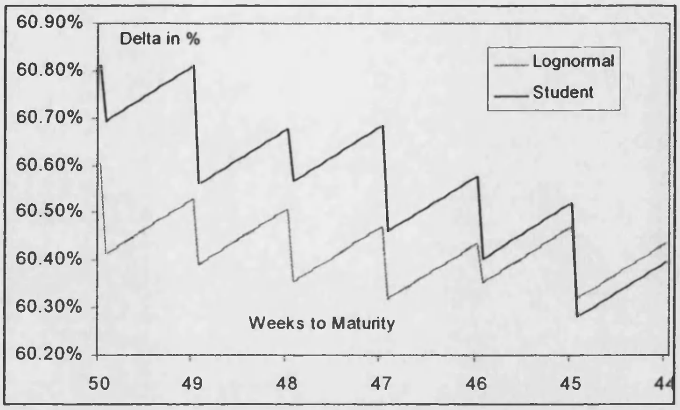

Figure 3.5: Evolution of the delta with time to maturity under different distributions

Delta hedging

The motivation for our numerical method is to examine the impact of fat-tailed distributions on the delta. In the case of the discrete Asian options, the delta jumps every time we cross a fixing date.

The comparative study of the delta evolution with lognormal density and the Student density shows that a fat-tailed distributions lead to higher jumps in the delta, a logical consequence of the fact that fat-tailed distributions imply more expensive prices and therefore larger drop of the price with the downfall of a fixing date. The difference in the delta is quite significant as shown by figure 3.5 and 3.6. The figure 3.5 show the evolution of the delta for a weekly Asian option far from the maturity of the option. The Student density taken here is the one with 22 degrees of freedom.

There is no rule concerning the difference between the delta for lognormal densities and for Student densities. In the figure 3.5, the option is 50 to 44 weeks before the expiration. In this particular case, the delta implied by the Student density is on overall more expensive. This is not the case when the option is close to the maturity as shown by figure 3.6 where there are only 10 to 1 week before

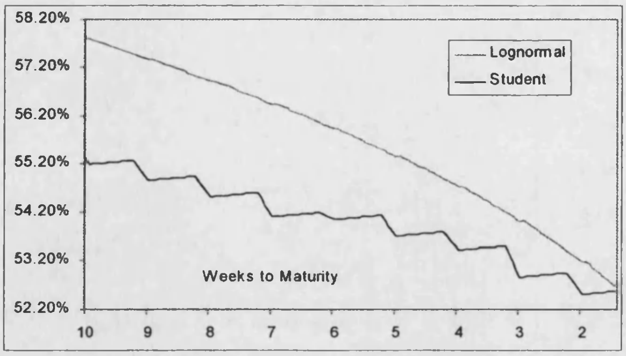

Figure 3.6: Evolution of the delta with time to maturity under different distributions when closed to the expiry

the maturity of the option. However, it is worth noting that on the average the delta is almost the same for the two densities. This suggests that for a long-run delta hedging, assuming normal returns is not too much inaccurate. However, for short run delta hedging, the assumptions on the return densities lead to very different hedging strategies.

3.5 Conclusion

In this chapter, we have seen that Fast Fourier Transform is an efficient way for pricing discrete Asian options with non-lognormal densities. The systematic recentering of intermediate densities enables to reduce the size of the grid so as to speed up the convergence. We show that the price of the Asian option should be more expensive with fat-tailed distributions. This indicates that approximation methods overpricing the Asian option incorporate, in a way, fat-tailed distribution. However, as far as the delta is concerned, fat-tailed distributions lead to very different hedging strategies, especially on the short run.

Our methodology raises many remarks. First, the Fast Fourier Transform technique enables to take into account volatility smile since, as an input, we can take returns'distribution derived by market data incorporating the smile effect. Second, the same approach can be applied with minor changes to basket and multi-asset options. Third, this methodology raises the issue of the way of deriving the density from market data properly.

Chapter 4

An Application of Malliavin Calculus to Continuous Time Asian Options Greeks

Summary of the chapter

This chapter extends the results on the Malliavin weighted scheme as described in the first chapter to the case of continuous Asian options. We give the necessary and sufficient conditions for a function to serve as a weight function introducing its generator. We discuss in greater detail the case of an option depending both on the continuous average as well as the maturity value of the underlying. We refer to this difficult case as the "complex Asian option". We conjecture indeed that these results should be adaptable to the case of the continuous lookback options as well.

4.1 Introduction

When handling sophisticated models with non standard and discontinuous payoff options, classical methods like lattice methods and numerical methods for solving partial differential equations like finite differences and finite elements, could be inefficient. The Monte-Carlo and Quasi-Monte-Carlo methods, often seen as last resort methods can overcome this technical difficulty. However, in the case of a strongly discontinuous payoff function, a well-known fact is their poor convergence to the exact solution, when computing the Greeks (price sensitivity to parameters like delta, gamma, rho and vega).

Traditionally, to speed up convergence, one relies on different more or less successful variance reduction techniques among which the most famous ones are antithetic variates, control variates, importance sampling, stratified sampling, Latin hypercube sampling and moment matching techniques. One uses as well deterministic methods based on low discrepancy sequences like Halton, Sobol, Faure sequences (see Glasserman and Yao (1992), Glynn (1989), and L'Ecuyer and Perron (1994)). Their straightforward use provide nonetheless little improvement when handling the Greeks for strong non-linear payoff functions.

The reason of this inefficiency lies in the way the Greeks are commonly computed. One estimates the Greeks by simply taking the finite difference of two particular simulation results. Denoting by $P(x)$ the option price for an underlying level of x, the different schemes can be classified into forward difference $(P(x+\varepsilon) - P(x))/\varepsilon$, central difference $(P(x+\varepsilon) - P(x-\varepsilon))/2\varepsilon$, or backward difference scheme $(P(x) - P(x-\varepsilon))/\varepsilon$. Despite a quasi antithetic technique implied by the substraction of terms $P(x+\varepsilon)$, $P(x)$, and $P(x-\varepsilon)$, this method embodies two different errors:

- discretization of the derivative function by a finite difference.
- imperfect estimation of the option prices $P(x+\varepsilon)$, $P(x)$, and $P(x-\varepsilon)$.

For discontinuous payoff functions, the main error is the first one. Recently, Fournié et al. (1999) (2000) and Benhamou (2000a) (rewritten as chapter 1 of this dissertation) have suggested a new methodology based on Malliavin calculus. The intuitive idea is to eliminate the need of taking the derivative of the payoff function, which is numerically approximated by a finite difference. They showed that we can transform the initial formula as an expectation of the discounted payoff function $e^{-\int_0^T r_s ds} f(X_T)$ multiplied by a suitable weight function referred in the literature as the Malliavin weights, denoted hereby *weight*:

$$Greek = \mathbb{E}^Q \left[weight.e^{-\int_0^T r_s ds} f(X_T) \right] \tag{4.1}$$

Their works were for options depending on a finite set of dates. The contribution of this chapter is to extend previous results to continuous-time Asian options. We give necessary and sufficient conditions for a function to serve as weight function by means of its generator. We give explicit solutions for the case of the Black diffusion.

The remainder of this chapter is organized as follows. In section 2, we laconically describe the mathematical framework of the chapter. In section 3, we investigate the case of Asian options. In section 4, we give explicit formula for the case of the Black diffusion. Section 5 shows by means of numerical results the efficiency of this method. We briefly conclude in section 6 giving possible extensions.

4.2 Notations and mathematical framework

For clarity, we assume a one dimensional diffusion of the underlying price process. Results can be easily extended to the multi-dimensional case. We consider a continuous-time trading economy with a limited horizon $t \in [0, T]$. Following Harrison and Kreps (1979), Harrison and Pliska (1981), the price of a contingent claim is calculated as the expected value of the discounted payoff value in the risk neutral probability measure Q. The uncertainty is characterized by a complete probability space (Ω, F, Q) where Ω is the state space, F is the σ-

algebra representing the measurable events. Information evolves according to the augmented right continuous complete filtration $\{F_t, t \in [0,T]\}$ generated by a standard one dimensional Brownian Motion $\{W_t, t \in [0,T]\}$. The underlying price process is defined as the Ito process $(X_t)_{t\in[0,T]}$ solution of the following stochastic differential equation:

$$dX_t = b\,(t, X_t)\,dt + \sigma(t, X_t)dW_t \tag{4.2}$$

with the initial condition

$$X_0 = x \tag{4.3}$$

$x \in \mathbb{R}$ is the initial value of our underlying. $b : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ represents the deterministic drift and $\sigma : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ the volatility structure of the process $(X_t)_{t\in[0,T]}$. We also assume that the deterministic drift and the volatility structure verify Lipschitz conditions and uniform ellipticity of σ.

We define the first variation process $(Y_t)_{t\in[0,T]}$ as the derivative of X_t with respect to the initial condition $Y_t = \frac{\partial}{\partial x}X_t$. As show in Nualart (1995) page Theorem 2.3.1 page 110, Malliavin calculus theory proves that the Malliavin derivative can be expressed as a function of the first variation process and the volatility structure $\sigma\,(t, X_t)$:

$$D_s X_t = \frac{Y_t}{Y_s}\sigma(s, X_s)1_{\{s\leq t\}} \;\; a.s. \tag{4.4}$$

4.3 Asian options

Since there is no closed formula for arithmetic Asian options as opposed to geometric ones, Asian options, sometimes called average options, are a prefered field for numerical solutions such as lattice methods, partial differential equations solving, convolution, Monte-Carlo and Quasi -Monte-Carlo methods (for an extensive survey on Asian options, see chapter 3 of this dissertation). However, these numerical methods have often low convergence for the Greeks. It is precisely this inefficiency that suggested us to extend the work of Fournié et al. and Benhamou to the case of the continuous-time Asian option.

Traditionally, the average in the Asian options can be computed as a discrete or continuous-time one. In the case of discrete Asian options, we can apply the analysis done by Benhamou (2000a) for a fast computation of the Greeks. In the case of continuous-time arithmetic Asian options, Fournié et al. (1999) gave a specific example of the weighting function generator for the delta. We provide here formulae for other Greeks with more general conditions and for any type of weight functions, specifying only the necessary and sufficient conditions for a function to serve as a weight function in derived formulae. We introduce the weight function generator. We give a detailed version in the case of the delta, generalizing to the case of the other Greeks. In the rest of this section, we also distinguish as well two type of Asian options:

- the simple Asian option, an option whose payoff is a function of the continuous-time average only. Classically, for a call, it is called the fixed strike option.
- the complex Asian option is depending on both the underlying and the continuous-time average. Classically, for a call, it is the floating strike option.

4.3.1 Simple Asian option

The simple continuous-time Asian option is an option only depending on the continuous-time average $\frac{1}{T}\int_0^T X_t dt$. Thus, its price can be written as the expected value of the discounted payoff:

$$P_x = \mathbb{E}_x\left[e^{-\int_0^T r_s ds} f\left(\int_0^T X_t dt\right)\right]$$

Like in the case of an option depending on a finite set of dates, we can show that the Greeks can be written as the expected value of the discounted payoff $f\left(\int_0^T X_t dt\right)$ times a suitable weight function *weight*:

$$Greeks = \mathbb{E}_x^Q\left[f\left(\int_0^T X_t dt\right) weight\right]$$

We also impose the existence of the Skorohod integral expressed as a non exploding condition (L_2 integrable) on the weighting function (4.5):

$$\mathbb{E}\left[(weight)^2\right] < +\infty \tag{4.5}$$

Writing the weight function *weight* as the Skorohod integral $\delta(w)$ of a stochastic function $w : \mathbb{R}_+ \to \mathbb{R} \quad s \to w_s$, we define the Skorohod integrand "the weight function generator". Interestingly, this generator can be characterized by necessary and sufficient conditions as stated in theorem 2.

Theorem 2 *Malliavin formula for the Greeks*
There exist necessary and sufficient conditions for a function w to serve as a weight function generator for the simulation of the Greeks. The first condition is the Skorohod integrability of this function (4.5). The second condition, different for each Greek and summarized in table 4.1 depends only on the underlying diffusion characteristics and not on the payoff function.

Proof: we give in the appendix section the proof for the delta section C.1 page 176. Proofs for the other Greeks are similar and are available upon request.□

4.3.2 Particular solutions

To get particular solution, we only need to find solutions that satisfy the conditions given in table 4.1. One can show the following results by checking that these solutions satisfy the necessary and sufficient conditions of table 4.1:

- Delta: one solution for the weight function generator is the one given by Fournié and al. (1999):

$$w_s = \frac{2Y_s^2}{\sigma(s, X_s)\int_0^T Y_t dt} \tag{4.6}$$

Greeks	Necessary and Sufficient conditions on the Malliavin Weights
delta	(4.M1) : $\mathbb{E}_x^Q\left[\int_{t=0}^T Y_t\left(\int_{s=0}^t \frac{\sigma(s,X_s)}{Y_s} w_s^{delta} ds\right) dt \mid \int_0^T X_t dt\right] = \mathbb{E}_x^Q\left[\int_0^T Y_t dt \mid \int_0^T X_t dt\right]$
gamma	(4.M2) : $\mathbb{E}_x^Q\left[\delta\left(w^{gamma}\right) \mid \int_0^T X_t dt\right]$ $\mathbb{E}_x^Q\left[\delta\left(w^{delta}\delta\left(w^{delta}\right)+\frac{\partial}{\partial x} w^{delta}\right) \mid \int_0^T X_t dt\right]$
"extended" rho	(4.M3) : $\mathbb{E}_x^Q\left[\int_{t=0}^T \int_{s=0}^t \frac{Y_t\sigma(s,X_s)}{Y_s} w_s^{rho} dtds \mid \int_0^T X_t dt\right] = \mathbb{E}_x^Q\left[\int_{t=0}^T \int_{s=0}^t \frac{Y_t\tilde{b}(s,X_s)}{Y_s} dsdt \mid \int_0^T X_t dt\right]$
"extended" vega	(4.M4) : $\mathbb{E}_x^Q\left[\int_{t=0}^T \int_{s=0}^t \frac{Y_t\sigma(s,X_s)}{Y_s} w_s^{vega} dsdt \mid \int_0^T X_t dt\right] = \mathbb{E}_x^Q\left[\left(\begin{array}{c}\int_{t=0}^T \int_{s=0}^t \frac{Y_t\tilde{\sigma}(s,X_s)}{Y_s} dW_s dt \\ -\int_{t=0}^T \int_{s=0}^t \frac{Y_t\frac{\partial}{\partial X}\sigma(s,X_s)\tilde{\sigma}(s,X_s)}{Y_s} dsdt\end{array}\right) \mid \int_0^T X_t dt\right]$

Table 4.1: Necessary and Sufficient conditions for the weight function Generators of a simple Asian option

Another solution is:

$$w_s = \frac{Y_s}{\sigma(s, X_s)} \frac{\int_0^T Y_t dt}{\int_0^T t Y_t dt} \tag{4.7}$$

Depending on the nature of the first variation process as well as the structure of volatility, it would be preferable to use either the first or the second solution.

- Rho: a solution for the weight function generator of the rho is given by

$$w_s = \frac{\widetilde{b}(s, X_s)}{\sigma(s, X_s)} \tag{4.8}$$

- Vega: a solution for the weight function generator is given by

$$w_s = \frac{Y_s}{\sigma(s,X_s)} \frac{\int_{t=0}^T \int_{s=0}^t \frac{Y_t \widetilde{\sigma}(s,X_s)}{Y_s} dW_s dt}{\int_{t=0}^T t Y_t dt} - \frac{\frac{\partial}{\partial X}\sigma(s, X_s)\widetilde{\sigma}(s, X_s)}{\sigma(s,X_s)} \tag{4.9}$$

4.3.3 Complex Asian option

The complex Asian option is an option depending both on the continuous-time average $\frac{1}{T}\int_0^T X_t dt$ and the underlying price at maturity, X_T. Its price can be written as the expected value of its discounted payoff:

$$P_x = \mathbb{E}_x \left[e^{-\int_0^T r_s ds} f\left(\int_0^T X_t dt, X_T \right) \right]$$

We can extend previous results of theorem 2. We restrict ourselves to the case of the delta. With minor changes, this can be adapted to all other Greeks.

Proposition 8 *Malliavin formula for the delta*
In the case of the complex Asian option, in addition to the non-exploding condition, the weight function generator for the delta should verify the conditions

$$\begin{aligned}
\mathbb{E}_x^Q \left[\int_{t=0}^T Y_t \left(\int_{s=0}^t \frac{\sigma(s,X_s)}{Y_s} w_s^{delta} ds \right) dt | \int_0^T X_t dt, X_T \right] &= \mathbb{E}_x^Q \left[\int_0^T Y_t dt | \int_0^T X_t dt, X_T \right] \\
\mathbb{E}_x^Q \left[Y_T \int_{s=0}^T \frac{\sigma(s,X_s)}{Y_s} 1_{\{s \le T\}} w_s^{delta} ds | \int_0^T X_t dt, X_T \right] &= \mathbb{E}_x^Q \left[Y_T | \int_0^T X_t dt, X_T \right]
\end{aligned}$$

Proof: given in the appendix section C.2 page 177.□

We can find particular solutions. We introduced two different discriminants Δ and $\widetilde{\Delta}$

$$\Delta = \frac{\int_{t=0}^{T} tY_t dt}{\int_{t=0}^{T} Y_t dt} - \frac{1}{2}T \quad , \quad \widetilde{\Delta} = \frac{\int_{t=0}^{T}\int_{s=0}^{t} \frac{(t^2-T^2)}{2} sY_s ds Y_t dt}{\left(\int_{t=0}^{T} Y_t dt\right)^2} \tag{4.10}$$

Proposition 9 *Particular solution*

When the discriminant Δ is not equal to zero, one particular solution for the weight function generator is given by:

$$w_s = \frac{Y_s}{\sigma(s, X_s)} \left(A + B \frac{Y_s}{\int_0^T Y_t dt} \right) \tag{4.11}$$

with the coefficients A and B verifying:

$$A = \frac{\int_0^T Y_t dt}{\int_0^T (2t - T) Y_t dt} \quad B = \frac{2\int_0^T (t - T) Y_t dt}{\int_0^T (2t - T) Y_t dt} \tag{4.12}$$

This solution is unique.

Corollary 4 *When the discriminant $\widetilde{\Delta}$ is not equal to zero, one particular solution for the weight function generator is given by:*

$$w_s = \frac{sY_s}{\sigma(s, X_s)} \left(\alpha + \beta \frac{Y_s}{\int_0^T Y_t dt} \right) \tag{4.13}$$

with the coefficients α and β verifying:

$$\alpha = \frac{\int_{t=0}^{T}\int_{s=t}^{T} sY_s ds Y_t dt}{\int_{t=0}^{T}\int_{s=0}^{t} \frac{(t^2-T^2)}{2} sY_s ds Y_t dt} \quad \beta = \frac{\int_{t=0}^{T} \frac{(t^2-T^2)}{2} Y_t dt \int_{t=0}^{T} Y_t dt}{\int_{t=0}^{T}\int_{s=0}^{t} \frac{(t^2-T^2)}{2} sY_s ds Y_t dt} \tag{4.14}$$

Proof: given in the appendix section C.3 page 178.□

The two solutions differ in the form of their basis since in one case it is $Y_s/\sigma(s, X_s)$ and in the other case it is rather $sY_s/\sigma(s, X_s)$.

4.4 Formula for a Black diffusion

In the case of the Black diffusion, formulae can be explicit. In this framework, the drift term $b(t, X_t)$ is equal to $r_t X_t$ which is the growth at the risk-free rate whereas the stochastic term $\sigma(t, X_t)$ is equal to $\sigma_t X_t$, with σ_t a deterministic time-dependent volatility. This leads to the following expression for the underlying process $X_t = xe^{\left(r-\frac{1}{2}\int_0^t \sigma_s^2 ds\right)t+\int_0^t \sigma_s dW_s}$ and for the first variation process to $Y_t = X_t/x$.

4.4.1 Simple Asian option

With these more restrictive assumptions, we can derive results for the simple continuous Asian option (the complex one leads to the same type of calculations, with lengthier formulae). Being a second order Greek, the calculation of the gamma requires two integration by parts and is consequently not as straightforward as first order Greeks. However, it can be shown that there is a proportionality between the vega and the gamma (see Benhamou (2000c)). The Gamma is subsequently obtained as the vega times this proportionality factor.

Delta

Proposition 10 *The particular solution for the weight function generator $w_s^1 = 2X_s/x\sigma_s \int_0^T X_t dt$ leads to the following weight function*

$$\frac{2\int_0^T \frac{X_s}{\sigma_s} dW_s}{x\int_0^T X_t dt} + \frac{1}{x} \tag{4.15}$$

Proof: the weight function is defined as the Skorohod integral of the weight function generator

$$Weight = \delta\left(w_s^1\right)$$

Using the property of the Skorohod operator that gives the Skorohod integral of a product: if F is a smooth random variable (Nualart notation (1995)) and u is an element of $L^2(\Omega \times [0,T])$, then the Skorohod integral can be expressed by means of the Malliavin derivative:

$$\delta(uF) = \delta(u)F - \int_0^T uD_tF dt$$

This leads in our case to

$$Weight = \frac{2\delta\left(\frac{X_s}{\sigma_s}\right)}{x\int_0^T X_t dt} - \int_0^T \frac{2X_s}{x\sigma_s} D_s\left(\frac{1}{\int_0^T X_t dt}\right) ds$$

Using the fact that the Malliavin derivative follows regular derivation rules, we get:

$$D_s\left(\frac{1}{\int_0^T X_t dt}\right) = -\frac{\int_0^T Y_t Y_s^{-1}\sigma_s Y_s 1_{\{s<t\}} dt}{\left(\int_0^T X_t dt\right)^2}$$

where in the last equation, we have used the fact that $D_s X_t = Y_t Y_s^{-1}\sigma_s Y_s 1_{\{s<t\}}$. We finally get

$$Weight = \frac{2\int_0^T \frac{X_s}{\sigma_s} dW_s}{x\int_0^T X_t dt} + \frac{1}{x}$$

□

Proposition 11 *The other solution for the weight function generator* $w_s = \frac{1}{\sigma_s x}\int_0^T Y_t dt / \int_0^T tY_t dt$ *leads to the following weight function*

$$\frac{1}{x}\frac{\int_{t=0}^T Y_t dt}{\int_{t=0}^T tY_t dt}\left(\int_{s=0}^T \frac{dW_s}{\sigma_s} + \frac{1}{x}\frac{\int_0^T t^2 Y_t dt}{\int_0^T tY_t dt}\right)$$

Proof: following the same procedure as in the proof of the proposition (4.15), we calculate the weight function as the Skorohod integral of the weight function generator:

$$
\begin{aligned}
&\delta\left(\frac{1}{\sigma_s x}\frac{\int_0^T Y_t dt}{\int_0^T tY_t dt}\right) \\
= \;& \frac{1}{x}\frac{\delta\left(\frac{1}{\sigma_s}\int_0^T Y_t dt\right)}{\int_0^T tY_t dt} + \frac{1}{x}\int_0^T X_t dt \int_{s=0}^T \frac{1}{\sigma_s}\frac{\int_{t=0}^T tD_s X_t dt}{\left(\int_0^T tX_t dt\right)^2} ds \\
= \;& \frac{1}{x}\frac{\int_{t=0}^T Y_t dt \int_{s=0}^T \frac{1}{\sigma_s} dW_s}{\int_0^T tY_t dt} + \frac{1}{x}\int_0^T X_t dt \frac{\int_{s=0}^T \int_{t=0}^T \frac{tY_t Y_s^{-1}\sigma_s Y_s 1_{\{s<t\}}}{\sigma_s} dt ds}{\left(\int_0^T tX_t dt\right)^2} \\
= \;& \frac{1}{x}\frac{\int_0^T Y_t dt}{\int_0^T tY_t dt}\left(\int_{s=0}^T \frac{dW_s}{\sigma_s} + \frac{1}{x}\frac{\int_0^T t^2 Y_t dt}{\int_0^T tY_t dt}\right)
\end{aligned}
$$

□

Rho

Interestingly, the condition given for the rho of a continuous Asian option weight function generator is the same as the one for an option depending on a finite set of dates as in Benhamou (2000a). Therefore, the same results can be applied leading to the two following propositions:

Proposition 12 *The extended rho as defined in Benhamou (2000a) can be expressed as*

$$\widetilde{rho} = \mathbb{E}^Q\left[e^{-rT} f\left(\int_0^T X_t dt\right)\int_0^{t_m} \frac{\widetilde{r}_t}{\sigma_t} dW_t\right]$$

Proof: we obtain the result by considering the particular solution of the weight function generator (4.8) and by using the fact that the weight function is defined as the Skorohod integral of the weight function generator.□

Proposition 13 *The classical rho is given by:*

$$Classical\ rho = \mathbb{E}^Q\left[e^{-rT} f\left(\int_0^T X_t dt\right)\left(\int_0^{t_m} \frac{dW_t}{\sigma_t} - t_m\right)\right]$$

Proof: See Benhamou (2000a) (rewritten as chapter 1 of this dissertation).□

Vega

Using the particular solution for the weight function generator (4.9) and assuming a perturbation defined as $\widetilde{\sigma}(t, X_t) = \widetilde{\sigma}_t X_t$, we get the following proposition.

Proposition 14 *The extended vega can be expressed as*

$$\mathbb{E}^Q\left[e^{-rT} f\left(\int_0^T X_t dt\right)\left(\delta\left(\frac{1}{\sigma_s}\frac{\int_{t=0}^T \int_{s=0}^t \widetilde{\sigma}_s Y_t dW_s dt}{\int_{t=0}^T tY_t dt}\right) + \int_0^T \widetilde{\sigma}_s dW_s\right)\right]$$

Proof: in the case of the Black diffusion, the weight function generator is defined as

$$w_s = \frac{1}{\sigma_s}\frac{\int_{t=0}^T \int_{s=0}^t \widetilde{\sigma}_s Y_t dW_s dt}{\int_{t=0}^T tY_t dt} + \widetilde{\sigma}_s$$

leading to the following weight function:

$$\delta(w_s) = \delta\left(\frac{1}{\sigma_s}\frac{\int_{t=0}^T \int_{s=0}^t \widetilde{\sigma}_s Y_t dW_s dt}{\int_{t=0}^T tY_t dt}\right) + \int_0^T \widetilde{\sigma}_s dW_s$$

□

Corollary 5 *In the case of the Black Scholes model, this leads to the following results*

$$\mathbb{E}^Q\left[e^{-rT} f\left(\int_0^T X_t dt\right)\left(\frac{\widetilde{\sigma}}{\sigma}\frac{\int_{s=0}^T \left(\int_{t=0}^T X_t W_t dt\right) dW_s}{\int_0^T tX_t dt} + \frac{\widetilde{\sigma}}{x}\int_0^T X_t W_t dt \frac{\int_{t=0}^T t^2 X_t dt}{\left(\int_0^T tX_t dt\right)^2} + \widetilde{\sigma} W_T\right)\right.$$

The traditional vega is then obtained for $\widetilde{\sigma} = 1$.

Proof: the Black Scholes assumptions are that the volatility is constant: $\sigma_s = \sigma$. The perturbation is defined as $\widetilde{\sigma}_s = \widetilde{\sigma}$. This leads to the following weight function

$$\begin{aligned} weight &= \frac{\widetilde{\sigma}}{\sigma}\delta\left(\frac{\int_0^T Y_t W_t dt}{\int_0^T tY_t dt}\right) + \widetilde{\sigma} W_T \\ &= \frac{\widetilde{\sigma}}{\sigma}\frac{\delta\left(\int_0^T X_t W_t dt\right)}{\int_0^T tX_t dt} + \frac{\widetilde{\sigma}}{\sigma}\int_0^T X_t W_t dt \int_{s=0}^T \frac{\int_{t=0}^T tD_s X_t dt}{\left(\int_0^T tX_t dt\right)^2} ds + \widetilde{\sigma} W_T \end{aligned}$$

Then using the fact that

$$\int_{s=0}^{T} \frac{\int_{t=0}^{T} tD_sX_tdt}{\left(\int_0^T tX_tdt\right)^2}ds = \frac{\sigma}{x}\frac{\int_{t=0}^{T} t^2X_tdt}{\left(\int_0^T tX_tdt\right)^2}$$

and

$$\delta\left(\int_{t=0}^{T} X_tW_tdt\right) = \int_{s=0}^{T}\left(\int_{t=0}^{T} X_tW_tdt\right)dW_s$$

leads to the final result. The traditional vega is the extended vega divided by the perturbed volatility $\widetilde{\sigma}$.□

Gamma

The Gamma can be calculated as the vega divided by $x^2\sigma T$ for the Black Scholes model and by $x^2\int_0^T \sigma_s dt$ in the Black model (see the remarks of chapter 2 page53 section 2.3.2 and Benhamou(2000c)).

4.4.2 Complex Asian

Specifying the terms in the solution of (4.11), the weight function for the delta of the complex Asian option is in this simple case given by:

$$w_s = \frac{1}{\sigma_s x}\left(A + B\frac{X_s}{\int_0^T X_tdt}\right) \tag{4.16}$$

with the coefficient A and B verifying:

$$A = \frac{1}{\int_0^T \frac{2tX_t}{\int_0^T X_tdt}dt - T} \qquad B = 1 - \frac{T}{\int_0^T \frac{2tX_t}{\int_0^T X_tdt}dt - T}$$

4.5 Numerical method

In this section, we examine the particular case of an Asian option with the following characteristics: a one year continuous time Asian option with risk-free interest rate $r = 3\%$, a null dividend rate, an initial underlying S_0 of 100, two strikes K_1, K_2 of 100 and 110 respectively and a volatility σ of 20%. We consider two different options:

- a corridor Asian option whose payoff is $1_{\left\{K_1<\int_0^T S_t dt<K_2\right\}}$
- an up and out Asian call $1_{\left\{\int_0^T S_t dt<K_2\right\}}(S_T - K_1)^+$

4.5.1 Simulation of the Gamma

We examine the Malliavin simulation on the gamma, since it is precisely in this particular case that the Malliavin weighted scheme should dramatically outperform finite difference. These results confirm the efficiency of the Malliavin weighted scheme compared with the standard finite difference. The figure 4.1 compares the two methods: Malliavin weighted scheme (black line) and the finite difference method (grey line). The Malliavin weighted scheme converges to the right answer quite fast with almost no oscillations, whereas the finite difference estimator fluctuates with a pseudo-periodicity around the correct value.

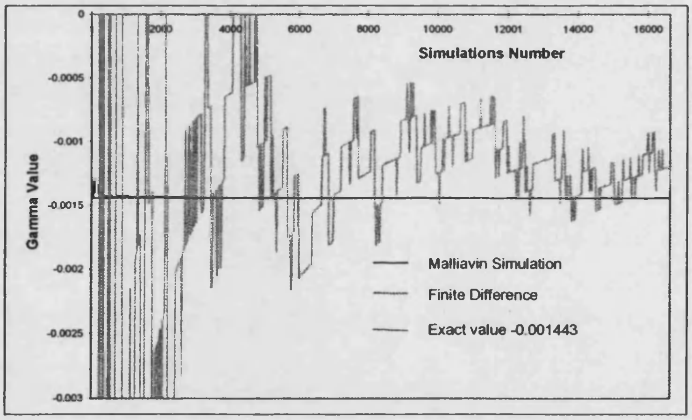

Figure 4.1: Efficiency of the Malliavin weighted scheme for the computation of the gamma of the first option

This advantage of the Malliavin weighted scheme is even more striking in the case of the second option as shown by figure 4.2. The Malliavin weighted scheme is given by the black line and the finite difference method by the grey one.

4.5.2 Comparison of variance

The purpose of the Malliavin weighted scheme is to reduce the variance of the simulations. We compare the variance implied by the Malliavin weighted scheme and the one of the finite difference in the case of the same two options for the delta, gamma, vega and rho. We found that Malliavin weighted scheme provides a quite efficient way of reducing the variance as shown in table 4.2. The number of simulation draws was 20,000 as indicated by N=20,000. Table 4.2 gives the ratio of simulation variances between finite difference and Malliavin-based simulation. When this ratio is bigger than 1, it means that the Malliavin method reduces the variance. Indeed, the numbers found are big numbers. Since the variance decreases roughly linearly in n, a ratio of 10 means that we need 10n draws with the finite difference method to get the same variance as the one obtained by the Malliavin method with only n draws. We found evidence that the Malliavin weighted scheme is more efficient for the first than the second option. The first option payoff is indeed more discontinuous. The ratios found are big numbers. For the gamma, it means that the same accuracy as 1,000 draws with the Malliavin weighted scheme is obtained with 14,000,000 draws for the first option and 7,000,000 for the second option when using the traditional finite difference method.

4.6 Conclusion

In this chapter, we have seen that using Malliavin calculus and its integration by parts formula, we can extend Malliavin calculus based formulae, for the Greeks, to the continuous-time Asian option. This enables us to smoothen the function to be estimated by the Monte Carlo or Quasi Monte Carlo simulation. This extension to Asian options is of particular interest since there is no closed formula for Asian options, even in the simple case of a geometric Brownian motion.

Option type	Variance ratio $\widehat{\sigma}^2_{\text{Finite Difference}}/\widehat{\sigma}^2_{\text{Malliavin}}$	delta	gamma	rho	vega
First option $1_{\{K_1<\int_0^T S_t dt<K_3\}}$	N=20,000	160	14256	69	7422
Second option $1_{\{\int_0^T S_t dt<K_3\}}(S_T - K_1)^+$	N=20,000	146	7210	47	5900

Table 4.2: Variance ratio between the Malliavin weighted scheme and the finite difference method

There are many possible extensions to this work. A first one is to study a local version of Malliavin weighted scheme, which smoothens the discontinuity at the kink and is based on a finite difference scheme enhanced by the common random numbers method everywhere else. A second one is to extend this methodology to continuous lookback options. However, the case of the continuous supremum of the underlying is not as easy to handle as the continuous time arithmetic average.

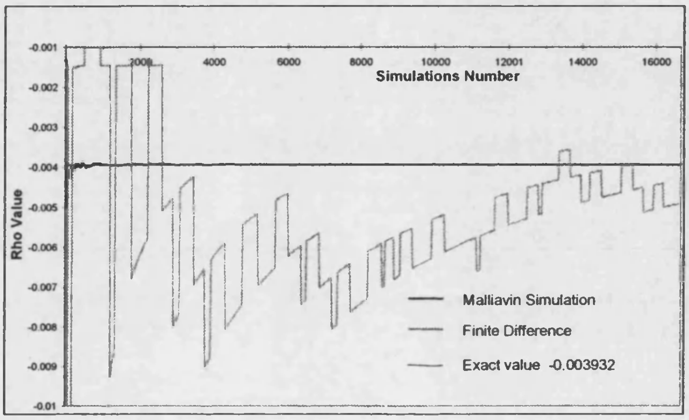

Figure 4.2: Efficiency of the Malliavin weighted for the computation of the gamma of the second option

Chapter 5

A Martingale Result for Convexity Adjustment in the Black Pricing Model

Summary of the chapter

This chapter explains how to calculate the convexity adjustment for interest rate derivatives, using martingale theory, when assuming a deterministic time dependent (often referred to as the Black (1976) pricing model) volatility. In this present chapter, we proceed as follows: first, we set up a proper no-arbitrage framework illustrated by a relationship between yield rate drift and bond price. Second, by making use of an approximation, we derive a closed formula with a specification of the error term. Earlier works (Brotherton-Ratcliffe and Iben (1993) or Hull (1997)) assumed constant volatility and could not specify the approximation error. As an application, we examine the convexity bias between CMS and forward swap rates.

5.1 Introduction

The motivation of this chapter is to provide a proper framework for the convexity adjustment formula, using martingale theory and no-arbitrage relationship. The use of the martingale theory initiated by Harrison, Kreps (1979) and Harrison, Pliska (1981) enables us to define an exact but non explicit formula for the convexity. We show that by making an approximation, we can rederive previous results, first discovered by Brotherton-Ratcliffe and Iben (1993) and later by Hull (1997) and Hart (1997). However, the approach hereby adopted has the great advantage to enable us to specify the error of the approximation. We extend results derived in the Black Scholes framework to time-dependent volatility, often referred to as the pricing model of Black (1976). This is more in agreement with the consideration of practitioners who commonly use time dependent volatility to best fit the market prices.

The convexity adjustment hereby derived is of considerable interest to measure the convexity adjustment required for a security paying only once a swap rate. The rate of this kind of security is named in the fixed income market as the CMS rate.

The formula, first discovered by Brotherton-Ratcliffe and Iben (1993) and later by Hull (1997), is an analytic approximation of the difference between the expected yield and the forward yield, collectively referred to as the convexity adjustment. It assumes a constant yield volatility σ. Brotherton-Ratcliffe and Iben (1993) show that the convexity adjustment for a yield bond was given by:

$$-\frac{1}{2}\left(y_0^f\right)^2\sigma^2 T\frac{h''\left(y_0^f\right)}{h'\left(y_0^f\right)} \tag{5.1}$$

where y_0^f denotes the value today of the forward bond yield, $h(y)$ the price of the bond that provides coupons equal to the forward bond yield and that is assumed to be a function of its yield y, $h'(y)$ and $h''(y)$ the first and second partial derivatives of the bond price $h(y)$ with respect to its yield. Hull (1997) shows that this convexity adjustment can be extended to derivatives with payoff depending on swap rates. Hart (1997) sharpened the approximation with a Taylor expansion up to the fourth term. However, all proofs, based on Taylor expansion, never referred to the error of the approximation. Deriving this error is precisely the motivation for this chapter. We show that formulae similar to (5.1) can be derived with an exact definition of the error term, when a proper no-arbitrage framework is assumed.

The remainder of this chapter is organized as follows. In section 2, we give some insight about convexity. In section 3, we derive convexity adjustment from a no-arbitrage proposition implied by martingale condition. We show how to derive a approached formula, with a control on the error term. Monte Carlo simulations confirm the efficiency of the approached closed formula. In section 4, we give a formula for the convexity adjustment required for a CMS rate. We conclude briefly and refer to future developments.

5.2 Insights about convexity

For fixed income markets, convexity has emerged as an intriguing and challenging notion. Taking correctly this effect into account could provide competitive advantage to financial institutions.

5.2.1 The definition of the convexity

One main difficulty is to give a unified framework for all the different meanings of convexity. Indeed, it is true that the notion of convexity refers to different situations, which can be sometimes seen as having almost nothing in common. It is sometimes used as the gamma ratio for interest rate options, as an indicator of risk for bonds portfolios, as a measurement of the curvature of some financial instruments or as a small adjustment quantity for a wide variety of interest rate derivatives. Indeed, convexity has become a synonym for small adjustment in fixed income markets, related somehow to the notion of mathematical convexity and more generally to a second order differentiation term. The situations which are of particular interest for practitioners can be classified into two types with different causes of adjustment:

- the bias due to correlation between the interest rate underlying the financial contract and the financing rate. An example is the bias between forward and futures contracts. This correlation, capitalized by the margin calls of the futures contract, leads to a more expensive (respectively cheaper) futures contract in the case of positive (respectively negative) correlation.
- the modified schedule adjustment. Even if the analysis is the same for the two sub-cases above, it is traditionally divided into two categories depending on the type of rates:
  - One-period interest rate (money-market rates, zero-coupon rate) and bond yield. An example is the difference between plain vanilla products and in-arrear ones, or in-advance ones. Another one is the differentiation between forward yield rate and expected bond yield. Furthermore, a modified formula for every type of path dependent interest rate option, like Asian options, multi-European options is required.
  - Swap rates. These products are called by the market CMS products (CMS: constant maturity swap). A convexity adjustment is required between forward swap rate and expected swap rate. This rate is traditionally called the CMS rate. Indeed, this analysis is very similar to the previous case. It comes as well from a modified schedule.

For practitioners, the two sub-cases have long been separated because they were concerning different products. As a result, they were seen as two types of adjustment. Indeed, the two required convexity adjustments are coming from a modified schedule of the rate.

In this chapter, we concentrate on the distinction between forward and expected bond yield as well as swap rate.

5.2.2 A rough model

As pointed out in our definition section, one should make a distinction between the convexity adjustment required between futures and forward contract (correlation convexity) and the modified schedule adjustment. As a general rule for the second type of situation, it is necessary to make a convexity adjustment when an interest rate derivative is structured so that it does not incorporate the natural time lag implied by the interest rate. This is the case obviously of in-arrears and in-advance products where the rate is observed and paid at the same time. This is as well the case of the CMS rate where the swap rate instead of being paid during the whole life of the swap is only paid once.

Let us now explain intuitively the convexity defined as the difference between forward rate and expected rate. We examine the case of bond but this applies to the swap rate as well. Since the relationship between bond price and the bond yield Y is non-linear, it is not correct to say that the expected yield is equal to the yield of a forward bond (also called the forward yield). Similarly, it is not correct to say that the expected swap rate should be equal to the forward swap rate.

This can be well understood by taking a two state world. The bond price can be either P_1, P_2 with equal probability $\frac{1}{2}$. The corresponding yields are Y_1, Y_2. In this binomial world, the expected price P_e is given by sum of the different possible prices multiplied by their corresponding probability: $P_e = \frac{1}{2}P_1 + \frac{1}{2}P_2$. The forward yield Y^f is the yield corresponding to the expected price P_e. The expected yield Y_e is the one given by the expected value of the yield $Y_e = \frac{1}{2}Y_1 + \frac{1}{2}Y_2$.

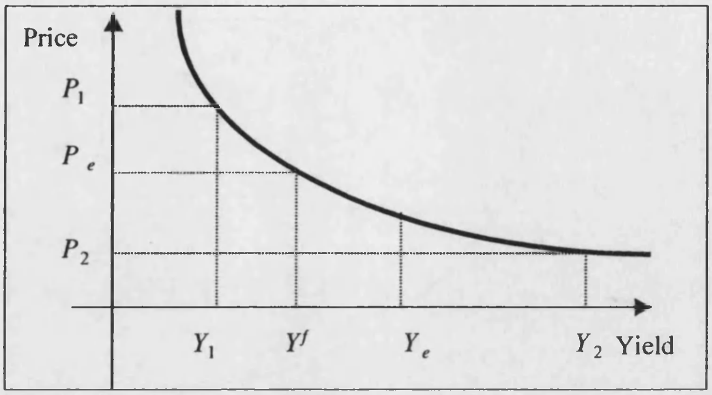

Figure 5.1: Convexity of the bond price with respect to its yield: Case of a simple two state model. This graphic shows that the expected yield denoted by Y_e is higher than the corresponding forward yield Y^f

However, since the relationship between price and yield is decreasing and convex, the two given yields, forward and expected one, are not equal and the expected yield Y_e is above the forward one as figure 5.1 shows it.

These results can be derived in a more general stochastic framework. From the Jensen inequality on convex functions, applied to the function $P(y)$, one sees that the forward price defined as the expected value of the price, under the risk neutral probability $\mathbb{E}(P(Y))$ should be higher than the bond price with a yield equal to the expected rate $P(\mathbb{E}(Y))$:

$$\mathbb{E}(P(Y)) > P(\mathbb{E}(Y))$$

Using the fact that the bond price is a decreasing function, we get that the expected bond rate defined as the expected value of the yield $\mathbb{E}(Y)$ is higher than the forward bond rate corresponding to the forward price $\mathbb{E}(P(Y))$ ($Y^f = P^{-1}(\mathbb{E}(P(Y)))$). The difference between the expected yield and the forward yield $Y^e - Y^f$ is called the convexity adjustment and defined by

$$Y^e - Y^f = \mathbb{E}(Y) - P^{-1}(\mathbb{E}(P(Y))) \tag{5.2}$$

With these rough modelling framework, we can already get interesting results.

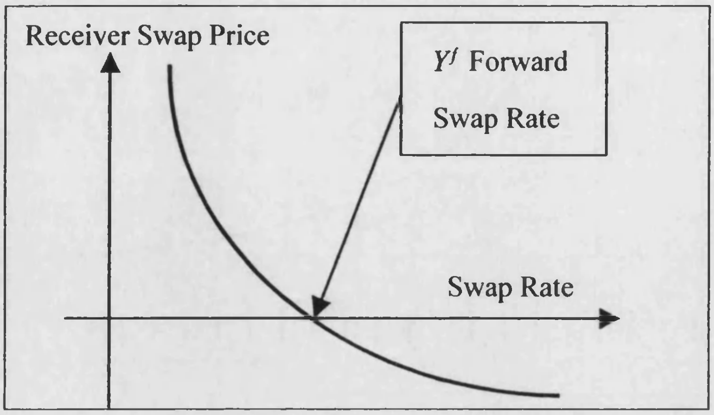

Figure 5.2: Convexity of the swap price with respect to its swap rate: The relationship between the receiver swap price and the swap rate is convex and decreasing. The only difference between swap and bond contract lies in the possible negative values of the receiver swap

When a bond or a security price is a convex function of the interest rate, the expected bond yield $\mathbb{E}(Y)$ is always above the forward bond yield $P^{-1}(\mathbb{E}(P(Y)))$.

This can as well applied to swap rates. Indeed, a receiver swap, swap where one receives the fixed rate and pays the floating one, is also a convex decreasing function of the swap rate. The only difference comes from the fact that the swap price contrary to the bond price can be negative. This is illustrated by figure 5.2. Since only hypotheses on the monotonicity and convexity of the function are required for deriving our result above (5.2), we conclude that the expected swap rate is above the forward swap rate.

As a general conclusion of this subsection, expected bond yields or swap rates should be higher than the corresponding forward for convex contracts and lower for concave ones.

5.2.3 Static hedge: locking the convexity

We saw above that the difference between the forward yield and the expected yield is due to the fact that the underlying bond price is a decreasing convex function of the yield. We can take advantage of this by a static hedge. Let us consider a continuous time trading economy. The uncertainty in this economy is characterized by the probability space (Ω, F, Q) where Ω is the state space, F is the $\sigma-$algebra representing measurable events, and Q is the risk neutral probability measure. We denote by y_t^f the value of the forward yield at time t. We denote by $h\left(y_t^f\right)$ the payoff of a security depending on the forward yield. We denote by y_0^f the value today of the forward yield. We denote by σ the constant volatility of the forward yield at time T when compared with the today forward yield. This means that the square difference between the forward yield at time T and the today value is proportional to the volatility times the time elapsed times the square of the today value of the forward yield:

$$\frac{\mathbb{E}_{Q_T}\left(\left(y_T^f - y_0^f\right)^2\right)}{\left(y_0^f\right)^2} = \sigma^2 T \tag{5.3}$$

The above analysis is carried out for yield bonds for clarity reasons. However, this can be easily adapted to swap rates. We consider the following portfolio:

- a forward contract on the forward yield with a strike at the today value of the forward yield. The payoff at time T is simply the difference between the forward yield at time T: y_T^f and the strike: today value of the forward yield y_0^f.

- a hedging portfolio composed of n forward contract(s) on the bond set at at-the-money strike. The payoff of the forward contract is therefore the difference between the non-linear security payoff at time T, $h\left(y_T^f\right)$ and the price if the yield were the value of the today forward yield, $h\left(y_0^f\right)$. This is a hedging portfolio since the variation of the forward contract on the forward yield y_T^f is offset by the variation of the forward contract on the bond. Since the forward contract is set at at-the-money strike, this contract is of zero value.

Since the value of the total portfolio is equal to the sum of its two components, with the second one of zero value, the total value of the portfolio is equal to the value of the first component of the portfolio, given at maturity time T by the expected difference between the forward yield and the today value of the forward yield, which is exactly the definition of the convexity adjustment. The value today of the total portfolio is therefore the convexity adjustment times the zero coupon maturing at time T. The determination of the convexity adjustment is consequently equivalent the one of the global portfolio. Its expression is given by the following proposition:

Proposition 15 *Convexity adjustment*
The value at time $t = 0$ of the portfolio denoted by P is given by

$$\frac{P}{B(0,T)} = -\frac{1}{2}\frac{h''\left(y_0^f\right)}{h'\left(y_0^f\right)}\left(y_0^f\right)^2\sigma^2 T \tag{5.4}$$

Proof: By means of a change of probability measure, from risk neutral to forward neutral probability measure, the price P of all the portfolio can be written as the expected value of the payoff under the forward neutral probability measure Q_T times the zero-coupon bond maturing at the payment time T:

$$P = B(0,T)\,\mathbb{E}_{Q_T}\left(\left(y_T^f - y_0^f\right) + n * \left(h\left(y_T^f\right) - h\left(y_0^f\right)\right)\right)$$

Using a Taylor expansion up to the second order around the today forward yield, we get that the payoff of the hedging portfolio at time T can be expressed as a simple function of the difference between the forward yield at time T : y_T^f and the today value of the forward yield y_0^f.

$$h\left(y_T^f\right) - h\left(y_0^f\right) = \left(y_T^f - y_0^f\right)h'\left(y_0^f\right) + \frac{1}{2}\left(y_T^f - y_0^f\right)^2 h''\left(y_0^f\right) + o\left(\left(y_T^f - y_0^f\right)^2\right)$$

we can assume that the difference between the value at time T of the forward yield y_T^f and its today value y_0^f is small since the forward yield at time T should be close to its initial value. The total value of the portfolio can therefore be expressed as a quadratic function of the difference between the value at time T of the forward yield y_T^f and its today value y_0^f

$$P = B(0,T)\,\mathbb{E}_{Q_T}\left[\left(y_T^f - y_0^f\right)\left(1 + nh'\left(y_0^f\right)\right) + \frac{1}{2}n\left(y_T^f - y_0^f\right)^2 h''\left(y_0^f\right)\right]$$

To eliminate the first order risk (role of our hedging strategy), the quantity of the hedging portfolio should exactly offset the variation of the forward contract(up to first order):

$$n = -\frac{1}{h'\left(y_0^f\right)}$$

The quantity n is positive and confirms that the second component of the global portfolio is a hedge against the variation of the first one. The value of the global portfolio is therefore coming only from the second order risk or gamma risk. Getting all the deterministic terms out of the expectation sign leads to the following expression:

$$P = -\frac{1}{2}B(0,T)\frac{h''\left(y_0^f\right)}{h'\left(y_0^f\right)}\mathbb{E}_{Q_T}\left[\left(y_T^f - y_0^f\right)^2\right]$$

Using the strong assumption (5.3) about the pseudo "volatility" σ, we get that the price of the total portfolio can be expressed as a function of the today value of the forward yield y_0^f and the parameter of "volatility" σ

$$P = -\frac{1}{2}B(0,T)\frac{h''\left(y_0^f\right)}{h'\left(y_0^f\right)}\left(y_0^f\right)^2\sigma T$$

which is exactly the result (5.4).□

5.3 Calculating the convexity adjustment

In this section, we show how to derive the convexity adjustment required when assuming a time-dependent volatility, an assumption similar to the Black model. The difference between our model and the Black model lies in the fact that in our model, the drift term is supposed to be stochastic. However, when we take a deterministic approximation of our drift term, our model becomes a standard Black model.

Our proof is based on the martingale theory. We obtain that the martingale condition implies a strong condition on the drift term of the forward yield. Making approximations, we obtain as a particular case (when the volatility is constant) the well-known formula for the convexity adjustment (5.1), obtained by Brotherton-Ratcliffe and Iben (1993) and later by Hull (1997). However, the motivation of this approach is to specify the error of the approximation. Monte Carlo simulations prove that the error is relatively small.

5.3.1 Pricing framework

We consider a continuous trading economy with a limited trading horizon $[0, \tau]$ for a fixed $\tau > 0$. The uncertainty in the economy is characterized by the probability space (Ω, F, Q) where Ω is the state space, F is the σ−algebra representing measurable events, and Q is the risk neutral probability measure uniquely defined in complete markets (Harrison, Kreps (1979) and Harrison, Pliska (1981)). We assume that information evolves according to the augmented right continuous complete filtration $\{F_t, t \in [0, \tau]\}$ generated by a standard one-dimensional Wiener Process $(W_t)_{t \in [0,\tau]}$.

We assume as well that the price at time t of the bond can be defined as a function $h(.)$ of the bond yield at time t, $\left(y_t^f\right)_{t<\tau} : h\left(y_t^f\right)$. The two stochastic variables $\left(y_t^f\right)_{t<\tau}$, $\left(h\left(y_t^f\right)\right)_{t<\tau}$ are supposed to be adapted to the information structure $(F_t)_{t \in [0,\tau]}$. We examine a bond security whose payoff is paid at time T. Following the work of El Karoui et al. (1995), the no-arbitrage condition and the markets' completeness assumption enable us to define a unique forward neutral probability measure Q_T, under which the price $h\left(y_t^f\right)$ is a martingale. Under this probability measure Q_T, the volatility of the forward yield rate y_t^f is supposed to have a deterministic volatility function depending only on the time, leading to the following diffusion:

$$\frac{dy_t^f}{y_t^f} = \mu_t dt + \sigma_t dW_t$$

where the drift term is stochastic. Since the volatility is supposed to be a deterministic function of time, this is sometimes referred to as the "Black" pricing model. However, the drift is stochastic. It is therefore different from the standard Black model where the drift is deterministic. We denote the zero coupon bond price at time t, maturing at time $T > t$ by $B(t,T)$. The following theorem gives us the necessary condition on the drift term for the price $h\left(y_t^f\right)$ to be a martingale.

5.3.2 Convexity adjustment formula

Theorem 3 *Convexity Adjustment formula*

Under the hypotheses mentioned above, the drift term should satisfy the following nonarbitrage condition

$$\mu_t = -\frac{\frac{1}{2}h''\left(y_t^f\right)\sigma_t^2 y_t^f}{h'\left(y_t^f\right)} \tag{5.5}$$

Proof: Ito lemma gives

$$dh\left(y_t^f\right) = h'\left(y_t^f\right) y_t^f \sigma_t dW_t + \left[y_t^f h'\left(y_t^f\right)\mu_t + \frac{1}{2}h''\left(y_t^f\right)\left(\sigma_t y_t^f\right)^2\right] dt$$

Under the forward neutral probability Q_T, the price of the bond $h\left(y_t^f\right)$ is a martingale. This implies that the drift term $y_t^f h'\left(y_t^f\right)\mu_t + \frac{1}{2}h''\left(y_t^f\right)\left(\sigma_t y_t^f\right)^2$ should be equal to zero, which leads to the necessary condition (5.5).□

We take the following definition of the convexity adjustment:

Definition 2 *The convexity adjustment is defined as the difference between the expected yield under the forward neutral probability measure and the forward yield, leading to the exact but non-explicit formula:*

$$\mathbb{E}_{Q_T}\left(y_T^f / F_0\right) - y_0^f \tag{5.6}$$

The above definition provides us with an exact but not tractable formula of the convexity adjustment. Assuming that the drift term can be approximated by its value with the forward yield equal to the today forward yield y_0^f, we get a closed and tractable formula given the following theorem:

Theorem 4 *Under the assumptions above, we can prove that the convexity adjustment for the expected bond yield can be approximated by*

$$y_0^f \left(e^{-\frac{1}{2}\frac{h''\left(y_0^f\right)}{h'\left(y_0^f\right)} y_0^f \int_0^T \sigma_t^2 dt} - 1 \right) \tag{5.7}$$

where $h'\left(y_0^f\right)$ and $h''\left(y_0^f\right)$ denote the first and second derivatives of the bond price with respect to its yield y taken at the point y_0^f.

Proof: Calculating the expected yield under the forward neutral probability gives:

$$\mathbb{E}_{Q_T}\left(y_T^f / F_0\right) = \mathbb{E}_{Q_T}\left(y_0^f e^{\int_0^T \left(\mu_t - \frac{1}{2}\sigma_t^2\right)dt + \int_0^T \sigma_t dW_t}\right)$$

And using that $\mu_t - \frac{1}{2}\sigma_t^2 \simeq -\frac{1}{2}\frac{h''\left(y_0^f\right)}{h'\left(y_0^f\right)} y_0^f \sigma_t^2 - \frac{1}{2}\sigma_t^2$ leads to

$$\mathbb{E}_{Q_T}\left(y_T^f / F_0\right) \simeq y_0^f \exp\left(-\frac{1}{2}\frac{h''\left(y_0^f\right)}{h'\left(y_0^f\right)} y_0^f \int_0^T \sigma_t^2 dt\right)$$

Consequently, using its definition (2), the convexity adjustment is given by the final result (5.7).□

An approximation of the theorem formula is then given by a Taylor expansion of the exponential up to the first order, leading to an extension, to time-dependent volatility, of the formula of Iben (1993)

$$-\frac{1}{2}\frac{h''\left(y_0^f\right)}{h'\left(y_0^f\right)} \left(y_0^f\right)^2 \int_0^T \sigma_t^2 \tag{5.8}$$

Corollary 6 *Black Scholes formula*

When the volatility is constant, the convexity adjustment derived here leads exactly to the one obtained by Brotherton-Ratcliffe and Iben (1993) and later by Hull (1997)

Proof: Using the approximation formula (5.8) with a constant volatility leads to the result.□

The calculation in the proof relies in the fact that the drift term μ_t can be approximated by the initial deterministic value equal to $\frac{-\frac{1}{2}h''\left(y_0^f\right)y_0^f\sigma_t^2}{h'\left(y_0^f\right)}$. This implies two interesting remarks. First, it means that the Black Scholes convexity adjustment used by markets is a very rough approximation formula when assuming a deterministic volatility. One assumes that the stochastic drift term is indeed deterministic. Second, this approximation is highly depending on the initial value of the forward yield rate y_0^f. If this forward yield rate is unstable, it might be more appropriate to use the historical average of past observations.

We can now specify the error term as the difference between our closed formula (5.7) and the formula (5.6). We can see that in the difference, the two terms, y_0^f simplify each other, leading to an error term given by:

$$\mathbb{E}_{Q_T}\left[y_0^f\left(e^{\int_0^T\left(\mu_t-\frac{1}{2}\sigma_t^2\right)dt+\int_0^T\sigma_t dW_t}-e^{-\frac{1}{2}\frac{h''\left(y_0^f\right)}{h'\left(y_0^f\right)}y_0^f\int_0^T\sigma_t^2 dt}\right)\right]$$

Using a change of probability measure (Girsanov theorem), we can see that this expression is, under a probability measure denoted by $\widetilde{Q}$, the difference between two terms, where the Radon Nykodim derivative of $\widetilde{Q}$ with respect to Q_T is given by $e^{\int_0^T\sigma_t dW_t-\frac{1}{2}\int_0^T\sigma_t^2 dt}$. Using the Taylor-Lagrange theorem, we get that there exists a parameter θ_t between 0 and t so that this difference of terms can be expressed as the difference between the two rates y_t and y_0^f times the derivatives of the exponential:

$$\mathbb{E}_{\widetilde{Q}}\left[y_0^f\left(e^{-\int_0^T\frac{\frac{1}{2}h''\left(y_{\theta_t}^f\right)y_{\theta_t}^f\sigma_t^2}{h'\left(y_{\theta_t}^f\right)}dt}\int_0^T g\left(y_{\theta_t}^f\right)dt\left(y_t^f-y_0^f\right)\right)\right]$$

where the function $g\left(\right)$ denotes the derivatives of the function $\frac{\frac{1}{2}h''\left(y\right)y}{h'\left(y\right)}\sigma_t^2$ with respect to y. To go further, we need to be more specific on the function h. This implies of course to specify more the diffusion equation of y. Without any further information, nothing very specific can be derived on the error term. Another way to measure the error term is by means of Monte Carlo simulations.

5.3.3 Monte Carlo simulations of the error

In the previous section, we have assumed that the forward yield rate y_t^f can be approximated by the today value of the forward yield y_0^f. In this subsection, we analyze, by means of Monte-Carlo simulations how big this error is in this approximation. We consider a derivative that provides the payoff equal to the one-year zero coupon rate in T years multiplied by a principal of 100. For the simplicity of the simulation, we take a constant volatility ($\sigma_t = \sigma$) equal to 20 % and a forward rate of 10%. Since our bond is a one year zero coupon, its payoff is equal to the discounted value of its unique coupon.

$$h(y) = \frac{1}{1+y}$$

The no-arbitrage condition (5.5) implies that the yield should have the following diffusion

$$\frac{dy_t^f}{y_t^f} = \frac{\sigma^2 y_t^f}{1+y_t^f}dt + \sigma dW_t$$

with the initial value $y_t^f = y_0^f$. The aim of the Monte Carlo simulation is to examine the quality of the approximation done for the convexity adjustment. We compute the expected yield $\mathbb{E}^Q[y_T]$ which is called theoretical yield in table 5.1 (calculated with a Sobol sequence Quasi-Monte Carlo with 20,000 draws) and compare it to the approximated formula for the convexity adjustment

$$y_0^f\, e^{-\frac{1}{2}\frac{h''\left(y_0^f\right)}{h'\left(y_0^f\right)}y_0^f\sigma^2 T} = y_0^f\, e^{\frac{y_0^f}{1+y_0^f}\sigma^2 T}$$

The results are given in table 5.1. These are simulations for different value of the expiry time T : 3, 5 and 10 years. It means that our derivative asset is paying the one-year zero-coupon rate determined at time T and paid at time T. The price of this derivative is therefore the forward rate with a convexity adjustment times the principal 100 discounted by the zero coupon bond maturing at time T.

Time T	Approached yield	Theoretical yield (MC)	Approached Price	Theoretical Price (MC)
3	10.1097	10.1099	7.59556	7.59572
5	10.1835	10.1844	6.32314	6.32373
10	10.3703	10.3888	4.83783	4.84646

Table 5.1: Results of the simulation for the expected rate. The simulations show that the rough approximation is quite valid.

The results show that the approximation is quite efficient and can therefore been used as a good estimator of the convexity adjustment required for the derivatives concerned.

5.4 CMS rate

5.4.1 Introduction

The CMS rate is the rate of a contract that pays only once the swap rate. Because a regular swap rate should be paid during the whole period, this product includes a modified schedule. The swap price is a convex function of the swap rate. Therefore, as explained in the first section, the expected swap rate should not be equal to the forward swap rate. The difference should be positive because of the convexity of the function.

This result can be proved in a very basic way. We want to calculate the expected value of an annual swap rate assumed to have n payments at date $T+i$ with $i = 1...n$. Let us denote by y_0^f the forward swap rate, and by y_t^f the swap rate at time t. A useful relationship between a receiver swap price with a fixed rate equal to the forward swap rate and the swap rate is the following: the receiver swap price, P_{Swap}, is equal to the difference between the forward swap rate and the swap rate times the swap sensitivity.

$$P_{Swap}(t) = \sum_{i=0}^{n} B(t, T+i)\left(y_0^f - y_t^f\right) \tag{5.9}$$

Therefore, introducing this quantity, we get that the expected swap rate can be calculated as:

$$\mathbb{E}_{Q_T}\left[y_T^f\right] = \mathbb{E}_{Q_T}\left[-\frac{\sum_{i=0}^{n} B(T, T+i)\left(y_0^f - y_T^f\right)}{\sum_{i=0}^{n} B(T, T+i)} + y_0^f\right]$$

Knowing that the swap sensitivity $\sum_{i=0}^{n} B(t, T+i)$ is positively correlated with the receiver swap price $\sum_{i=0}^{n} B(t, T+i)\left(y_0^f - y_t^f\right)$ for every time t, we get that the two variables, the opposite of the inverse of the sensitivity of the swap $-\frac{1}{\sum_{i=0}^{n} B(T,T+i)}$ and the receiver swap $\sum_{i=0}^{n} B(T, T+i)\left(y_0^f - y_T^f\right)$ are positively correlated. A simple result is that when two stochastic variables X_1 and X_2 are positively correlated, the expectation of their product is bigger than the product of their expectation

$$\mathbb{E}[X_1 X_2] \geq \mathbb{E}[X_1]\,\mathbb{E}[X_2]$$

In the case of a strictly positive correlation, the inequality is strict. Since the forward swap is exactly at the money (fixed rate equal to the forward rate), its expected value should be equal to zero. This leads to the final result that the expected swap rate should be higher than the corresponding forward swap:

$$\mathbb{E}_{Q_T}[y_T] > y_0^f$$

5.4.2 Hedging strategy

The hedging point of view is interesting as well. If an investor who is long a CMS rate hedges it like a forward swap rate, he will make almost surely profit. Let us show how to make an arbitrage in this situation. The hedging strategy should cost today exactly the discounted swap rate $y_0^f B(0, T)$.

Take the following strategy. An investor is long a CMS rate which maturity is denoted by T, with an underlying swap rate of an n years maturity. He hedges it as if the CMS contract were giving him the forward swap rate. A hedging strategy is to replicate synthetically the forward swap rate:

- be long the corresponding forward receiver swap with an amount equal to the inverse of the forward swap sensitivity $\frac{1}{\sum_{i=1}^{n}\mathbb{E}_{Q_T}[B(T,T+i)]}$ with a fixed rate equal to the forward swap rate. Since the receiver swap has its fixed rate equal to the forward rate, the value today of this swap is 0.

- be short at the same time a risk-free zero coupon bond maturing at time T with an investment amount equal to the forward swap rate y_0^f. The value of this zero coupon bond is today $y_0^f B(0,T)$.

We verify that the today's hedge cost is the discounted forward swap rate $y_0^f B(0,T)$. Let us now examine our total portfolio. It is long a CMS rate, long a forward receiver swap, short a zero coupon. The total value Π_T of the portfolio at time T is:

$$\Pi_T = \left[\left(y_T^f - y_0^f\right) + \frac{P_{Swap}(T)}{\sum_{i=1}^{n}\mathbb{E}_{Q_T}[B(T,T+i)]}\right]$$

using again the useful relationship between swap price and swap rate (5.9), we get

$$\Pi_T = \left[\frac{-P_{Swap}(T)}{\sum_{i=1}^{n} B(T,T+i)}\right] + \left[\frac{P_{Swap}(T)}{\sum_{i=1}^{n}\mathbb{E}_{Q_T}[B(T,T+i)]}\right]$$

Using the fact that being short the receiver swap is equivalent to being the corresponding payer swap, the first position is exactly long a stochastic amount $\frac{1}{\sum_{i=1}^{n} B(T,T+i)}$ of a payer swap. Denoting by $P_{P\_Swap}(T)$ the price of the payer swap, we get that our total portfolio can be decomposed into two sub-portfolios:

- portfolio 1: the sum of the CMS rate and the zero coupon bond times the forward swap rate. Its value at time T is equal to a payer swap $P_{P\_Swap}(T)$ with a stochastic amount $\frac{1}{\sum_{i=1}^{n} B(T,T+i)}$

- portfolio 2: the forward receiver swap with an amount equal to the inverse of the forward swap sensitivity $\frac{1}{\sum_{i=1}^{n}\mathbb{E}_{Q_T}[B(T,T+i)]}$.

Let us examine different scenarios for the interest rates.

- If the swap rate realized at time T is exactly the forward swap rate, the two portfolios have zero value.

- If the swap rate y_T^f is above the forward swap rate $-y_0^f$, the portfolio 1 increases because of two things. First, this is because the payer swap ends in the money. Second, it increases as well because the sensitivity of the swap has decreased. This is in turn equivalent to an increase of the inverse of the swap sensitivity. In contrast, the portfolio 2 decreases only because the receiver swap ends out of the money. This offsets only the profit realized on the payer swap. Therefore, in this case, the total portfolio will increase.

- If the swap rate is below the forward swap rate, the payer swap ends out of the money whereas the receiver swap ends in the money by the same amount. However, the loss on the payer swap of the portfolio 1 is offset by the decrease of the inverse of the swap sensitivity, leading again, to a positive value for the total portfolio.

As a conclusion, we can see that whenever the swap rate is above or below the forward swap rate, our total portfolio ends in the money. This positive value is due to the convexity effect. We see on this example that the static hedge does not hedge against the convexity term. Since this effect is depending obviously on the importance of the move between the swap rate and the forward one, in either directions, this should be related somehow to the volatility. A hedging strategy against the convexity term should therefore have a volatility component. This can be done with options like swaptions. However, since swaptions are not perfect substitute for the convexity term, the hedge needs to be re-evaluated dynamically. Many questions remain unanswered. Which option should one take and more specifically which option maturity and strike should one choose? These questions are depending mainly on the market type.

5.4.3 Pricing CMS rate

The price of a bond that gives the forward swap rate at each coupon date, and with no principal exchanged at the end of the swap is given by

$$h(y) = \frac{F}{(1+y)^{T_1}} + ... + \frac{F}{(1+y)^{T_n}}$$

This leads to the following calculation for the convexity adjustment denoted by CA (equation (5.8))

$$CA = \frac{1}{2} \frac{(T_1+1)\,T_1 \frac{F}{\left(1+y_0^f\right)^{T_1+2}} + ... + (T_n+1)\,T_n \frac{F}{\left(1+y_0^f\right)^{T_n+2}}}{T_1 \frac{F}{\left(1+y_0^f\right)^{T_1+1}} + ... + T_n \frac{F}{\left(1+y_0^f\right)^{T_n+1}}} \left(y_0^f\right)^2 \int_0^T \sigma_t^2 dt$$

This shows us that it is only because of some volatility on the swap rate that the CMS rate is different from the forward swap rate. Our result shows that the influence of the volatility is linear in the volatility of the whole process $\int_0^T \sigma_t^2 dt$.

5.5 Conclusion

In this chapter, we have seen that using martingale theory enables us to give a more robust proof of the convexity adjustment formula in the Black framework.

Looking for a definition of convexity, we classified the convexity adjustments into two categories: a correlation convexity, futures versus forward contracts and a modified schedule convexity, mainly the rest of the convexity adjustments. We explain on a static hedge the origin of the convexity. We derive convexity adjustment from a no-arbitrage proposition implied by a martingale condition. This enables us to give a definition of the convexity adjustment, with no approximation. Then making approximation, we show how to get a tractable closed formula, which encompassed previous results. We specify the error term between the approximated closed formula and the exact but non-explicit formula. We show that under certain conditions, this error term can be bounded by a "modified" Laplace Transform of the yield variable. Monte Carlo simulations show that the error is relatively small. One can consider the approximate formula a good estimate of the convexity adjustment.

There are many possible extensions to this chapter. The first one is to relax the hypothesis of a Black diffusion. This is more in agreement with the use of term structure models by financial institutions. However, the problem turns to be non-linear and complex. Its solving requires sophisticated approximation techniques like Wiener chaos, Cramers-Moyal expansion or the theory of stochastic perturbation (see Benhamou (2000b) reprinted as chapter 6 for a discussion and a solution by means of Wiener Chaos). A second development concerns the pricing of in-arrear derivatives. These derivatives are well-known for their convexity component. An approximate pricing can be obtained by using forward rates modified by the correct convexity adjustment, as explained in this chapter. Another extension to this work should be to compare the no-arbitrage dynamic approach with hedging models like the one developed by Besseminder (1991) and Neuberger (1999) among others. Last but not least, the same methodology could be applied to the convexity adjustment of futures against forward contracts, fact that has been studied empirically by French (1983), Park and Chen (1985) and Viswanath (1989) and that is still little explored.

Chapter 6

Pricing Convexity Adjustment with Wiener Chaos

Summary of the chapter

This chapter presents an approximated formula for the convexity adjustment of Constant Maturity Swap rates, using Wiener Chaos expansion, for multi-factor lognormal zero-coupon models. We derive closed formulae for CMS bonds and swaps and apply these results to various well-known one-factor models (Ho and Lee (1986), Amin and Jarrow (1992), Hull and White (1990), Mercurio and Moraleda (1996)). Quasi Monte Carlo simulations confirm the efficiency of the approximation. Its precision relies on the importance of second and higher order terms.

6.1 Introduction

Due to the main role of interest rates swap rates in the determination of long term rates, it has been of great relevance to develop exotic options that incorporate swap rates. This has led to new products that use the rate of a Constant Maturity Swap (CMS) as an underlying rate. These are very diverse, ranging from CMS swaps and bonds to more complicated ones like CMS swaptions, caps and other traditional exotic fixed income derivatives. These CMS derivatives are tailored instruments for trading the steepening or flattening of the yield curve, since one receives/pays, in the future, the swap rate (long term rate) and finances/borrows himself/herself with money market rates (short term rates). Even if there are other products for a trade on the steepening or flattening of the yield curve, like in-arrear derivatives, CMS derivatives have become more popular because they are more leveraged than their competitor derivatives and correspond to long duration investment.

A main limitation for pricing and hedging these derivatives has been the inability to get closed formula within a standard term-structure yield curve model. Usually, practitioners compare the CMS rate with the forward swap rate of the same maturity. Since in the CMS case, the investor pays/receives the swap rate only once, whereas in the case of the forward swap, during the whole life of the swap, this modified schedule leads to a difference between the two rates, classically called convexity adjustment. The term convexity refers to the convexity of a receiver swap prices with respect to the swap rate. Traditionally, this adjustment is calculated assuming that swap rates behave according to the Black Scholes (1973) hypotheses.

There has been extensive research for the so called Black Scholes convexity adjustment. Brotherton-Ratcliffe and Iben (1993) first derived an analytic approximation for the convexity adjustment in the case of bond yield. Other works completed the initial formula: Hull (1997) (extension to swap rates), Hart (1997) (better precision approximation), Kirikos and al (1997) (extension to a

Hull and White yield curve) and recently Benhamou (1999) (estimation of the approximation error).

However, when assuming that interest rates follow a diffusion process different from the Black-Scholes and Hull and White's ones, using the convexity adjustment in the Black Scholes setting is irrelevant. Indeed, since nowadays, almost all financial institutions rely on more realistic multi-factor term-structure models, the traditional formula looks old-fashioned and inappropriate. In this chapter, we offer a solution to this problem. Using approximations based on Wiener Chaos expansion, we provide an approximated formula for the convexity adjustment when assuming a multi-factor lognormal zero coupon model (Heath Jarrow hypotheses). This is consistent with most common term structure models.

The remainder of this chapter is organized as follows. In section 2, we explain the intuition of the convexity adjustment as well as the products based on CMS rates. In section 3, we give explicit formulae of a coupon paying a CMS rate when assuming a lognormal zero coupon bond model. In section 4, we explicit formulae for different term-structure models and compare the closed form results with the ones given by a Quasi Monte Carlo method. We conclude briefly in section 5. In appendix (section E.3 page 185), some key results on Wiener chaos expansion are presented as well as the proof of the approximation's theorem.

6.2 Convexity: intuition and CMS products

In this section, we explain intuitively the nature of the convexity adjustment as well as the CMS products.

6.2.1 Convexity of Swap rates

In the modern derivatives industry, two risks have emerged as intriguing and challenging for the management and control of secondary market risk: for equity

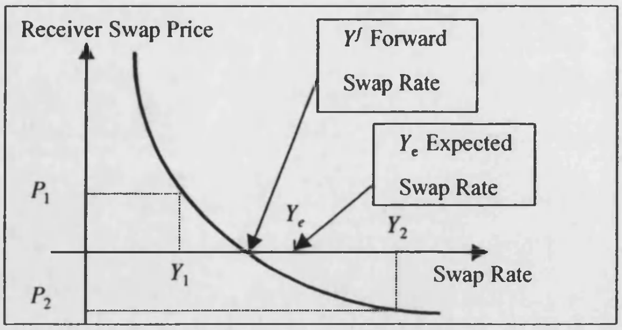

Figure 6.1: Convexity of the swap rate. In this graphic, we see that the convexity of the receiver swap price with respect to the swap rate leads to a higher expected swap rate than the forward swap rate, corresponding to a zero swap price.

derivatives, it has been the volatility smile and for fixed income derivatives, the convexity adjustment. Taking correctly these effects into account can provide competitive advantage for financial institutions.

Our chapter focuses on swap rates. Since the receiver swap price is a convex function of the swap rate, it is not correct to say that the expected swap is equal to the forward swap rate, defined as the rate at which the forward swap has zero value. This can be seen with the figure 6.1.

6.2.2 CMS derivatives

Since their early creation in 1981, interest rates swap contracts have grown very rapidly. The swap market represents now hundreds of billions of dollars each year. Reasons given to justify the existence of swaps range from market misperceptions, rationing and taxes (Bicksler and Chen (1986)) to agency problems (Wall and Pringle (1989)) and signalling (Titman (1992)). Subsequently, investors have been and are potentially looking for new instruments to risk-manage and hedge their positions as well as to speculate on the steepening or flattening of the yield curve. Indeed, the main interest of investors has turned out to be speculation. Even if other products like in-arrear derivatives enable to trade the flattening or the steepening of the yield curve, CMS derivatives are of particular interest since they are highly leveraged.

CMS derivatives are called CMS because they use a Constant Maturity Swap rate as the underlying rate. They are very diverse ranging from CMS swaps, CMS bonds to CMS swaptions and all other types of CMS exotics. Two major products are mainly traded over the counter: CMS swap and CMS bond. Logically, a CMS swap is an agreement to exchange a fixed rate for a swap rate, the latter referring to a swap of constant maturity. Assuming that our CMS swap starts in five years, is annual and is based on a swap rate of five year maturity, this typical contract will be the following: in five years, the investor will receive the swap rate of the swap starting in five years from today maturing in ten years. The investor will pay in return a fixed rate agreed in advance in the contract. One year later, that is in six years from today, the investor will receive the swap rate of the swap starting this time in six years from today maturing in eleven years. Again, the investor will pay the fixed rate. We see that at each payment, the investor receives a swap rate of a different swap. All the swap have in common to be settled at the date of the payment and to have the same maturity. A CMS bond is very similar to a CMS swap. It is a bond with coupons paying a swap rate of constant maturity. Therefore a CMS bond is exactly equal to the swap leg paying the swap rate. Since the swap leg paying the swap rate can be decomposed into each different payment, to price the CMS swap or CMS bond, we only need to price one payment of a swap rate. The value of a swap rate paid only once is called CMS rate value. The difference in value between the forward swap rate and this CMS rate is called the convexity adjustment.

Indeed, other CMS derivatives can be priced using forward rates increased by the convexity adjustment. The rest of the chapter will concentrate on the pricing of the CMS rate. Knowing these rates, one can use them to plug it into derivatives pricing formula to get an approached value of the CMS derivatives.

6.2.3 CMT bond and CMS swap

We consider a continuous trading economy with a trading interval $[0, \tau]$ for a fixed $\tau > 0$. The uncertainty in the economy is characterized by the probability space (Ω, F, Q) where Ω is the state space, F is the $\sigma-$algebra representing measurable events, and Q is the risk neutral probability measure uniquely defined in complete markets with no-arbitrage (Harrison, Kreps(1979) and Harrison, Pliska (1981)). We assume that information evolves according to the augmented right continuous complete filtration $\{F_t, t \in [0, \tau]\}$ generated by a standard (initialized at zero) $k-$dimensional Wiener Process (or Brownian motion). Let $(r_t)_{t<\tau}$ be the continuous spot rate, $B(t, T)_{t<\tau, T<\tau}$ the price at time t of a default-free forward zero coupon maturing at time T and $(y_T)_{T<\tau}$ the swap rate at time T. These three stochastic variables are supposed to be adapted to the information structure $(F_t)_{t\in[0,\tau]}$.

Referring to each coupon by the subscript variable i, the i^{th} coupon of a CMS bond pays the swap rate y_{T_i}, with a constant maturity specified in the contract, determined at a fixing date T_i often equal (eventually prior) to the payment date T_i^p. Therefore, the coupon value at time T_i^p is the swap rate times the nominal $y_{T_i}N$ while, at the fixing time, it is this value discounted by the forward zero coupon : $B(T_i, T_i^p)\, y_{T_i}N$. Assuming the no-arbitrage condition in a complete market, the value of one coupon C_i at time zero is obtained as the expectation under the risk neutral probability measure Q of the discounted payoff:

$$C_i = \mathbb{E}_Q\left[e^{-\int_0^{T_i} r_s ds} B(T_i, T_i^p)\, y_{T_i} N\right] \tag{6.1}$$

The total value at time zero of a N-nominal bond with m coupons with value at time zero $(C_i)_{i=1..m}$, with payment dates $(T_i^p)_{i=1..m}$, provided that the nominal N is paid at the end date T_m^p, is given by:

$$CMS\_Bond = \sum_{i=1}^{m} C_i + N * B(0, T_m^p) \tag{6.2}$$

In an interest rate CMS receiver swap, the fixed rate is received and the Constant Maturity Swap rate is paid. The different payment dates are also noted $T_1^p, ...T_m^p$.

The fixed leg valuation is easy. Its total value, denoted by V_F, is equal to the sum of all the discounted cash flows equal to the fixed rate R_{fixed}:

$$V_F = \sum_{i=1}^{m} R_{fixed} B\left(0, T_i^p\right)$$

The fixing dates for the swap rates are denoted by $T_1, ...T_m$. The CMS leg can be valuated as the sum of all the different coupons with value at time T_i, y_{T_i} and paid at time T_i^p. Its total value as of $t = 0$, denoted by V_{CMS}, is the sum of individual swap rate coupons:

$$V_{CMS} = \sum_{i=1}^{m} E_Q\left[e^{-\int_0^{T_i} r_s ds} B\left(T_i, T_i^p\right) y_{T_i}\right] \tag{6.3}$$

The price of the CMS swap is the difference of price between the two legs: $V_F - V_{CMS}$ for a receiver CMS swap and the opposite for a payer CMS swap. As a consequence, the rate R_{CMS_swap}, called the CMS swap rate, is the one which makes the value of the two legs equal:

$$R_{CMS\_swap} = \frac{V_{CMS}}{\sum_{i=1}^{m} B\left(0, T_i^p\right)} \tag{6.4}$$

The term of the denominator is classically called the sensitivity of the swap. The CMS swap rate is consequently the value of the CMS leg over the sensitivity of the swap.

As a conclusion of this subsection, CMS swap or CMS bonds are valued exactly with the same procedure. One needs to determine the exact value of a coupon paying the CMS rate. To calculate explicitly these quantities, we need to specify our interest rate model.

6.3 Calculating the convexity adjustment

In this section, we explain how to price the convexity adjustment with an approximated formula based on a Wiener Chaos expansion. Indeed, techniques based on perturbation theory or Kramers Moyal expansion could have also been used. Moreover, a recursive use of the Ito lemma gives exactly the same results. However, the framework given by Wiener Chaos expansion is much more powerful and leads to a straightforward calculation instead of very tedious ones.

6.3.1 Pricing framework

We assume that default-free zero coupon bonds are modelled by a lognormal k-multi-factor model, with a k-dimensional deterministic volatility vector denoted by $V(t,T) = (v_1(t,T),...,v_k(t,T))'$ verifying the Novikov condition $\forall T < \tau, \quad e^{\frac{1}{2}\int_0^T \|V(s,T)\|^2 ds} < +\infty$. This enables us to use probability measure change since this condition is sufficient for the Girsanov theorem. The default-free $T-$ maturity zero coupon bond price at time t is denoted by $B(t,T)$ and it is defined as the unique strong solution of the stochastic differential equation given under the risk neutral probability Q by:

$$\frac{dB(t,T)}{B(t,T)} = r_t dt + \langle V(t,T), dW_t \rangle \tag{6.5}$$

with $\langle V(t,T), dW_t \rangle = \sum_k v_k(t,T) dW_t^k$. The initial condition expresses that at maturity, the zero coupon bond is equal to the unity coupon $B(T,T) = 1$. Using traditional results (El-Karoui et al(1995)), we can define the forward neutral probability at time t, Q^t either by means of its Radon Nykodym derivatives with respect to the risk neutral probability measure or by the fact that $\widetilde{dW_s} = dW_s - V(s,t)ds$ is a standard Brownian motion under Q^t. We get that under this new probability measure, the bond price solution of the equation (6.5) can be written as a normalized Doleans martingale times the value of the forward zero coupon bond at time zero:

$$B(t,T) = \frac{B(0,T)}{B(0,t)} e^{\int_0^t \langle V(s,T)-V(s,t), \widetilde{dW_s} \rangle - \frac{1}{2}\int_0^t (\|V(s,T)-V(s,t)\|^2) ds} \tag{6.6}$$

To price a CMS swap/bond, we need to determine the value of one coupon, knowing that the total value of the swap/bond is the sum of the individual swap coupons. The core of the pricing problem is to determine the value at time zero,

Π_0, of a contingent claim that at a payment time T, gives the swap rate y_T fixed at time T, of a vanilla interest rate swap. The underlying interest rate swap has n equally separated payment dates : $T_1, ...T_n$. We ignore the issue of credit risk for valuating interest rate swap as described in Cooper and Mello (1991) and Duffie and Huang (1994). As proved for example in Musiela Ruttkowski(1997) page 389 equation (16.4)) the no-arbitrage condition gives a simple expression of the swap rate y_T with respect to the zero coupon bonds $(B(T, T_i))_{i=0..n}$

$$y_T = \frac{B(T, T_0) - B(T, T_n)}{\sum_{i=1}^{n} B(T, T_i)} \tag{6.7}$$

We then adopt the following definition of the CMS rate:

Definition 3 *The CMS rate is the expected value under the forward risk neutral probability measure at the payment time T of the swap rate y_T*

$$CMS\_Rate = \mathbb{E}_{Q_T}(y_T) \tag{6.8}$$

When the payment date T^p is different from the fixing date T^f, the above formula is modified in $CMS\_Rate = \mathbb{E}_{Q_{T^p}}(y_{T_f})$

The guiding idea of the chapter is to obtain an approximate formula for the expression above, by means of Wiener Chaos expansion. Let us introduce some notation. We call B_{T_i} the forward zero coupon bond:

$$B_{T_i} = \frac{B(0, T_i)}{B(0, T)}$$

Let the forward volatility $V_s^{(T,T_i)}$ be the volatility of a $T-$forward zero coupon maturing at time T_i:

$$V_s^{(T,T_i)} = V(s, T_i) - V(s, T)$$

Let $C(T_i, T_j)$ denote the (symmetric) correlation term between the return of the zero coupon bonds (mathematically between the logarithm of zero coupon bonds)

$$C(T_i, T_j) = \int_0^T \langle V_s^{(T,T_i)}, V_s^{(T,T_j)} \rangle \, ds$$

and K the sensitivity of the forward swap defined as the sum of the forward zero coupon bonds $K = \sum_{i=1}^{n} B_{T_i}$.

Definition 4 *Convexity adjustment CA is the difference between the CMS rate and the value today of the forward swap rate:*

$$CA = CMS\ rate - y^{forward} \tag{6.9}$$

The value today of the forward swap rate is given by the equation (6.7) with the time considered being zero leading to $y^{forward} = \frac{B_{T_0} - B_{T_n}}{K}$.

6.3.2 Closed formulae

The chapter's main result is the following approximation theorem. By means of approximations based on Wiener chaos, we can get a closed formula for the CMS rate.

Theorem 5 *Under the above assumptions, the convexity adjustment denoted CA can be expressed as a sum of correlation terms, plus an error term expressed with Landau notation as an $O\left(\|V_s(.,.)\|^4\right)$:*

$$CA = \begin{pmatrix} \frac{\sum_{i=1}^{n} B_{T_i}\left(B_{T_n} C(T_i,T_n) - B_{T_0} C(T_i,T_0)\right)}{K^2} \\ + y^{forward} \frac{\sum_{i,j=1}^{n} B_{T_i} B_{T_j} C(T_i,T_j)}{K^2} \end{pmatrix} + O\left(\|V_s(.,.)\|^4\right) \tag{6.10}$$

Proof: see section E.3 page 185.□

This theorem shows us that the convexity adjustment on a swap rate is a simple function of correlation terms. Interestingly, it is a linear function of the forward swap rate $y^{forward}$. The terms $B_{T_i} B_{T_n} C(T_i, T_n)$ (respectively $B_{T_i} B_{T_0} C(T_i, T_0)$) can be interpreted as the convexity adjustment between the zero coupon bonds $B(T, T_i)$ and $B(T, T_n)$ (respectively $B(T, T_i)$ and $B(T, T_0)$) as the following proof states:

Proposition 16 *The convexity adjustment CA between two zero coupons bonds can be expressed in terms of the correlation term through:*

$$\begin{aligned} CA &= \mathbb{E}^{Q_T}\left[B(T,T_i)\,B(T,T_j)\right] - \mathbb{E}^{Q_T}\left[B(T,T_i)\right]\mathbb{E}^{Q_T}\left[B(T,T_i)\right] \quad (6.11) \\ &= B_{T_i}B_{T_j}C(T_i,T_j) + O\left(\|V_s(.,.)\|^4\right) \end{aligned}$$

Proof: Plugging in the expression of the zero coupon bond (6.6), the convexity adjustment can is given by:

$$CA = B_{T_i}B_{T_j}\mathbb{E}^{Q_T}\left[e^{\int_0^t\left\langle V_s^{(T,T_i)}+V_s^{(T,T_j)},\widetilde{dW}_s\right\rangle - \frac{1}{2}\int_0^t\left(\left\|V_s^{(T,T_i)}\right\|^2+\left\|V_s^{(T,T_j)}\right\|\right)ds} - 1\right]$$

Using the fact that $e^{\int_0^t\langle f(s),\widetilde{dW}_s\rangle - \frac{1}{2}\int_0^t(\|f(s)\|^2)ds}$ is a martingale for any deterministic function $f(.)$, this expression simplifies to $B_{T_i}B_{T_j}\left(e^{\int_0^t\left\langle V_s^{(T,T_i)},V_s^{(T,T_i)}\right\rangle ds} - 1\right)$, which leads to the result (6.11) when taking a Taylor expansion up to the first order.□

Corollary 7 *When the underlying CMS swap is a spot CMS swap; then: $T = T_0$ and the formula simplifies to*

$$CA = \begin{pmatrix} \frac{B_{T_n}\sum_{i=1}^n B_{T_i}C(T_i,T_n)}{K^2} \\ +y^{forward}\frac{\sum_{i,j=1}^n B_{T_i}B_{T_j}C(T_i,T_j)}{K^2} \end{pmatrix} \tag{6.12}$$

Proof: When the CMS swap is a spot CMS swap, the correlation term $C(T_0,T_i)$ (convexity term due to the fact that we have a forward swap) becomes zero.□

In this latter case, equation (6.12), the convexity adjustment is always positive. This result can be easily derived within an elementary term structure model (since we notice that the rate of a forward bond should always be above the forward rate). Put another way, for this CMS, it is pure convexity.

The previous results are approximation formulae. Specifying the error term as the difference between the intractable expression of the convexity adjustment and the closed formula obtained by Wiener Chaos, we can stipulate an upper bound for the error term. Indeed, using Wiener Chaos expansion, we conclude that the error term is dominated by the following quantity $O_3 = O\left(\left(\int_{s_1\leq s_2\leq s_3\leq T}\left\|V_{s_1}^{(T,T_i)}\right\|^2 ... \left\|V_{s_3}^{(T,T_i)}\right\|^2 ds_1...ds_3\right)^{1/2}\right)$. This indicates that our approximation is all the more efficient than the volatility is small.

6.3.3 Extension

It turns out that some CMS rates present a delayed adjustment. This case is more complicated to handle. However, the same methodology gives a closed formula for the price.

Theorem 6 *In the case of a payment date T^p different from the fixing time T, the above expression gets additional terms due to delayed adjustment. The convexity adjustment is then given by:*

$$CA = \left(\begin{array}{c} \frac{\sum_{i=1}^n B_{T_i}\left(B_{T_n}C(T_i,T_n)-B_{T_0}C(T_i,T_0)\right)}{K^2} \\ +y^{forward}\left(\frac{\sum_{i,j=1}^n B_{T_i}B_{T_j}(C(T_i,T_j)-C(T_i,T^p))}{K^2}\right)\end{array}\right) \tag{6.13}$$

Proof: The proof goes along the same lines as the one of theorem (5) and can be done using thee same techniques. □

Corollary 8 *The convexity adjustment can also be expressed as:*

$$CA = \frac{\sum_{i,j=1}^n B_{T_i}B_{T_j}\left(\begin{array}{c} B_{T_n}\left(C\left(T_i,T_n\right)-C\left(T_i,T_j\right)+C\left(T_i,T^p\right)\right) \\ -B_{T_0}\left(C\left(T_i,T_0\right)-C\left(T_i,T_j\right)+C\left(T_i,T^p\right)\right)\end{array}\right)}{K^3} \tag{6.14}$$

The interpretation is simple. This formula expresses the convexity adjustment as the difference of correlation terms. Since these terms are small, this suggests already that the convexity adjustment is small. This a posteriori justifies our approximate method where we truncate the Wiener Chaos expansion after the second order. Indeed, the theoretical justification of this truncation can be found as well in the theorem of Pawula which states that a positive transition probability, the Kramers-Moyal expansion (similar to the Wiener Chaos one) may be stopped either after the first term or after the second term. If it does not stop after the second term, it must contain an infinite number of terms.

For the interpretation of this convexity adjustment, we assume that the correlation term $C(T_i, T_j)$ is an increasing function of both T_i and T_j. Let us assume that the payment date T^p is prior to the payment dates of the underlying swap $(T_i)_{i=1..n}$, i.e., $T_i > T^p$ for every i. Consequently, the first term in the right hand side of equation (6.13) $y^{forward}S_1$, of the same sign as $\sum_{i=1}^{n} B_{T_i}B_{T_j}\left(C(T_i, T_j) - C(T_i, T^p)\right)$ is positive. The other term is closely connected to the sign of

$$\sum_{i=1}^{n} B_{T_i}\left(B_{T_n}\left(C(T_i, T_n) - C(T^p, T_n)\right) - B_{T_0}\left(C(T_i, T_0) - C(T^p, T_0)\right)\right)$$

This leads to think that this expression, expressed as a difference, should be relatively small and in many cases, smaller than the first correction term. In the case it is non positive, it should be slightly negative. This result is of great significance since it states that under non-classical conditions, the expected swap rate can be lower than its corresponding forward swap rate, mainly due to a negative delayed adjustment.

6.4 Application and results

In this section, we apply the formula to different types of stochastic interest rate models.

6.4.1 Application to different models

In this section, we apply our closed formula to various one-factor interest rates models. Therefore, for all of them, the number of factors k equals 1.

Ho and Lee model

Among the early one-factor interest rate term-structure models, the Ho and Lee (1986) model was originally stated in the form of a binomial tree of bond prices. After the appearance of the Heath Jarrow Morton formalism, this model has been rewritten in the form of a diffusion of the zero coupons bonds:

$$\frac{dB(t,T)}{B(t,T)} = r_t dt + \sigma (T-t)\, dW_t$$

It has been observed that the volatility of zero coupons bonds was decreasing with time. This model assume a linear decrease. The forward volatility as well as the correlation were consequently taken to be of the form:

$$V_s^{(T,T_i)} = \sigma (T_i - T)$$

and $C(T_i,T_j) = \sigma^2 (T_i - T)(T_j - T)\dot{T}$. The convexity adjustment formula (6.13) can then be expressed as a function of forward zero coupon and the volatility:

$$convexity = \begin{pmatrix} \sigma^2 \frac{\left(\sum_{i=1}^n B_{T_i} T(T_i - T^p)\right)\left(B_{T_n}(T_n - T) - B_{T_0}(T_0 - T)\right)}{K^2} \\ + y^{forward} \sigma^2 \frac{\sum_{i,j=1}^n B_{T_i} B_{T_j} (T_i - T)(T_j - T^p) T}{K^2} \end{pmatrix}$$

Amin and Jarrow model

The purpose of the Amin and Jarrow (1992) model was to take into account a phenomenon called the volatility hump. Basically, the volatility of zero-coupon bonds is first increasing and then decreasing. Amin and Jarrow offered to model the volatility as a second order polynomial given by $\sigma_0 (T-t) + \sigma_1 \frac{(T-t)^2}{2}$. This leads to the following expression for the zero coupons bonds diffusion

$$\frac{dB(t,T)}{B(t,T)} = r_t dt + \left(\sigma_0 (T-t) + \sigma_1 \frac{(T-t)^2}{2} \right) dW_t$$

The forward volatility is expressed as a second order polynomial expression of the different maturities $V_s^{(T,T_i)} = \left(\sigma_0 (T_i - T) + \sigma_1 \frac{\left[(T_i-t)^2 - (T-t)^2\right]}{2} \right)$ whereas the correlation term, which is more complicated, is expressed in this particular case as a sum of four terms:

$$C(T_i, T_j) = A_1 + A_2 + A_3 + A_4$$

with

$$\begin{aligned}
A_1 &= \sigma_0^2 (T_i - T)(T_i - T) T \\
A_2 &= \sigma_0 \sigma_1 (T_i - T) \tfrac{1}{2} \left[\tfrac{T_j^3 - (T_j - T)^3}{3} - \tfrac{T^3}{3} \right] \\
A_3 &= \sigma_0 \sigma_1 (T_j - T) \tfrac{1}{2} \left[\tfrac{T_i^3 - (T_i - T)^3}{3} - \tfrac{T^3}{3} \right] \\
A_4 &= \sigma_1^2 (T_i - T)(T_i - T) \left[\tfrac{1}{4} \left(T T_i T_j + \tfrac{1}{3} T^3 \right) \right]
\end{aligned}$$

The convexity is then calculated using the convexity adjustment formula (6.13).

Hull and White model

This model represents a significant breakthrough compared to the Ho&Lee model. It is a one factor model, extendable to a two-factor or more ones, that enables both to incorporate deterministic mean-reverting features and to allow perfect matching of an arbitrary yield curve. It has become very popular among practitioners since one can derive closed forms for vanilla interest rates derivatives like cap/floor and swaption (one factor version). This implies a quick calibration. The time-dependent volatility version of this model has been advocated to be unstable and is consequently not used in practice. We will give here the convexity adjustment for the classic Hull and White (1990) model with a constant volatility σ and constant mean reverting parameter λ. In this model, in its formulation on zero coupons bonds, the latter follow a diffusion given by:

$$\frac{dB(t,T)}{B(t,T)} = r_t dt + \sigma \frac{1 - e^{-\lambda(T-t)}}{\lambda} dW_t.$$

The volatility structure is realistic since it is decreasing with time. It does not allow for the hump and this can be seen as the main drawback of this model. In this case,

$$V(s,t) = \sigma \frac{1 - e^{-\lambda(T-t)}}{\lambda}$$

and the forward volatility is given by $V_s^{(T,T_i)} = \sigma \frac{e^{-\lambda T} - e^{-\lambda T_i}}{\lambda} e^{\lambda s}$ whereas the correlation term becomes:

$$C(T_i, T_j) = \sigma^2 \frac{1 - e^{-\lambda (T_j - T)}}{\lambda} \frac{1 - e^{-\lambda (T_i - T)}}{\lambda} \frac{1 - e^{-2\lambda T}}{2\lambda}$$

It is worth noticing that this model implies a lower correlation between the different rates than the Ho&Lee model. As for the convexity adjustment, we get the following formula:

$convexity = HW_1 + HW_2$ with:

$$HW_1 = \sigma^2 y^{forward} \frac{\sum_{i,j=1}^n B_{T_i} B_{T_j} \frac{1-e^{-2\lambda T}}{2\lambda} \frac{1-e^{-\lambda(T_i - T)}}{\lambda} \left(\frac{e^{-\lambda(T^p - T)} - e^{-\lambda \left(T_j - T\right)}}{\lambda} \right)}{K^2}$$

$$HW_2 = \sigma^2 \frac{\sum_{i=1}^n B_{T_i} \frac{1-e^{-2\lambda T}}{2\lambda} \left(\frac{e^{-\lambda(T^p - T)} - e^{-\lambda(T_i - T)}}{\lambda} \right) \left(B_{T_n} \frac{1-e^{-\lambda(T_n - T)}}{\lambda} - B_{T_0} \frac{1-e^{-\lambda(T_0 - T)}}{\lambda} \right)}{K^2}$$

or for the simplified version $T = T_0 = T^p$

$$HW_1 = \sigma^2 \frac{\sum_{i=1}^n B_{T_i} \frac{1-e^{-2\lambda T}}{2\lambda} \frac{1-e^{-\lambda(T_i - T)}}{\lambda}}{K^2} \sum_{j=1}^n B_{T_j} \left(\frac{1 - e^{-\lambda(T_j - T)}}{\lambda} \right) y^{forward}$$

$$HW_2 = \sigma^2 \frac{\sum_{i=1}^n B_{T_i} \frac{1-e^{-2\lambda T}}{2\lambda} \left(\frac{1-e^{-\lambda(T_i - T)}}{\lambda} \right)}{K^2} \left(B_{T_n} \frac{1 - e^{-\lambda(T_n - T)}}{\lambda} \right)$$

Mercurio and Moraleda model

Last but not least, we examine the case of the Mercurio and Moraleda (1996) model. This model has been introduced like the Amin and Jarrow model to take into account the volatility hump. Mercurio and Moraleda (1996) suggested to use a combination of Ho and Lee and Hull and White volatility form to get another volatility in which the hump would be modelled more realistically while still analytical tractable. This leads to the following diffusion for the zero coupons bonds:

$$\frac{dB(t,T)}{B(t,T)} = r_t dt + \sigma \left(\frac{1 - e^{-\lambda(T-t)}}{\lambda} + \gamma \left(\frac{1 - e^{-\lambda(T-t)}}{\lambda^2} - \frac{(T-t)\, e^{-\lambda(T-t)}}{\lambda} \right) \right) dW_t$$

In this particular case, the volatility structure takes the following form: $V_s^{(T,T_i)} = g(s,T_i) + f(s,T_i)$

$$g(s,T_i) = \sigma \frac{e^{-\lambda T} - e^{-\lambda T_i}}{\lambda} e^{\lambda s}$$
$$f(s,T_i) = \gamma\sigma \left(\frac{(T_i - s)e^{-\lambda(T_i - s)} - (T-s)e^{-\lambda(T-s)}}{\lambda} + e^{\lambda s} \frac{e^{-\lambda T} - e^{-\lambda T_i}}{\lambda^2} \right)$$

and

$$C(T_i,T_j) = M_{21} + M_{22} + M_{23} + M_{24}$$
$$M_{21} = \sigma^2 \frac{1-e^{-\lambda\left(T_j - T\right)}}{\lambda} \frac{1-e^{-\lambda(T_i - T)}}{\lambda} \frac{1-e^{-2\lambda T}}{2\lambda}$$
$$M_{22} = \int_0^T f(s,T_i) f(s,T_j)\, ds$$
$$M_{23} = \int_0^T g(s,T_i) f(s,T_j)\, ds$$
$$M_{24} = \int_0^T g(s,T_j) f(s,T_i)\, ds$$

or after simplifying:

$$\begin{aligned} M_{22} &= \psi(T_i,T_j) \\ M_{23} &= \alpha(i)\beta(j) \\ M_{24} &= \alpha(j)\beta(i) \end{aligned}$$

$$\alpha(i) = \gamma\sigma^2 \frac{1-e^{-\lambda(T_i - T)}}{\lambda}$$

$$\beta(j) = \left(\begin{array}{c} \left(\frac{T_j e^{-\lambda\left(T_j - T\right)} - T}{\lambda} \frac{1-e^{-2\lambda T}}{2\lambda} + \frac{1-e^{-\lambda\left(T_j - T\right)}}{\lambda} \frac{2\lambda T - 1 + e^{-2\lambda T}}{4\lambda^2} \right) \\ + \frac{1-e^{-\lambda\left(T_j - T\right)}}{\lambda^2} \frac{1-e^{-2\lambda T}}{2\lambda} \end{array} \right)$$

$$\psi(T_i,T_j) = (\gamma\sigma)^2 \left(\begin{array}{c} \left(\frac{1-e^{-\lambda(T_i - T)}}{\lambda}\right) \left(\frac{1-e^{-\lambda\left(T_j - T\right)}}{\lambda}\right) \left(\frac{2\lambda^2 T^2 - 2T\lambda + 1 - e^{-2T\lambda}}{4\lambda^3}\right) \\ + \left(\frac{T_i e^{-\lambda(T_i - T)} - T}{\lambda} + \frac{1-e^{-\lambda(T_i - T)}}{\lambda^2} \right) \left(\frac{1-e^{-\lambda\left(T_j - T\right)}}{\lambda} \right) \frac{2T\lambda - 1 + e^{-2T\lambda}}{4\lambda^2} \\ + \left(\frac{T_j e^{-\lambda\left(T_j - T\right)} - T}{\lambda} + \frac{1-e^{-\lambda\left(T_j - T\right)}}{\lambda^2} \right) \left(\frac{1-e^{-\lambda(T_i - T)}}{\lambda} \right) \frac{2T\lambda - 1 + e^{-2T\lambda}}{4\lambda^2} \\ + \left(\begin{array}{c} \frac{T_i e^{-\lambda(T_i - T)} - T}{\lambda} \\ + \frac{1-e^{-\lambda(T_i - T)}}{\lambda^2} \end{array} \right) \left(\begin{array}{c} \frac{T_j e^{-\lambda\left(T_j - T\right)} - T}{\lambda} \\ + \frac{1-e^{-\lambda\left(T_j - T\right)}}{\lambda^2} \end{array} \right) \frac{1-e^{-2T\lambda}}{2\lambda} \end{array} \right)$$

The convexity is then calculated using the convexity adjustment formula (6.13)

6.4.2 Results for a standard contract

In this section, we gave some results for the case of a Ho and Lee model, a one factor Hull and White model, and a Mercurio and Moraleda model. We compared them to the results we got from a Quasi Monte Carlo simulation with 10,000 random draws. The difference between the two was negligible. These results are summarized in the four tables given in the appendix section: table E.1, E.2, E.3 and E.4. Interestingly, convexity adjustment differ from one model to another but all adjustment are approximately the same.

6.5 Conclusion

In this chapter, we have seen that Wiener Chaos theory provides closed formulae which are very good approximations of the correct result. The interesting point is that this methodology is quite general and could also be applied to many other products where the payoff function is a non linear function of lognormal variables.

Indeed, there are many extensions to this chapter, like the one in deriving the convexity adjustment between futures and forwards contracts. A second development, quite promising, is to apply Wiener chaos technique to other option pricing problems.

General Conclusion and Future Research

The final word of this dissertation is that these new techniques, Malliavin calculus and Wiener chaos expansion, have turned out to be extremely powerful for different pricing problems where no explicit formula could be found. The interest of the Malliavin calculus theory is that it imposes few restrictions on the payoff function. Therefore, it can handle very general situations. We have shown in this dissertation how to use it properly to get very efficient Monte Carlo schemes for the computation of the Greeks. We have contrasted this approach with a new numerical technique based on Fast Fourier Transform. The Wiener chaos expansion theory enables to calculate approximations of small perturbations. We have used it to get accurate pricing formulae for the convexity adjustment.

We conjecture that these methods are going to be more and more used in the coming following years. Many possible developments and extensions have been already mentioned in the preceding chapters like the extension of the Malliavin weighted scheme to continuous lookback options, to stochastic interest rate models, the advanced study of local type Malliavin weighted scheme, the extension of the Wiener Chaos expansion theory to other type of term structure models.

Appendix A

Malliavin calculus

This appendix summarizes the main results needed for this dissertation. It is intended for the reader who is unfamiliar with this domain of stochastic calculus. A rough idea of Malliavin calculus is that Malliavin calculus is the calculus of variations in a stochastic framework. Or, comparing with the deterministic framework, Ito calculus would correspond to the ordinary derivative in infinitesimal calculus while the Malliavin derivative on Wiener Space to the Frechet derivative on a function space. This theory was initiated by Malliavin and further developed by Stroock, Bismut, Watanabe and others. The original motivation was to provide a probabilistic proof of Hormander's sum of squares theorem. One of the important conclusions is the existence of the adjoint operator of the Malliavin derivative called the Skorohod integral which has the elegant property to be an extension of the Ito integral for non adapted process. The great advantage of this theory is also that it allows the formulation of solutions of stochastic differential equations in those cases where the solution is not adapted to the Brownian filtration. For further reference see Nualart (1995). All the definitions and propositions below are taken from this book.

A.1 Key Results

Let $(\Omega, \mathcal{F}, P)$ be the classical Wiener space which supports a standard d-dimensional Wiener process (Brownian motion) $\{W_t = (W_t^1, ..., W_t^d), t \in [0,T]\}$. Let $(L^2([0,T]))^d$ be the real separable Hilbert space of d-dimensional real functions defined on $[0,T]$ with squared integrable, with its canonical scalar product $\langle .,. \rangle$ and its norm $\|\|$. We denote by W the isonormal Gaussian process (Nualart (1995) definition 1.1.1 pp 4) associated with $(L^2([0,T]))^d$ and defined as for h element of

$$W(h) \equiv \left(W^1(h), ..., W^d(h)\right)$$

where the sign $\equiv$ stands for a definition with

$$\begin{aligned} W(h) &\equiv \int_0^T h(t)dW_t \\ W^j(h) &\equiv \int_0^T h^j(t)dW_t^j \end{aligned}$$

where the symbol $\equiv$ stands for a definition We denote by $C_p^\infty(.)$ (respectively $C_b^\infty(.)$) the set of infinitely continuously differentiable functions f such that f and all its partial derivatives have polynomial growth (respectively bounded). Let $\mathcal{S}$ be the set of stochastic functions F of the form:

$$F = f(\int_0^T h_1(t)dW_t, ..., \int_0^T h_n(t)dW_t)$$

where f belongs to $C_p^\infty(\mathbb{R}^{nd})$ $f(x_{11}, ..., x_{d1}, ..., x_{1n}, ..., x_{dn})$ and $h_1, ..., h_n$ to $C_b^\infty(L^2([0,T])^d)$. $\mathcal{D}^{1,2}$ designates the Banach space, completion of $\mathcal{S}$ with respect to the norm

$$\| F \|_{1,2} = (\mathbb{E}(F^2))^{1/2} + (\mathbb{E}(\int_0^T (D_tF)^2 dt)^{1/2}$$

We then have the interesting definition of the Malliavin derivative (Nualart (1995) definition 1.2.1 pp 24):

Definition 5 *The Malliavin derivative of a stochastic function of a type mentioned above is the stochastic process* $\{D_tF, t \in [0,T]\}$ *or equivalently the random* *gradient* $DF = \left(D^1F, ..., D^NF\right)$ *given by*

$$D_t^jF = \sum_{i=1}^{n} \frac{\partial f}{\partial x_{ji}}(\int_0^T h_1(t)\ dW_t, ..., \int_0^T h_n(t)\ dW_t)h_i(t), \qquad t \in [0,T] \ \textit{a.s.}$$

The intuitive idea behind this Frechet derivative in Wiener space is to differentiate F with respect to the chance parameter $w \in \Omega$.

The two major facts that are heavily used in the present dissertation are the following ones:

- the integration by parts formula and its link to the Skorohod integral which turns out to be an extension of the Ito integral (Nualart (1995) Chapter I I.2, I.3) and
- the formula of the Malliavin derivative of the solution of a stochastic differential equation (Nualart (1995) Chapter II II.2, II.3) with respect to its first variation process.

The integration by parts is closely linked to the Skorohod integral. Indeed, the adjoint of the closed unbounded linear operator $D : \mathcal{D}^{1,2} \to L^2\left(\Omega \times [0,T]\right)^d$ is usually denoted by δ and is called the Skorohod integral. Its domain $Dom\left(\delta\right)$ can be characterized as the set of measurable process $u \in L^2\left(\Omega \times [0,T]\right)$ such that there exists a positive constant C for which

$$\mathbb{E}\left(\int_0^T D_tFu_tdt\right) \leq C\left(u\right)\ \|F\|_{1,2.}$$

for all $F \in D^{1,2}$. Then the Skorohod integral for $u \in Dom\left(\delta\right)$ is the square-integrable stochastic variable determined by the duality relation

$$\mathbb{E}\left(F\delta\left(u\right)\right) = \mathbb{E}\left(\int_0^T D_tFu_tdt\right) \qquad \forall F \in D^{1,2} \tag{A.1}$$

This defines the integration by parts which is at the core of our proof. The major result concerning the Skorohod integral is its link to the Ito integral. The Skorohod integral turns out to be an extension of the classical Ito integral and even it allows the integration of processes that are not necessarily adapted.

Therefore if u is an adapted process in $L^2(\Omega \times [0,T])$ then we have (Nualart (1995) Proposition 1.3.4 pp 41)

$$\begin{aligned}\delta(u) &= \int_0^T u_t dW_t \\ &= \sum_{i=1}^{d} \int_0^T u_t^i dW_t^i\end{aligned}$$

We have also the classical chain rule useful to the work in hands:

Proposition 17 *chain rule (Nualart (1995) Proposition 1.2.2 pp 29)*

Let $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a function continuously differentiable with bounded partial derivatives. Suppose that $F = (F_1, ..., F_n)$ is a random vector whose components belong to the space $D^{1,2}$. Then $\varphi(F) \in D^{1,2}$ and

$$D_t\varphi(F) = \sum_{i=1}^{n} \nabla_i\varphi(F)\ D_tF_i \qquad t \in [0,T] \text{ a.s.}$$

The chain rule can be extended with only a Lipschitz condition (Nualart (1995) Proposition 1.2.3 pp 30)

Proposition 18 *Let F be a F_T-adapted stochastic process of $D^{1,2}$, then $\forall u \in Dom(\delta)$*

$$\delta(Fu) = F\delta(u) - \int_0^T D_tF\ u(t)\ dt$$

The second important fact concerns the Malliavin derivative of the solution of a stochastic differential equation:

Proposition 19 *Derivative of a function being a solution of a Stochastic Differential Equation with respect to its initial condition*

Let $\{X_t, t \in [0,T]\}$ be a vector with value in $\mathbb{R}^n$ solution of the following SDE

$$\begin{aligned} dX_t &= b(t,X_t).dt + \sigma(t,X_t).dW_t \qquad &\text{(A.2)} \\ X_0 &= x, \qquad x \in \mathbb{R}^n & \end{aligned}$$

or written in the integral form

$$X_t = x + \int_0^t b(s, X_s).ds + \int_0^t \sigma(s, X_s).dW_s$$

We also supposed that $b()$, $\sigma()$ are globally Lipschitz functions with linear growth and continuously differentiable. Then $(X_t, t \in [0,T])$ belongs to $D^{1,2}$ (Nualart (1995) Theorem 2.2.1 pp 102-104) and

$$D_r X_t = \left(\sigma(r, X_r) + \int_r^t \nabla_2 b(s, X_s) D_r X_s ds + \int_r^t \nabla_2 \sigma(s, X_s) D_r X_s dW_s\right) 1_{\{r \leq t\}}$$

Thus $(D_r X_t)_{t \geq r}$ is a solution of the following Stochastic Differential Equation:

$$\begin{aligned} dZ_t &= \nabla_2 b(t, X_t) Z_r dr + \nabla_2 \sigma(t, X_t) Z_t dW_t \quad r \leq t \\ Z_{t=r} &= \sigma(r, X_r) \end{aligned}$$

One of the key mathematical Malliavin derivative function for our dissertation is the derivative function of the process with respect to its initial condition (useful for the delta and gamma computation). The derivative of our process $\{X_t, t \in [0,T]\}$ noted $\{Y_t = \frac{\partial X_t}{\partial x}, t \in [0,T]\}$ with respect to its initial condition x is a solution of the following SDE :

$$\begin{aligned} dY_t &= \nabla_2 b(t, X_t) Y_t \, dt + \nabla_2 \sigma(t, X_t) Y_t \, dW_t \\ Y_0 &= I_n \end{aligned}$$

where I_n is the $\mathbb{R}^n$ identity matrix. The interesting link between the two processes $\{X_t, t \in [0,T]\}$ and $\{Y_t, t \in [0,T]\}$ is then given by the following proposition (Nualart (1995) equation 2.38 page 109)

$$D_s X_t = Y_t Y_s^{-1} \sigma(s, X_s) \; 1_{\{s \leq t\}} a.s. \tag{A.3}$$

(See also Elworthy (1992)). One important consequence is, when σ is moreover hypoelliptic, that for $s \leq t$

$$Y_t = D_s X_t Y_s \sigma^{-1}(s, X_s)$$

that is

$$
\begin{aligned}
Y_t &= \int_0^T \frac{Y_t}{T} ds \\
&= \frac{1}{T} \int_0^T D_s X_t Y_s \sigma^{-1}(s, X_s) ds
\end{aligned}
$$

A.2 The Elworthy formula

The motivation of chapter 1 is very similar to the one used to derived the Elworthy formula (Elworthy (1982) and (1992)). The Elworthy formula is summarized by the following proposition:

Proposition 20 *Elworthy formula*
If $(X_t, t \in [0,T])$ is the solution of the SDE (Stochastic Differential Equation)

$$X_t = x + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s$$

with $b()$, $\sigma()$ globally Lipschitz functions with linear growth and continuously differentiable, then for any function $\varphi \in C_b^1(\mathbb{R}^n)$ and for any $t > 0$, we have

$$
\begin{aligned}
\mathbb{E}\left[\partial_i \varphi(X_t)\right] &= \mathbb{E}\left[\varphi(X_t)\, \delta\left(\frac{1}{t}\sigma^{-1}(s, X_s) Y_s\right)\right] \\
&= \mathbb{E}\left[\varphi(X_t) \frac{1}{t} \int_0^t \sigma^{-1}(s, X_s) Y_s dW_s\right]
\end{aligned}
$$

Proof: see Elworthy (1982) and (1992).□

Appendix B

Technical Proofs of Chapter 1

B.1 Proof of the proposition 1 concerning the link between first variation process and sensibility of the underlying to perturbation page 36

The proof is only given for the Z_t^{rho} process. It is identical for Z_t^{vega}. To prove proposition (1), we first show that the process $\left(Z_t^{rho}\right)_{t\in[0,T]}$ verifies a stochastic differential equation (B.1). Since the two process $\left(Z_t^{rho}\right)_{t\in[0,T]}$ and $\left(\int_0^t Y_t Y_s^{-1}\widetilde{b}(s,X_s)\,ds\right)_{t\in[0,T]}$ verify the same SDE (B.1) and have the same initial conditions, they are equal according to the stochastic version of the Cauchy Lipschitz theorem.□

We now prove the lemma about the stochastic differential equation (B.1):

Lemma 1 *Under the assumption of continuous differentiability of b,σ with bounded derivatives, the process $\left(Z_t^{rho}\right)_{t\in[0,T]}$ defined by (1.16) is the unique solution of the following stochastic differential equation*

$$dZ_t = \left(\widetilde{b}(t,X_t) + Z_t b'(t,X_t)\right)dt + Z_t\sigma'(t,X_t)\,dW_t \tag{B.1}$$

with initial condition $Z_0 = 0_n$.

Proof: Solving the diffusion equation (1.12) with the initial condition $\left(X_{t=0}^{\varepsilon,rho}=x\right)$ gives

$$X_t^{\varepsilon,rho}=x+\int_0^t\left[b\left(s,X_s^{\varepsilon,rho}\right)+\varepsilon\widetilde{b}\left(s,X_s^{\varepsilon,rho}\right)\right]ds+\int_0^t\sigma\left(s,X_s^{\varepsilon,rho}\right)dW_s$$

For $\varepsilon\neq 0$

$$\frac{X_t^{\varepsilon,rho}-X_t}{\varepsilon}$$

$$=\int_0^t\left(\widetilde{b}\left(s,X_s^{\varepsilon,rho}\right)+\frac{b\left(s,X_s^{\varepsilon,rho}\right)-b\left(s,X_s\right)}{\varepsilon}\right)ds+\int_0^t\frac{\sigma\left(s,X_s^{\varepsilon,rho}\right)-\sigma\left(s,X_s\right)}{\varepsilon}dW_s$$

Using the hypothesis that b,σ are continuously differentiable with bounded derivative, as well as the continuity of $X_s^{\varepsilon,rho}$ in ε with its limit equal to the non-perturbed process $(X_t)_{t\in[0,T]}$, we can show that the Gateau derivative $\left(Z_t^{rho}\right)_{t\in[0,T]}$ of the drift-perturbed underlying process $\left\{X_t^{\varepsilon,rho},t\in[0,T]\right\}$along the direction $\widetilde{b}$ can be expressed as:

$$Z_t^{rho}=\int_0^t\left(\widetilde{b}\left(s,X_s\right)+Z_s^{rho}b'\left(s,X_s\right)\right)ds+\int_0^t Z_s^{rho}\sigma'\left(s,X_s\right)dW_s$$

which in its differential form is exactly equal to the result. The uniqueness is then given by the stochastic version of the Cauchy Lipschitz theorem.□

B.2 Proof of the delta formula (1.M1)

In this section, we prove that the weight function for the delta should satisfy necessary and sufficient conditions. The proof is given for the case of a stochastic interest rate depending both on time and the underlying. As a special case, we derive the necessary and sufficient conditions given in table 1.1 when the interest rate is only a function of time. For the sake of simplicity, we denote in this section w^{delta} by w,and the derivative with respect to the second variable by a prime . The part of the proof based on integration by parts is quite short and goes along the same line as the one of Elworthy (1992). The technical difficulty here is to justify rigorously the use of weaker assumptions. The proof can be divided into three major steps:

- first preliminary: weaker conditions on the payoff function f: show that if the result holds for any function of C_K^∞ (set of infinitely differentiable functions with compact support), it also holds for any element of L^2.
- second preliminary: interchange of the order of differentiation and expectation: show that one can interchange the order of differentiation and expectation.
- integration by parts:
  - necessary condition so as to fulfil an integration by parts.
  - sufficient condition so as to fulfil an integration by parts.

Step 1: Weaker assumptions

The idea of the first technical point is the following: taking f as an element of L^2 is the same as assuming f infinitely differentiable with a compact support. It is based on a density argument using Cauchy Schwartz inequality and the continuity of the expectation operator.

More precisely, let assume the result is true for any function of C_K^∞ (set of infinitely differentiable functions with compact support). Let f be now only in L^2. Using the density of $C_K^\infty[0,T]$ in L^2, there exists a sequence $(f_n)_{n\in N}$ of C_K^∞ elements that converges to f in L^2. Let 's denote by $u(x) = \mathbb{E}_x^Q[F]$ and $u_n = \mathbb{E}_x^Q[F_n]$ the prices associated with the discounted payoff functions F and F_n and x as the starting point of the underlying security price. Since L^2 convergence implies L^1 convergence, we know that the set of functions u_n converges simply to the function u.

$$\forall x \in \mathbb{R} \quad u_n(x) \underset{n\to\infty}{\to} u(x)$$

Since the result is true for payoff functions element of C_K^∞, the derivative of the u_n function can be written as the expectation of the discounted payoff function f_n times a suitable "Malliavin" weight $\delta(w)$ defined as the Skorohod integral of a function w:

$$\frac{\partial}{\partial x} u_n(x) = \mathbb{E}_x^Q [F_n \delta(w)]$$

Let's denote by g the function obtained as the expectation of the discounted payoff function f times the Malliavin weight $\delta(w)$: $g(x) \equiv \mathbb{E}_x^Q [F\delta(w)]$. By Cauchy-Schwartz inequality, we get:

$$\left| g(x) - \frac{\partial}{\partial x} u_n(x) \right| = \left| \mathbb{E}_x^Q [(F - F_n) \delta(w)] \right| \leq h(x) \epsilon_n(x) \qquad \text{(B.2)}$$

with

$$h(x) = \mathbb{E}_x^Q \left[(\delta(w))^2 \right]^{1/2} \quad \epsilon_n(x) = \mathbb{E}_x^Q \left[(F - F_n)^2 \right]^{1/2}$$

By definition, the L^2 convergence of u_n means $\epsilon_n(x)$ converges simply to zero as n tends to infinity. Therefore we already know that the function sequence $\left(\frac{\partial}{\partial x} u_n\right)_{n \in N}$ converges simply to the function g. By property of Lebesgue compactness and the fact that the functions F and F_n are continuous and that $h(x)$ is bounded (non-explosive condition (1.22)), inequality (B.2) proves that this convergence is uniform on any compact subsets K of $\mathbb{R}$..

We conclude using the fact that if a sequence of functions $(u_n)_{n \in \mathbb{N}}$ converges simply to a function u and the sequence of function's derivative $\left(\frac{\partial}{\partial x} u_n\right)_{n \in \mathbb{N}}$ converges uniformly to a function g on any compact subsets of $\mathbb{R}$, the limit function u is continuously differentiable with its derivative equal to the limit function of the sequence of function's derivative $\left(\frac{\partial}{\partial x} u_n\right)_{n \in \mathbb{N}}$ leading to the final result:

$$\frac{\partial}{\partial x} \mathbb{E}_x^Q [F] = \mathbb{E}_x^Q [F\delta(w)]$$

□

Second step: Interchanging the order of expectation and differentiation

The second technical point is to show that we can interchange the order of expectation and differentiation (using the dominated convergence theorem).

More precisely, since because of the first step, f is assumed to be element of C_K^∞ and therefore is continuously differentiable with bounded derivative, we have

$$\frac{F^{x+h}-F^x}{\|h\|}-\frac{\left\langle \frac{\partial}{\partial x}F,h\right\rangle}{\|h\|}\underset{\|h\|\to 0.}{\longrightarrow}0 \quad a.s.$$

An elementary calculation gives us

$$\frac{\partial}{\partial x}F=\left(\begin{array}{l}\sum_{i=1}^m e^{-\int_0^T r(s,X_s^x)ds}\partial_i f\left(X_{t_1}^x,X_{t_2}^x,...,X_{t_m}^x\right)\frac{\partial}{\partial x}X_{t_i}^x\\ -F\int_0^T r'\left(s,X_s^x\right)\frac{\partial}{\partial x}X_s^x ds\end{array}\right)$$

Since f has bounded derivative,first, $\frac{\left\langle \frac{\partial}{\partial x}F,h\right\rangle}{\|h\|}$ is uniformly integrable in h and second, by the Taylor-Lagrange theorem,

$$\left\|\frac{F^{x+h}-F^x}{\|h\|}\right\|\leq M\sum_{i=1}^m\frac{\left\|X_{t_i}^{x+h}-X_{t_i}^x\right\|}{\|h\|}$$

Using the result that $\sum_{i=1}^m\frac{\left\|X_{t_i}^{x+h}-X_{t_i}^x\right\|}{\|h\|}$ is uniformly integrable in h (See Theorem 2.4 pp 362 Chapter IX Stochastic Differential Equations, Revuz and Yor (1994)) leads to the uniform integrability in h of $\left\|\frac{F^{x+h}-F^x}{\|h\|}\right\|$

This, in turn, tells us that $\frac{F^{x+h}-F^x}{\|h\|}-\frac{\left\langle \frac{\partial}{\partial x}F,h\right\rangle}{\|h\|}$ is uniformly integrable in h. Since it converges a.s. to zero, the dominated convergence theorem gives us that it converges also to zero in L^1. We conclude that

$$\frac{\partial}{\partial x}u\left(X^x\right)=\mathbb{E}_x^Q\left[\frac{\partial}{\partial x}F\right] \tag{B.3}$$

□

Integration by parts:

Necessary condition: In this subsection, we examine the necessary condition to be satisfied by the weight function. The delta is defined as the derivative of the price function with respect to the initial condition x

$$delta=\frac{\partial}{\partial x}\mathbb{E}_x^Q\left[e^{-\int_0^T r(s,X_s^x)ds}f\left(X_{t_1}^x,X_{t_2}^x,...,X_{t_m}^x\right)\right] \tag{B.4}$$

Assuming the delta can be written with a weight function leads to

$$\begin{aligned} delta &= \mathbb{E}_x^Q\left[e^{-\int_0^T r(s,X_s^x)ds} f\left(X_{t_1}^x, X_{t_2}^x, ..., X_{t_m}^x\right)\delta(w)\right] \\ &= E_x^Q\left[\left\langle D_t\left(e^{-\int_0^T r(s,X_s^x)ds} f\left(X_{t_1}^x, X_{t_2}^x, ..., X_{t_m}^x\right)\right), w(t)\right\rangle\right] \end{aligned}$$

Using the property of Malliavin derivatives for compound functions, this can be written as:

$$\mathbb{E}_x^Q\left[\begin{array}{c} e^{-\int_0^T r(s,X_s^x)ds}\sum_{i=1}^m \partial_i f\left(X_{t_1}^x, X_{t_2}^x, ..., X_{t_m}^x\right)\int_{t=0}^T D_t X_{t_i}^x w(t)\,dt \\ -F\int_{t=0}^T\int_{s=0}^T \frac{\partial}{\partial X} r\left(s, X_s^x\right) D_t X_s w(t)\,dsdt \end{array}\right]$$

Using the relationship between the Malliavin derivative and the first-variation process (1.10), we can replace the expression of $D_t X_u$ $u \geq t$ in the equation above, leading to

$$\mathbb{E}_x^Q\left[\begin{array}{c} e^{-\int_0^T r(s,X_s^x)ds}\sum_{i=1}^m \partial_i f\left(X_{t_1}^x, X_{t_2}^x, ..., X_{t_m}^x\right) \\ \int_0^T Y_{t_i} Y_t^{-1}\sigma(t, X_t^x) w(t)\, 1_{\{t\leq t_i\}}dt \\ -F\int_{s=0}^T\int_{t=0}^T \frac{\partial}{\partial X} r\left(s, X_s^x\right) Y_s Y_t^{-1}\sigma(t, X_t^x) w(t)\, t 1_{\{t\leq s\}}dtds \end{array}\right]$$

On the other hand, the delta is defined as the derivative of the price function with respect to the initial condition x. Using (1.10) and the second step's results (B.3), we can change the LHS of (B.4) to

$$\begin{aligned} delta &= \mathbb{E}_x^Q\left[\begin{array}{c} e^{-\int_0^T r(s,X_s^x)ds}\sum_{i=1}^m \partial_i f\left(X_{t_1}^x, X_{t_2}^x, ..., X_{t_m}^x\right)\frac{\partial}{\partial x} X_{t_i} \\ -F\int_0^T r'\left(s, X_s^x\right)\frac{\partial}{\partial x} X_s ds \end{array}\right] \\ &= \mathbb{E}_x^Q\left[\begin{array}{c} e^{-\int_0^T r(s,X_s^x)ds}\sum_{i=1}^m \partial_i f\left(X_{t_1}^x, X_{t_2}^x, ..., X_{t_m}^x\right) Y_{t_i} \\ -F\int_0^T r'\left(s, X_s^x\right) Y_s ds \end{array}\right] \end{aligned}$$

At this stage, equalling the two expressions of delta gives us:

$$\begin{aligned} &\mathbb{E}_x^Q\left[\begin{array}{c} e^{-\int_0^T r(s,X_s^x)ds}\sum_{i=1}^m \partial_i f\left(X_{t_1}^x, X_{t_2}^x, ..., X_{t_m}^x\right) \\ \int_0^T Y_{t_i} Y_t^{-1}\sigma(t, X_t) w(t)\, 1_{\{t\leq t_i\}}dt \\ -F\int_{s=0}^T\int_{t=0}^T r'\left(s, X_s^x\right) Y_s Y_t^{-1}\sigma(t, X_t) w(t)\, t 1_{\{t\leq s\}}dtds \end{array}\right] \\ = &\mathbb{E}_x^Q\left[\begin{array}{c} e^{-\int_0^T r(s,X_s^x)ds}\sum_{i=1}^m \partial_i f\left(X_{t_1}^x, X_{t_2}^x, ..., X_{t_m}^x\right) Y_{t_i} \\ -F\int_0^T r'\left(s, X_s^x\right) Y_s\, ds \end{array}\right] \end{aligned}$$

Using the fact that this should hold for any f and any function $r(.,.)$, we get that the following two quantities should be equal on any functions measurable, leading to conditions expressed with conditional expectations (where, to simplify notations, the x in superscript has been omitted):

$$\begin{aligned}
&\mathbb{E}_x^Q\left[e^{-\int_0^T r(s,X_s^x)ds}\int_0^{t_i}\frac{Y_{t_i}\sigma(t,X_t)}{Y_t}w(t)\,dt|X_{t_1},...,X_{t_m}\right] \\
=\ &\mathbb{E}_x^Q\left[e^{-\int_0^T r(s,X_s^x)ds}Y_{t_i}|X_{t_1},...,X_{t_m}\right] \qquad \forall i=1...m
\end{aligned} \tag{B.5}$$

$$\begin{aligned}
&\mathbb{E}_x^Q\left[\int_{s=0}^T\int_{t=0}^T r'(s,X_s)\,Y_sY_t^{-1}\sigma(t,X_t)w(t)\,t1_{\{t\leq s\}}dtds|X_{t_1},...,X_{t_m}\right] \\
=\ &\mathbb{E}_x^Q\left[\int_0^T r'(s,X_s)\,Y_sds|X_{t_1},...,X_{t_m}\right]
\end{aligned} \tag{B.6}$$

This is exactly (1.M1) when the interest rate is a only function of the time □

Sufficient condition: Let's assume that there exists a function w that verifies the two equations (B.5) and (B.6) and its Skorohod integral is L_2 integrable, the above proof can be conducted backwards:

$$\begin{aligned}
delta &= \frac{\partial}{\partial x}\mathbb{E}_x^Q\left[\left(e^{-\int_0^T r(s,X_s)ds}f\left(X_{t_1},X_{t_2},...,X_{t_m}\right)\right)\right] \\
&= \mathbb{E}_x^Q\left[\begin{array}{l}\sum_{i=1}^m e^{-\int_0^T r(s,X_s)ds}\partial_i f\left(X_{t_1},X_{t_2},...,X_{t_m}\right)\frac{\partial}{\partial x}X_{t_i} \\ -\left(e^{-\int_0^T r(s,X_s)ds}f\left(X_{t_1},X_{t_2},...,X_{t_m}\right)\right)\int_0^T r'(s,X_s)\,\frac{\partial}{\partial x}X_sds\end{array}\right]
\end{aligned}$$

then using the conditions (B.5) and (B.6), we get

$$\begin{aligned}
&= \mathbb{E}_x^Q\left[\begin{array}{c}\sum_{i=1}^m e^{-\int_0^T r(s,X_s)ds}\partial_i f\left(X_{t_1},X_{t_2},...,X_{t_m}\right) \\ \int_0^T Y_{t_i}Y_t^{-1}\sigma(t,X_t)w(t)\,1_{\{t\leq t_i\}}dt \\ -F\int_{s=0}^T\int_{t=0}^T r'(s,X_s)\,Y_sY_t^{-1}\sigma(t,X_t)\,w(t)\,t1_{\{t\leq s\}}dtds\end{array}\right] \\
&= \mathbb{E}_x^Q\left[\begin{array}{c}e^{-\int_0^T r(s,X_s)ds}\sum_{i=1}^m \nabla_i f\left(X_{t_1},X_{t_2},...,X_{t_m}\right)\int_{t=0}^T D_tX_{t_i}w(t)\,dt \\ -F\int_{t=0}^T\int_{s=0}^T r'(s,X_s)\,D_tX_sw(t)\,dsdt\end{array}\right]
\end{aligned}$$

and finally the expression of the Malliavin derivative in terms of the first variation process, leads to

$$\mathbb{E}_x^Q\left[\left\langle D_t\left(e^{-\int_0^T r(s,X_s)ds}f\left(X_{t_1},X_{t_2},...,X_{t_m}\right)\right),w(t)\right\rangle\right]$$

that is:

$$delta = \mathbb{E}_x^Q \left[F\delta\left(w\right)\right]$$

where in the last step, we made use of the integration by parts formula. □

B.3 Proof of the gamma formula (1.M2)

Necessary and sufficient condition

The proof goes along the same lines as for the delta case, so we omit to give all details of it. We assume that f is continuously twice differentiable with bounded first and second order derivatives. To remind that the generator w^{delta} does depend on x, we adopt an explicit notation w_x^{delta}.

$$\begin{aligned}
\Gamma &= \Delta_x \mathbb{E}_x^Q [F] \\
&= \frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}\mathbb{E}_x^Q [F]\right) \\
&= \frac{\partial}{\partial x}\left(\mathbb{E}_x^Q \left[F\delta\left(w_x^{delta}\right)\right]\right) \\
&= \mathbb{E}_x^Q\left[\frac{\partial}{\partial x}F\delta\left(w_x^{delta}\right)\right] + \mathbb{E}_x^Q\left[F\frac{\partial}{\partial x}\delta\left(w_x^{delta}\right)\right]
\end{aligned}$$

Using the fact that one could invert the Skorohod integral operator $\delta(.)$ and the differential operator $\frac{\partial}{\partial x}$ (thanks to a mathematical argument based on dominated convergence theorem), we get

$$\begin{aligned}
\Gamma &= \mathbb{E}_x^Q\left[F\left(\delta\left(w_x^{delta}\right)\delta\left(w_x^{delta}\right) + \delta\left(\frac{\partial}{\partial x}w_x^{delta}\right)\right)\right] \\
&= \mathbb{E}_x^Q\left[F\delta\left(w^{delta}\delta\left(w_x^{delta}\right) + \frac{\partial}{\partial x}w_x^{delta}\right)\right]
\end{aligned}$$

where in the last inequality we used the linearity of the Skorohod integral operator. Since this should hold for any F, the necessary and sufficient condition is

$$\mathbb{E}^Q_{x,X_{t_1},\ldots,X_{t_m}}\left[\delta\left(w^{gamma}\right)\right] = \mathbb{E}^Q_{x,X_{t_1},\ldots,X_{t_m}}\left[\delta\left(w^{delta}\delta\left(w^{delta}\right) + \frac{\partial}{\partial x}w^{delta}\right)\right]$$

□

Particular solution

In this section, we prove the formula for the particular form of weight, we have already encountered, namely:

$$w_t^{delta} = Y_t \sigma^{-1}(t, X_t)\lambda_t$$

with

$$\int_0^T \lambda_t 1_{\{t \leq t_i\}} dt = 1 \qquad \forall i = 1...m$$

Using the result on the Gamma weight function, a sufficient condition on the Malliavin weight is the equality:

$$\delta\left(w^{gamma}\right) = \delta\left(\frac{\partial}{\partial x} w^{delta}\right) + \delta\left(w^{delta}\right)\delta\left(w^{delta}\right)$$

with

$$\delta\left(w^{delta}\right)\delta\left(w^{delta}\right) = \left(\int_{t=0}^T \lambda_t \sigma^{-1}(t, X_t) Y_t dW_t\right)\left(\int_{u=0}^T \lambda_u \sigma^{-1}\left(X_u\right) Y_u dW_u\right)$$

which can be expressed in terms of the square of the simple integral:

$$\delta\left(w^{delta}\right)\delta\left(w^{delta}\right) = \left(\int_0^T \lambda_t \sigma^{-1}(t, X_t) Y_t dW_t\right)^2 - \int_0^T \mathbb{E}\left[\left(\lambda_t \sigma^{-1}(t, X_t) Y_t\right)^2\right] ds$$

The term $\frac{\partial}{\partial x} w^{delta}$ can be calculated as the sum of two terms:

$$\frac{\partial}{\partial x} w^{delta} = \lambda_t \left(\partial_2\left(\sigma^{-1}\right)(t, X_t) Y_t + \sigma^{-1}(t, X_t)\frac{\partial}{\partial x} Y_t\right)$$

We then use the following equation:

$$\partial_2\left(\sigma^{-1}\right)(t, X_t) = -\sigma^{-2}(t, X_t)\sigma'\left(t, X_t\right) Y_t$$

and we use for the second term that

$$\begin{aligned}\frac{\partial}{\partial x} Y_t &= \int_0^t Y_t Y_s b''\left(s, X_s\right) ds + \int_0^t Y_t Y_s \sigma''\left(s_1, X_{s_1}\right) dW_s \\ &\quad - \int_0^t Y_t \sigma'\left(s, X_s\right) Y_s \sigma''\left(s, X_s\right) ds\end{aligned}$$

giving:

$$\begin{aligned}\partial_x\left(w^{delta}\right) &= \int_{s_2 0}^T \int_{s=0}^{s_2} \lambda_{s_2} \sigma^{-1}(s_2, X_{s_2}) Y_{s_2} Y_{s_1}\left(b''\left(s_1, X_{s_1}\right) - \sigma'\left(s_1, X_{s_1}\right)\sigma''\left(s_1, X_{s_1}\right)\right) ds_1 dW_{s_2} \\ &\quad + \int_{s_2=0}^T \lambda_{s_2}\sigma^{-1}(s_2, X_{s_2}) Y_{s_2}\int_{s_1=0}^{s_2} Y_{s_1}\sigma''\left(s_1, X_{s_1}\right) dW_{s_1} dW_{s_2}\end{aligned}$$

We then conclude that the Malliavin weight is given by (1.25).□

B.4 Proof of the rho formulae (1.M3)

Since the rho is a perturbation on the drift term, we can prove the formula by means of a measure change (Girsanov theorem). We can also use the technique developed in the delta proof.

Standard Proof

The following proof is based on Malliavin calculus. As explained in the chapter 1 section on extended Greeks 1.2.3, page 33, we are perturbing our process along a perturbation direction given by the function $\widetilde{b}(.,.)$. We denote by $\left\{X_t^{\varepsilon,rho}, t \in [0,T]\right\}$ the perturbed underlying process following equation (1.12) and the unmodified initial condition ($X_0^{\varepsilon,rho} = x$). We denote by $\left(Z_t^{rho}\right)_{t\in[0,T]}$ the Gateau derivative of the drift-perturbed underlying process $\left\{X_t^{\varepsilon,rho}, t \in [0,T]\right\}$ along the direction $\widetilde{b}$, defined as the limit in L^2, uniformly with respect to the time t and given by equation (1.16). To find a necessary condition for the weight function, we apply the same methodology as the one described for the computation of the delta. We assume therefore that we can write rho, defined as in (1.14) as the expectation of the discounted payoff function F times a suitable weighting function $weight^{rho}$

$$\begin{aligned} rho &= \mathbb{E}_x^Q\left[F\delta\left(w^{rho}\right)\right] \\ &= \left.\frac{\partial}{\partial\varepsilon}P_{rho}^{\varepsilon}(x)\right|_{\varepsilon=0,\widetilde{b}\ given} \end{aligned} \tag{B.7}$$

Transforming equation (B.7) leads to

$$\begin{aligned} rho &= \mathbb{E}_x^Q\left[F\delta\left(w^{rho}\right)\right] \\ &= \mathbb{E}_x^Q\left[\left\langle D_tF, w^{rho}\right\rangle\right] \end{aligned} \tag{B.8}$$

$$= \mathbb{E}_x^Q\left[\int_{t=0}^{T}\left(\begin{array}{c} \sum_{i=1}^m e^{-\int_0^T r(s,X_s)ds}\nabla_i f\left(X_{t_1}, X_{t_2}, ..., X_{t_m}\right) D_tX_{t_i} \\ -F\int_{s=0}^T \frac{\partial}{\partial X} r\left(s, X_s\right) D_tX_s\, ds \end{array}\right) w^{rho}(t)\, dt\right] \tag{B.9}$$

Using relationship (1.10), we write:

$$= \mathbb{E}_x^Q \left[\int_{t=0}^{T} \left(\begin{array}{c} \sum_{i=1}^{m} e^{-\int_0^T r(s,X_s)ds} \nabla_i f(X_{t_1}, X_{t_2}, ..., X_{t_m}) \\ Y_{t_i} Y_t^{-1} \sigma(t, X_t) 1_{\{t \leq t_i\}} \\ -F \int_{s=0}^{T} \frac{\partial}{\partial X} r(s, X_s) Y_s Y_t^{-1} \sigma(t, X_t) 1_{\{t \leq s\}} \, ds \end{array} \right) w^{rho}(t) \, dt \right] \tag{B.10}$$

$$= \mathbb{E}_x^Q \left[\left(\begin{array}{c} \sum_{i=1}^{m} e^{-\int_0^T r(s,X_s)ds} \nabla_i f(X_{t_1}, X_{t_2}, ..., X_{t_m}) \\ \int_{t=0}^{T} Y_{t_i} Y_t^{-1} \sigma(t, X_t) 1_{\{t \leq t_i\}} w^{rho}(t) \, dt \\ -F \int_{s=0}^{T} \frac{\partial}{\partial X} r(s, X_s) \left(\int_{t=0}^{T} Y_s Y_t^{-1} \sigma(t, X_t) 1_{\{t \leq s\}} \, w^{rho}(t) \, dt \right) ds \end{array} \right) \right] \tag{B.11}$$

On the other hand, equation (1.14) leads to

$$rho = \mathbb{E}_x^Q \left[\begin{array}{c} \sum_{i=1}^{m} e^{-\int_0^T r(s,X_s)ds} \nabla_i f(X_{t_1}, X_{t_2}, ..., X_{t_n}) Z_t^{rho} \\ - \left(e^{-\int_0^T r(s,X_s)ds} f(X_{t_1}, X_{t_2}, ..., X_{t_n}) \right) \int_{s=0}^{T} \frac{\partial}{\partial X} r(s, X_s) Z_s^{rho} ds \end{array} \right] \tag{B.12}$$

using the proposition 1 page 36 with the equation (1.18), we get

$$= \mathbb{E}_x^Q \left[\begin{array}{c} \sum_{i=1}^{m} e^{-\int_0^T r(s,X_s)ds} \nabla_i f(X_{t_1}, X_{t_2}, ..., X_{t_n}) \\ \int_{t=0}^{T} Y_{t_i} Y_t^{-1} \widetilde{b}(t, X_t) 1_{\{t < t_i\}} dt \\ -F \int_{s=0}^{T} \frac{\partial}{\partial X} r(s, X_s) \int_{t=0}^{T} Y_s Y_t^{-1} \widetilde{b}(t, X_t) 1_{\{t < s\}} dt ds \end{array} \right] \tag{B.13}$$

This should be verified for any f, any process $(X_t)_{t \in [0,T]}$, any process $r(.,.)$. Thus, we find that the necessary conditions are

$$\begin{aligned} & \mathbb{E}_{x, X_{t_1}, ..., X_{t_m}}^Q \left[Y_{t_i} \int_0^{t_i} \frac{\sigma(t, X_t)}{Y_t} w^{rho}(t) \, dt \right] \\ = \; & \mathbb{E}_{x, X_{t_1}, ..., X_{t_m}}^Q \left[Y_{t_i} \int_0^{t_i} \frac{\widetilde{b}(t, X_t)}{Y_t} dt \right] \end{aligned} \tag{B.14}$$

equation (M3) and the conditions for stochastic interest rate models

$$\begin{aligned} & \mathbb{E}_{x, X_{t_1}, ..., X_{t_m}}^Q \left[\int_{0=t \leq s}^{T} r'(s, X_s) \frac{Y_s \sigma(t, X_t)}{Y_t} w^{rho}(t) \, dt \, ds \right] \\ = \; & = \mathbb{E}_{x, X_{t_1}, ..., X_{t_m}}^Q \left[\int_{0=t \leq s}^{T} r'(s, X_s) \frac{Y_s \widetilde{b}(t, X_t)}{Y_t} dt ds \right] \end{aligned} \tag{B.15}$$

In fact, we can go backwards as well. Rho is defined by equation (B.12), which is equivalent to (B.13). Assuming (B.14) and (B.15), (B.13) is equivalent to (B.11), (B.10), (B.9), (B.8) leading to the result.□

Probabilistic proof

This proof requires stronger assumptions. In this part we assume that the Novikov condition (B.16) is satisfied so as to be able to use Girsanov's theorem.

$$\mathbb{E}\left[e^{\frac{\varepsilon^2}{2}\int_0^T \left\|\sigma^{-1}\left(t,X_t^{\varepsilon,rho}\right)\widetilde{b}\left(t,X_t^{\varepsilon,rho}\right)\right\|^2 dt}\right] < +\infty \tag{B.16}$$

The idea of this probabilistic proof is the following: in the special case of rho, the perturbation of the diffusion equation is only on the deterministic term and not on the stochastic term. Therefore, a judicious change of measure can remove the drift term. In this new probability measure, the process follows the same diffusion equation. The calculation of the Malliavin formula turns out to be a change of measure. This proof is inspired by Fournié et al. (1999).

The proof is based on the following lemma:

Lemma 2 *Change of measure*
Introducing the new probability measure Q^ε defined by its Radon-Nikodym derivative with respect to the risk-neutral probability measure Q:

$$\left.\frac{dQ^\varepsilon}{dQ}\right|_T = e^{\varepsilon\int_0^T \sigma^{-1}\left(t,X_t^{\varepsilon,rho}\right)\widetilde{b}\left(t,X_t^{\varepsilon,rho}\right)dW_t - \frac{\varepsilon^2}{2}\int_0^T \left\|\sigma^{-1}\left(t,X_t^{\varepsilon,rho}\right)\widetilde{b}\left(t,X_t^{\varepsilon,rho}\right)\right\|^2 dt} \tag{B.17}$$

we get that $(X_t^\varepsilon)_{t\in[0,T]}$ follows the same SDE under Q^ε as $(X_t)_{t\in[0,T]}$ under Q

Proof: Justified by Novikov condition (B.16), the Girsanov theorem enables to rewrite the perturbed equation (1.12) into a regular one:

$$dX_t^{\varepsilon,rho} = b\left(t,X_t^{\varepsilon,rho}\right)dt + \sigma\left(t,X_t^{\varepsilon,rho}\right)d\widetilde{W_t} \tag{B.18}$$

with $\widetilde{W_t}$ a $Q^\varepsilon-$Brownian motion given by:

$$d\widetilde{W_t} = dW_t + \varepsilon\sigma^{-1}\left(t,X_t^{\varepsilon,rho}\right)\widetilde{b}\left(t,X_t^{\varepsilon,rho}\right)dt$$

□

Lemma 3 *The Radon-Nikodym derivative* $\frac{dQ}{dQ^\varepsilon}$ *is differentiable in* $\varepsilon = 0$ *and:*

$$\lim_{\varepsilon \to 0} \frac{\frac{dQ}{dQ^\varepsilon} - 1}{\varepsilon} = -\int_0^T \sigma^{-1}(t, X_t)\,\widetilde{b}(t, X_t)\,dW_t \text{ in } L^2 \tag{B.19}$$

Proof: Writting

$$M_{\varepsilon,s} \equiv e^{-\varepsilon \int_0^s \sigma^{-1}\left(t, X_t^{\varepsilon,rho}\right)\widetilde{b}\left(t, X_t^{\varepsilon,rho}\right)dW_t - \frac{\varepsilon^2}{2}\int_0^s \left\|\sigma^{-1}\left(t, X_t^{\varepsilon,rho}\right)\widetilde{b}\left(t, X_t^{\varepsilon,rho}\right)\right\|^2 dt}$$

We have that $M_{\varepsilon,s}$ is the solution of the SDE (by the use of Ito lemma):

$$dM_{\varepsilon,t} = -\varepsilon\sigma^{-1}\left(t, X_t^{\varepsilon,rho}\right)\widetilde{b}\left(t, X_t^{\varepsilon,rho}\right)M_{\varepsilon,t}dW_t \tag{B.20}$$

with initial condition

$$M_{\varepsilon,t=0} = 1 \tag{B.21}$$

(B.17) gives us

$$\left.\frac{dQ}{dQ_\varepsilon}\right|_T = M_{\varepsilon,s}$$

Using dominated convergence, since $b(.,.)$, $\sigma(.,.)$, $\widetilde{b}(.,.)$, and $b + \varepsilon\widetilde{b}(.,.)$ are supposed to satisfy linear growth conditions, we obtain:

$$\frac{\partial}{\partial\varepsilon}M_{\varepsilon,t} = -\sigma^{-1}\left(t, X_t^{\varepsilon,rho}\right)\widetilde{b}\left(t, X_t^{\varepsilon,rho}\right)M_{\varepsilon,t}$$

Since $b(.,.)$, $\sigma(.,.)$, $\widetilde{b}(.,.)$, and $b + \varepsilon\widetilde{b}(.,.)$ for every $\varepsilon \in [0,1]$ are supposed to satisfy the global Lipschitz and linear growth conditions, by Theorem 2.4 pp 362 Chapter IX Stochastic Differential Equations of Revuz and Yor (1994) or by Theorem 2.9 pp 289 Chapter V 5.2 Strong Solution of Karatzas and Shreeve (1988), the processes $\left\{X_t^{\varepsilon,rho}, \varepsilon \in [0,1], t \in [0,T]\right\}$ (respectively $\{M_{\varepsilon,t}, \varepsilon \in [0,1], t \in [0,T]\}$) defined as the continuous strong solution of the SDE associated (1.12) (respectively (B.20)) with unmodified initial condition ($X_0^{\varepsilon,rho} = x$) (respectively (B.21)) belong to $L^2([0,1] \times [0,T] \times \Omega)$ and converge in the sense of the L^2 norm to the no perturbed process $\{X_t, t \in [0,T]\}$ respectively $\{1, t \in [0,T]\}$ as ε tends to zero. Using continuity of $\sigma^{-1}(.,.)$, $\widetilde{b}(.,.)$, we get

$$\lim_{\varepsilon \to 0}\frac{\partial}{\partial\varepsilon}M_{\varepsilon,t} = -\sigma^{-1}(t, X_t)\,\widetilde{b}(t, X_t)$$

Using the continuity and differentiability of the exponential, we get that

$$\frac{1}{|\varepsilon|}\left[M_{\varepsilon,T}-1+\varepsilon\sigma^{-1}(t,X_t)\widetilde{b}(t,X_t)\right]$$

is uniformly integrable. By dominated convergence, we conclude that the limit of $\frac{\frac{dQ}{dQ^\varepsilon}-1}{\varepsilon}$ (in the sense of L^2 norm) is given by:

$$\begin{aligned}\lim_{\varepsilon\to0,\varepsilon\neq0}\frac{\frac{dQ}{dQ^\varepsilon}-1}{\varepsilon} &= \lim_{\varepsilon\to0,\varepsilon\neq0}\frac{M_{\varepsilon,T}-1}{\varepsilon}\\ &= -\sigma^{-1}(t,X_t)\widetilde{b}(t,X_t)\end{aligned}$$

□

Final Proof:

Let us denote by F^ε_{rho} the perturbed discounted payoff function.

$$F^\varepsilon_{rho}=e^{-\int_0^T r\left(s,X_s^{\varepsilon,rho}\right)ds}f\left(X_{t_1}^{\varepsilon,rho},X_{t_2}^{\varepsilon,rho},\ldots,X_{t_n}^{\varepsilon,rho}\right)$$

Therefore

$$\begin{aligned}rho &\equiv \left.\frac{\partial}{\partial\varepsilon}P^\varepsilon_{rho}(x)\right|_{\varepsilon=0}\\ &= \lim_{\varepsilon\to0,\varepsilon\neq0}\frac{\mathbb{E}_x^Q\left[F^\varepsilon_{rho}\right]-\mathbb{E}_x^Q\left[F\right]}{\varepsilon}\end{aligned}$$

But on the first term of the numerator, we can change our probability measure:

$$\mathbb{E}_x^Q\left[F^\varepsilon_{rho}\right]=\mathbb{E}_x^{Q^\varepsilon}\left[F^\varepsilon_{rho}\frac{dQ}{dQ^\varepsilon}\right]$$

by lemma (2), this is equivalent to

$$\mathbb{E}_x^{Q^\varepsilon}\left[F^\varepsilon_{rho}\frac{dQ}{dQ^\varepsilon}\right]=\mathbb{E}_x^Q\left[F\frac{dQ}{dQ^\varepsilon}\right]$$

leading to

$$\frac{P^\varepsilon_{rho}(x)-P(x)}{\varepsilon}=\mathbb{E}_x^Q\left[F\frac{\frac{dQ}{dQ^\varepsilon}-1}{\varepsilon}\right]$$

therefore

$$\begin{aligned}&\left|\frac{P^\varepsilon_{rho}(x)-P(x)}{\varepsilon}-\mathbb{E}_x^Q\left[F\int_0^T\sigma^{-1}(t,X_t)\widetilde{b}(t,X_t),dW_t\right]\right|\\ \leq&\left\|F\left(\frac{\frac{dQ^\varepsilon}{dQ}-1}{\varepsilon}-\int_0^T\sigma^{-1}(t,X_t)\widetilde{b}(t,X_t),dW_t\right)\right\|_{L^1}\\ \leq&\ \|F\|_{L^2}\left\|\frac{\frac{dQ^\varepsilon}{dQ}-1}{\varepsilon}-\int_0^T\sigma^{-1}(t,X_t)\widetilde{b}(t,X_t),dW_t\right\|_{L^2}\end{aligned}$$

The last equation is justified by Cauchy Schwartz inequality. We trivially conclude using lemma (3) result (B.19)

$$\lim_{\varepsilon\to 0,\varepsilon\neq 0}\frac{P_t^{\varepsilon}(x)-P_t(x)}{\varepsilon} = \mathbb{E}_x^Q\left[F\int_0^T \sigma^{-1}(t,X_t)\widetilde{b}(t,X_t),dW_t\right] \text{ in } L^2$$

□

B.5 Proof of the vega formula (1.M4)

The method is very similar to the one used in the rho calculation except that we cannot use a measure change to find the Malliavin weight. As explained in the chapter 1 section on extended Greeks 1.2.3, page 33, we are assuming that we are perturbing our process along the direction given by the function $\widetilde{\sigma}(.,.)$. We also assume hypoellipticity of the perturbed diffusion term (1.11). We denote by $\{X_t^{\varepsilon,vega}, t\in[0,T]\}$ the perturbed underlying process following equation (1.13) with the unmodified initial condition ($X_0^{\varepsilon,vega}=x$). We denote by $(Z_t^{vega})_{t\in[0,T]}$ the Gateau derivative of the volatility-perturbed underlying process $\{X_t^{\varepsilon,vega}, t\in[0,T]\}$ along the direction $\widetilde{\sigma}$ and defined as the limit in L^2, uniformly with respect to the time t as summarized by equation (1.17). To find a necessary condition for the weight function, we apply the same methodology as the one described for the computation of the delta or rho. We assume therefore that we can write vega, defined as in (1.15) as the expectation of the discounted payoff function F times a suitable weight function $weight^{vega}$

$$\begin{aligned} vega &= \mathbb{E}_x^Q\left[F\delta\left(w^{vega}\right)\right] \\ &= \left.\frac{\partial}{\partial\varepsilon}P^{\varepsilon}_{vega}(x)\right|_{\varepsilon=0,\widetilde{b}\ given} \end{aligned} \tag{B.22}$$

Transforming equation (B.22)

$$vega = \mathbb{E}_x^Q\left[\int_{t=0}^T\left(\begin{array}{c}\sum_{i=1}^m e^{-\int_0^T r(s,X_s)ds}\nabla_i f\left(X_{t_1},X_{t_2},...,X_{t_m}\right)D_tX_{t_i}\\ -F\int_{s=0}^T\nabla_2 r\left(s,X_s\right)D_tX_s\,ds\end{array}\right)w^{vega}(t)\,dt\right] \tag{B.23}$$

Which using (1.10), is written as:

$$\mathbb{E}_x^Q\left[\int_{t=0}^{T}\left(\begin{array}{c}\sum_{i=1}^{m} e^{-\int_0^T r(s,X_s)ds}\nabla_i f\left(X_{t_1},X_{t_2},...,X_{t_m}\right)\\ Y_{t_i}Y_t^{-1}\sigma(t,X_t)1_{\{t\leq t_i\}}\\ -F\int_{s=0}^{T}\nabla_2 r\left(s,X_s\right)Y_sY_t^{-1}\sigma(t,X_t)\,1_{\{t\leq s\}}\;ds\end{array}\right)w^{vega}(t)\;dt\right] \tag{B.24}$$

$$=\mathbb{E}_x^Q\left[\left(\begin{array}{c}\sum_{i=1}^{m} e^{-\int_0^T r(s,X_s)ds}\nabla_i f\left(X_{t_1},X_{t_2},...,X_{t_m}\right)\\ \int_{t=0}^{T}Y_{t_i}Y_t^{-1}\sigma(t,X_t)1_{\{t\leq t_i\}}w^{vega}(t)\;dt\\ -F\int_{s=0}^{T}\nabla_2 r\left(s,X_s\right)\left(\int_{t=0}^{T}Y_sY_t^{-1}\sigma(t,X_t)\,1_{\{t\leq s\}}\;w^{vega}(t)\;dt\right)\;ds\end{array}\right)\right] \tag{B.25}$$

On the other hand equation (1.15), we have

$$vega \;=\; \mathbb{E}_x^Q\left[\begin{array}{c}\sum_{i=1}^{m} e^{-\int_0^T r(s,X_s)ds}\nabla_i f\left(X_{t_1},X_{t_2},...,X_{t_n}\right)Z_t^{vega}\\ -\left(e^{-\int_0^T r(s,X_s)ds}f\left(X_{t_1},X_{t_2},...,X_{t_n}\right)\right)\int_{s=0}^{T}\nabla_2 r\left(s,X_s\right)Z_s^{vega}ds\end{array}\right] \tag{B.26}$$

and using the lemma (1) equation (1.19), we get

$$=\mathbb{E}_x^Q\left[\begin{array}{c}\sum_{i=1}^{m} e^{-\int_0^T r(s,X_s)ds}\nabla_i f\left(X_{t_1},X_{t_2},...,X_{t_n}\right)\\ \int_{t=0}^{T}Y_{t_i}Y_t^{-1}\widetilde{\sigma}\left(t,X_t\right)1_{\{t<t_i\}}dW_t-\int_0^T Y_{t_i}Y_t^{-1}\nabla_2\sigma\left(s,X_s\right)\widetilde{\sigma}\left(s,X_s\right)1_{\{t<t_i\}}ds\\ -F\int_{s=0}^{T}\nabla_2 r\left(s,X_s\right)\left(\begin{array}{c}\int_{t=0}^{T}Y_sY_t^{-1}\widetilde{\sigma}\left(t,X_t\right)1_{\{t<s\}}dW_t\\ -\int_0^T Y_sY_t^{-1}\nabla_2\sigma\left(t,X_t\right)\widetilde{\sigma}\left(t,X_t\right)1_{\{t<s\}}dt\end{array}\right)ds\end{array}\right] \tag{B.27}$$

This should be verified for any f, any process $(X_t)_{t\in[0,T]}$, any process $r(.,.)$. Thus, the following two necessary should hold

$$\mathbb{E}_{x,X_{t_1},...,X_{t_m}}^Q\left[Y_{t_i}\int_0^{t_i}\frac{\sigma(t,X_t)}{Y_t}w^{vega}(t)\,dt\right] \tag{B.28}$$

$$=\;\mathbb{E}_{x,X_{t_1},...,X_{t_m}}^Q\left[\begin{array}{c}\int_0^{t_i}\frac{\widetilde{\sigma}(t,X_t)Y_{t_i}}{Y_t}dW_t\\ -\int_0^{t_i}\sigma'\left(s,X_s\right)\frac{\widetilde{\sigma}(s,X_s)Y_{t_i}}{Y_t}ds\end{array}\right]$$

Greeks	Weighting Function
delta	$\int_0^T a(t) \frac{Y_t}{\sigma(t,X_t)} dW_t$
"extended" rho	$\int_0^T \frac{1}{\sigma(t,X_t)} \widetilde{b}(t,X_t) dW_t$
"extended" vega	$\delta\left(\frac{Y_t}{\sigma(t,X_t)} \widetilde{a}(t) \sum_{i=1}^m \left(\frac{Z_{t_i}^{vega}}{Y_{t_i}} - \frac{Z_{t_{i-1}}^{vega}}{Y_{t_{i-1}}}\right) 1_{\{t_{i-1} \leq t < t_i\}}\right)$

Table B.1: Summary of Fournié et al. Results

$$\begin{aligned} & \mathbb{E}^{Q}_{x,X_{t_1},\ldots,X_{t_m}}\left[\int_{0=t\leq s}^{T} r'(s,X_s)\frac{Y_s\sigma(t,X_t)}{Y_t}w^{vega}(t)\,dtds\right] \qquad \text{(B.29)} \\ = \; & \mathbb{E}^{Q}_{x,X_{t_1},\ldots,X_{t_m}}\left[\begin{array}{c} \int_{0=t\leq s}^{T} r'(s,X_s) \\ \left(\begin{array}{c} \frac{Y_s}{Y_t}\widetilde{\sigma}(t,X_t)\,dW_t \\ -\frac{Y_s}{Y_t}\sigma'(t,X_t)\widetilde{\sigma}(t,X_t)\,ds \end{array}\right) ds \end{array}\right] \end{aligned}$$

In fact, we can go backwards as well. Rho is defined by equation (1.15), which is equivalent to (B.26) and (B.27). Assuming (B.28) and (B.29), (B.27) is equivalent to (B.25), (B.24), (B.23) leads to the result (B.22).□

B.6 Summary of Fournié et al. particular solutions

Founie et al. proved that the weight function could be written in the case of adapted processes as some Ito integral. Let us define

$T_m = \left\{a \in L^2[0,T] \mid \int_0^{t_i} a(t)\,dt = 1 \;\forall i = 1...m\right\}$

and $\widetilde{T}_m = \left\{\widetilde{a} \in L^2[0,T] \mid \int_{t_{i-1}}^{t_i} \widetilde{a}(t)\,dt = 1 \;\forall i = 1...m\right\}$. Their results are summarized in the table 1.1 , where the symbol δ stands for the Skorohod integral and a is an element of T_m, $\widetilde{a}$ an element of $\widetilde{T}_m$.

Greeks	Supplementary conditions
delta	$\mathbb{E}^{Q}_{x,X_{t_1},\dots,X_{t_m}}\left[\int_{0=t\leq s}^{T} r'(s,X_s)\frac{Y_s\sigma(t,X_t)}{Y_t}w^{delta}(t)\,dtds\right] = \mathbb{E}^{Q}_{x,X_{t_1},\dots,X_{t_m}}\left[\int_{0=s}^{T} r'(s,X_s)\,Y_s ds\right]$
gamma	the extension to this case is included in the equality of the delta
"extended" rho	$\mathbb{E}^{Q}_{x,X_{t_1},\dots,X_{t_m}}\left[\int_{0=t\leq s}^{T} r'(s,X_s)\frac{Y_s\sigma(t,X_t)}{Y_t}w^{rho}(t)\,dt\,ds\right] = \mathbb{E}^{Q}_{x,X_{t_1},\dots,X_{t_m}}\left[\int_{0=t\leq s}^{T} r'(s,X_s)\frac{Y_s\widetilde{b}(t,X_t)}{Y_t}dtds\right]$
"extended" vega	$\mathbb{E}^{Q}_{x,X_{t_1},\dots,X_{t_m}}\left[\int_{0=t\leq s}^{T} r'(s,X_s)\frac{Y_s\sigma(t,X_t)}{Y_t}w^{vega}(t)\,dtds\right] = \mathbb{E}^{Q}_{x,X_{t_1},\dots,X_{t_m}}\left[\begin{array}{c}\int_{0=t\leq s}^{T} r'(s,X_s)\\ \left(\begin{array}{c}\frac{Y_s}{Y_t}\widetilde{\sigma}(t,X_t)\,dW_t\\ -\frac{Y_s}{Y_t}\sigma'(t,X_t)\,\widetilde{\sigma}(t,X_t)\,ds\end{array}\right)ds\end{array}\right]$

Table B.2: Supplementary conditions for models with risk-free rate depending on the underlying

B.7 Second conditions for stochastic interest models

Appendix C

Technical Proofs of chapter 4

The proof is given for the delta formula (4.M1). However, similar methods lead to the one for the gamma, formula (4.M2), the rho, (4.M3) and the vega, (4.M4).

For the sake of clarity of the proofs given below, we take a discount factor equal to 1. The risk-free interest rate is deterministic. The discount factor is therefore a multiplicative constant. Consequently, it does not change the results.

C.1 Simple Asian option delta: necessary and sufficient conditions (Formula (4.M1))

The proof goes along the same lines as the one given in Benhamou (2000a). First of all, using the density of the set of infinitely differentiable functions with compact support C_K^∞ into the set of square integrable functions, $L^2[0,T]$ as well as the continuity of the expectation operator and the Cauchy-Schwartz inequality, we can prove that if the results hold for a payoff function element of C_K^∞, the result is true for any function of $L^2[0,T]$. Second, using dominated convergence theorem, we can justify the interchange of order between the expectation and the differential operator as well as the interchange of the differential operator and the integral operator. We get therefore that a weight function w should satisfy

$$\delta = \mathbb{E}\left[f\left(\int_0^T X_t dt\right)\delta(w)\right]$$

The RHS is also equivalent to the following expressions:

$$\begin{aligned}
&\mathbb{E}\left[f\left(\int_0^T X_t dt\right)\delta(w)\right] \\
= \ &\mathbb{E}\left[f'\left(\int_0^T X_t dt\right)\int_{t=0}^T \int_{s=0}^T D_s X_t dt w_s ds\right] \\
= \ &\mathbb{E}\left[f'\left(\int_0^T X_t dt\right)\int_{t=0}^T \int_{s=0}^T Y_t Y_s^{-1}\sigma(s, X_s) 1_{\{s \le t\}} w_s dt ds\right]
\end{aligned}$$

where in the last equation, we have used the fact that $D_s X_t = Y_t Y_s^{-1}\sigma(s, X_s)1_{\{s\leq t\}}$ formula (4.4). The Left Hand Side (LHS) or the delta is defined as the gradient with respect to the initial condition, leading to the following developments:

$$\begin{aligned}
&\frac{\partial}{\partial x}\mathbb{E}\left[f\left(\int_0^T X_t dt\right)\right] \\
= \ &\mathbb{E}\left[f'\left(\int_0^T X_t dt\right)\int_{t=0}^T Y_t dt\right]
\end{aligned}$$

A weight function should verify that the two expressions derived should be equal for all functions f of L^2:

$$\begin{aligned}
&\mathbb{E}\left[f'\left(\int_0^T X_t dt\right)\int_{t=0}^T \int_{s=0}^T Y_t Y_s^{-1}\sigma(s, X_s) 1_{\{s \le t\}} w_s dt ds\right] \\
= \ &\mathbb{E}\left[f'\left(\int_0^T X_t dt\right)\int_{t=0}^T Y_t dt\right]
\end{aligned}$$

which is equivalent to the equality of the terms inside the conditional expectation.□

C.2 Complex Asian option delta: necessary and sufficient conditions

Using the same arguments as for the simple Asian option case, we can justify the interchange of integral and derivative operator and vice versa. The Right

Hand Side (RHS) is also equivalent to the following expressions:

$$\begin{aligned}
&\mathbb{E}\left[f\left(\int_0^T X_t dt, X_T\right)\delta(w)\right] \\
=\ &\mathbb{E}\left[\left\langle \frac{\partial}{\partial x_1} f\left(\int_0^T X_t dt, X_T\right)\int_0^T D_s X_t dt + \frac{\partial}{\partial x_2} f\left(\int_0^T X_t dt, X_T\right) D_s X_T, w_s \right\rangle\right]
\end{aligned}$$

using the fact that $D_s X_t = Y_t Y_s^{-1}\sigma(s, X_s)1_{\{s\leq t\}}$, this leads to

$$\begin{aligned}
&\mathbb{E}\left[f\left(\int_0^T X_t dt, X_T\right)\delta(w)\right] \\
=\ &\mathbb{E}\left[\left(\begin{array}{c} \frac{\partial}{\partial x_1} f\left(\int_0^T X_t dt, X_T\right)\int_{t=0}^T Y_t \left(\int_{s=0}^T Y_s^{-1}\sigma(s, X_s)1_{\{s\leq t\}} w_s ds\right) dt \\ +\frac{\partial}{\partial x_2} f\left(\int_0^T X_t dt, X_T\right) Y_T \left(\int_{s=0}^T Y_s^{-1}\sigma(s, X_s)1_{\{s\leq T\}} w_s ds\right) \end{array}\right)\right]
\end{aligned}$$

The Left Hand Side (LHS) is equivalent to the following expressions:

$$\begin{aligned}
&\frac{\partial}{\partial x}\mathbb{E}\left[f\left(\int_0^T X_t dt, X_T\right)\right] \\
=\ &\mathbb{E}\left[\frac{\partial}{\partial x_1} f\left(\int_0^T X_t dt, X_T\right)\int_0^T Y_t dt + \frac{\partial}{\partial x_2} f\left(\int_0^T X_t dt, X_T\right) Y_T\right]
\end{aligned}$$

A weight function should satisfy that the two expressions derived should be equal for all functions f of L^2:

$$\begin{aligned}
&\mathbb{E}\left[\left(\begin{array}{c} \frac{\partial}{\partial x_1} f\left(\int_0^T X_t dt, X_T\right)\int_{t=0}^T Y_t \left(\int_{s=0}^T Y_s^{-1}\sigma(s, X_s)1_{\{s\leq t\}} w_s ds\right) dt \\ +\frac{\partial}{\partial x_2} f\left(\int_0^T X_t dt, X_T\right) Y_T \left(\int_{s=0}^T Y_s^{-1}\sigma(s, X_s)1_{\{s\leq T\}} w_s ds\right) \end{array}\right)\right] \\
=\ &\mathbb{E}\left[\frac{\partial}{\partial x_1} f\left(\int_0^T X_t dt, X_T\right)\int_0^T Y_t dt + \frac{\partial}{\partial x_2} f\left(\int_0^T X_t dt, X_T\right) Y_T\right]
\end{aligned}$$

which is equivalent to the equality of the conditional expectation of inside terms.□

C.3 Particular solutions for the Complex Asian delta

To examine a particular solution, let us write $G_s = \sigma(s, X_s)\, w_s / Y_s$. We obtain that a stronger condition of the necessary and sufficient conditions is given by the equality of term inside the conditional expectation:

$$\begin{aligned}
\mathbb{E}_x^Q\left[\int_{t=0}^T Y_t\left(\int_{s=0}^t \frac{\sigma(s,X_s)}{Y_s} w_s^{delta} ds\right) dt \middle| \int_0^T X_t dt, X_T\right] &= \mathbb{E}_x^Q\left[\int_0^T Y_t dt \middle| \int_0^T X_t dt, X_T\right] \\
\mathbb{E}_x^Q\left[Y_T \int_{s=0}^T \frac{\sigma(s,X_s)}{Y_s} 1_{\{s\leq T\}} w_s^{delta} ds \middle| \int_0^T X_t dt, X_T\right] &= \mathbb{E}_x^Q\left[Y_T | \int_0^T X_t dt, X_T\right]
\end{aligned}$$

$$\frac{\int_{s=0}^T G_s \int_{t=s}^T Y_t dt}{\int_0^T Y_t dt} ds = 1 \qquad \text{(C.1)}$$

$$\int_{s=0}^T G_s ds = 1 \qquad \text{(C.2)}$$

C.3.1 First particular solution

If we assume a particular form of the function G with:

$$G_s = a + b\frac{Y_s}{\int_{t=0}^T Y_t dt}$$

we get

$$\int_{s=0}^T \frac{\int_{t=s}^T Y_t dt}{\int_{t=0}^T Y_t dt} G_s ds = a\frac{\int_{t=0}^T tY_t dt}{\int_{t=0}^T Y_t dt} + \frac{1}{2}b$$

where in the last equation, we have used Fubini theorem for the integration. We get then:

$$\int_{s=0}^T G_s ds = aT + b$$

The solution should therefore verify:

$$\begin{aligned}
a\frac{\int_{t=0}^T tY_t dt}{\int_{t=0}^T Y_t dt} + \frac{1}{2}b &= 1 \\
aT + b &= 1
\end{aligned}$$

The discriminant of this system is

$$\Delta = \frac{\int_{t=0}^T tY_t dt}{\int_{t=0}^T Y_t dt} - \frac{1}{2}T$$

and the solutions satisfy:

$$\begin{aligned}
a2\int_{t=0}^T tY_t dt + b\int_{t=0}^T Y_t dt &= 2\int_{t=0}^T Y_t dt \\
aT + b &= 1
\end{aligned}$$

in the particular case $\Delta \neq 0$, we get

$$a = \frac{\int_{t=0}^{T} Y_t dt}{2\int_{t=0}^{T} tY_t dt - T\int_{t=0}^{T} Y_t dt} \quad \text{(C.3)}$$

$$b = \frac{2\left(\int_{t=0}^{T} tY_t dt - T\int_{t=0}^{T} Y_t dt\right)}{2\int_{t=0}^{T} tY_t dt - T\int_{t=0}^{T} Y_t dt} \quad \text{(C.4)}$$

which is exactly the final result (4.11) with the two constants a and b defined by the conditions (C.3) and (C.4) respectively.□

C.3.2 Second particular solution

If we assume a particular form of

$$G_s = \alpha s + \beta \frac{sY_s}{\int_{t=0}^{T} Y_t dt}$$

we get the same type of system:

$$\alpha \frac{\int_{t=0}^{T} \frac{t^2}{2} Y_t dt}{\int_{t=0}^{T} Y_t dt} + \beta \frac{\int_{t=0}^{T}\int_{s=0}^{t} sY_s ds Y_t dt}{\left(\int_{t=0}^{T} Y_t dt\right)^2} = 1$$

$$\alpha \frac{T^2}{2} + \beta \frac{\int_{s=0}^{T} sY_s ds}{\int_{t=0}^{T} Y_t dt} = 1$$

The discriminant of this system is

$$\Delta = \frac{\int_{t=0}^{T}\int_{s=0}^{t} \frac{(t^2-T^2)}{2} sY_s ds Y_t dt}{\left(\int_{t=0}^{T} Y_t dt\right)^2}$$

The solution when the discriminant is not equal to zero is

$$\alpha = \frac{\int_{t=0}^{T}\int_{s=t}^{T} sY_s ds Y_t dt}{\int_{t=0}^{T}\int_{s=0}^{t} \frac{(t^2-T^2)}{2} sY_s ds Y_t dt}$$

$$\beta = \frac{\int_{t=0}^{T} \frac{(t^2-T^2)}{2} Y_t dt \int_{t=0}^{T} Y_t dt}{\int_{t=0}^{T}\int_{s=0}^{t} \frac{(t^2-T^2)}{2} sY_s ds Y_t dt}$$

leading to the solution (4.13).□

Appendix D

Inverse of the cumulative distribution of the Student density

The general algorithm (given in Abramovitz and Stegun (1970)) for computing the inverse t_p of the cumulative distribution of the Student density, with n degrees of freedom is given below with $0 < p < 1$ and with x_p the inverse of the cumulative distribution of the normal density. It is very useful for generating random number distributed according to a Student density. $N(0,1)$:

$$
\begin{aligned}
t_p &= x_p + \frac{g_1(x_p)}{n} + \frac{g_2(x_p)}{n^2} + \frac{g_3(x_p)}{n^3} + \frac{g_4(x_p)}{n^4} \\
g_1(x) &= \frac{1}{4}\left(x^3 + x\right) \\
g_2(x) &= \frac{1}{96}\left(5x^5 + 16x^3 + 3x\right) \\
g_3(x) &= \frac{1}{384}\left(3x^7 + 19x^5 + 17x^3 - 15x\right) \\
g_4(x) &= \frac{1}{92160}\left(79x^9 + 776x^7 + 1482x^5 - 1920x^3 - 945x\right)
\end{aligned}
$$

Appendix E

Wiener Chaos and Convexity

E.1 Introduction to Wiener Chaos

Introduced in finance by Lacoste (1996) (in an article about transaction costs) and by Brace and Musiela (1995), Wiener Chaos expansion could be intuitively thought of the generalization of Taylor's expansion to stochastic processes with some martingale considerations. This representation of stochastic processes initially proved for the Brownian motion by Wiener (1938) and later for Levy process (see Ito 1956) has been recently refocused, motivated by the contemporary development of the Malliavin calculus theory and its application not only to probability theory but also to mechanics, economics and finance (1995).

More precisely, we present in this section the basic properties of the chaotic representation for a given fundamental martingale. Let M be a square-integrable martingale with respect to an appropriate filtration called F_t with deterministic Doob Meyer brackets $\langle M \rangle_t$ (defined through the requirement that $(M_t^2 - \langle M \rangle_t)$ be a martingale). The latter property is vital for obtaining the chaotic orthogonal representation of the space $\mathcal{L}^2(\mathcal{F}_\infty)$. Let

$$C_n = \{(s_1, ..., s_n) \in \mathbb{R}^n, 0 < s_1 < ... < s_n < t\}$$

be the set of strictly increasingly-ordered n-uplets. Let $(\Phi_n)_{n \in \mathcal{N}}$ be the morphisms from $\mathcal{L}^2(C_n)$ to $\mathcal{L}^2(\mathcal{F}_\infty)$

$$\begin{aligned}\Phi_n(f) &: \mathcal{L}^2(C_n) \to \mathcal{L}^2(\mathcal{F}_\infty) \\ \Phi_n(f) &= \int_{0 \le s_1 \le .. s_i \le ... s_n \le T} f(s_1, ..., s_n)\, dM_{sn}\, dM_{s_1}\end{aligned}$$

The interesting property of the series of the images of $\mathcal{L}^2(C_n)$ by the morphisms $(\Phi_n)_{n \in \mathcal{N}}$ is the orthogonal decomposition of the space $\mathcal{L}^2(\mathcal{F}_\infty)$.

$$\mathcal{L}^2(\mathcal{F}_\infty) = \overset{\perp}{\underset{n}{\oplus}} \Phi_n\left(\mathcal{L}^2(C_n)\right)$$

This fundamental decomposition of the space $\mathcal{L}^2(\mathcal{F}_\infty)$ into sub-spaces, called M-chaos sub-spaces leads to the interesting representation of any function F of $\mathcal{L}^2(\mathcal{F}_\infty)$ in a form of a series of terms resulting from the orthogonal projection of the function F on the series of M-chaos sub-spaces.

$$F = \sum_n \Phi_n(f) = \sum_n \int_{C_n} f_n(s_1, ..., s_n)\, dM_{s_n} dM_{s_1}$$

where $f_n \in L^2(C_n)$. Deriving the Wiener Chaos expansion of a function f element of $L^2(F_\infty)$ is very simple as the following theorem proves it:

E.2 Theorem and proposition

Theorem 7 *Decomposition in Wiener Chaos*
Let $D^n F$ represent the nth derivative of function F according to its second variable. The M-chaos decomposition of the process $(F(t, M_t))_{t \ge 0}$ gives, for all $t \ge 0$,

$$F(t, M_t) = \mathbb{E}[F(t, M_t)] + \sum_{n=1}^{\infty} \mathbb{E}[D^n F(t, M_t)] \int_{C_n} dM_{s_n} ... dM_{s_1}$$

Proof : See Lacoste (1996) Theorem 3.1 p 201.

The following two propositions refer to important and useful results about Wiener Chaos.

Proposition 21 *Orthogonality of the different chaos*
The fundamental properties used are the orthogonality of the different chaos. Let $f_n \in L^2(C_n)$ and $f_m \in L^2(C_m)$ and let $(M)_{t\in\mathbb{R}^+}$ be a martingale process defined as in the previous section

$$\begin{aligned} &\mathbb{E}\left[\int_{C_n} f_n(s_1,...,s_n)\,dM_{s_n}dM_{s_1}\int_{C_m} f_m(s_1,...,s_m)\,dM_{s_m}dM_{s_1}\right] \\ = &\;\delta_{n,m}\int_{C_n} f_n(s_1,...,s_n)\,f_m(s_1,...,s_m)\,ds_1...ds_n \end{aligned}$$

with $\delta_{n,m}$ the Kronecker delta.

$$\begin{aligned} \delta_{n,m} &= 1 \quad if \quad n=m \\ &= 0 \quad otherwise \end{aligned}$$

The other result we use is the decomposition of a geometric Brownian motion (or a Doleans martingale).

Proposition 22 *Wiener Chaos decomposition of a geometric multi-dimensional Brownian motion*
The geometric multi-dimensional Brownian motion denoted by A_{T_k} can be expanded as the Hilbertian sum of orthogonal terms called Wiener Chaos of order i, denoted by I_i:

$$A_{T_k} = e^{\int_0^T\left\langle V_s^{(T,T_k)},d\widetilde{W}_s\right\rangle-\frac{1}{2}\int_0^T\left\|V_s^{(T,T_k)}\right\|^2 ds} \tag{E.1}$$

$$= \sum_{i=0}^{\infty} I_i(V,T,T_k) \tag{E.2}$$

with

$$I_0(V,T,T_k) = 1$$

$$I_{i,i>0}(V,T,T_k) = \frac{\int_{s_1\leq s_2\leq...\leq s_i\leq T}\left\langle V(s_1,T,T_k),d\widetilde{W}_{s_1}\right\rangle...\left\langle V(s_i,T,T_k),d\widetilde{W}_{s_i}\right\rangle}{i!}$$

Proof: see either Øksendal (1997) exercise p1.2.d. page 19 or Lacoste (1996) page 201 Theorem 3.1.□

E.3 Proof of the theorem

This appendix section gives the proof of therorem 5.

E.3.1 Finding the convexity adjustment

We remind some notations for the proof. We denote by K the sensitivity of the forward swap, $K = \sum_{i=1}^{n} B_{T_i}$. We also postulate that a zero coupon bond can be written as a normalized Doleans martingale times its value at time zero, leading to the following notation: $B_T^{(T,T_i)} = B_{T_i} A_{T_i}$ with $A_{T_i} = e^{\int_0^T \left\langle V_s^{(T,T_i)}, d\widetilde{W}_s \right\rangle - \frac{1}{2}\int_0^T \left\| V_s^{(T,T_i)} \right\|^2 ds}$ and $B_{T_i} = \frac{B(0,T_i)}{B(0,T)}$. We need to calculate the following quantity:

$$\Pi_0 = B(0,T)\,\mathbb{E}_{Q_T}\left(\frac{B_{T_0}A_{T_0} - B_{T_n}A_{T_n}}{\sum_{i=1}^{n} B_{T_i}A_{T_i}} \right)$$

Using the linearity of the expectation operator, we get that the above expression can be separated into two terms:

$$\frac{\Pi_0}{B(0,T)} = B_{T_0}\mathbb{E}_{Q_T}\left(\frac{A_{T_0}}{\sum_{i=1}^{n} B_{T_i}A_{T_i}} \right) - B_{T_n}\mathbb{E}_{Q_T}\left(\frac{A_{T_n}}{\sum_{i=1}^{n} B_{T_i}A_{T_i}} \right)$$

Using the technical lemma (by means of Wiener chaos expansion) proved below, we get that the two expectations can be approximated by the following expression:

$$\mathbb{E}_{Q_T}\left(\frac{A_{T_j}}{\sum_{i=1}^{n} B_{T_i}A_{T_i}} \right) = \frac{1}{K} - \frac{\sum_{i=1}^{n} B_{T_i} C(T_j, T_i)}{K^2} + \frac{\sum_{i,k=1}^{n} B_{T_i} B_{T_k} C(T_i, T_k)}{K^3} + O_3$$

with the signification of O_3 explained in the technical lemma. Rearranging the term, we get that the price of the expected swap rate could be written as a simple expression

$$\begin{aligned} \frac{\Pi_0}{B(0,T)} &= \frac{B_{T_0} - B_{T_n}}{K} + \frac{\sum_{i=1}^{n} B_{T_0} B_{T_i} C(T_0, T_i) - B_{T_n} B_{T_i} C(T_n, T_i)}{K^2} \\ &\quad + \frac{\sum_{i,k=1}^{n} B_{T_i} B_{T_k} C(T_i, T_k)}{K^2} \end{aligned}$$

which leads to the final result.□

E.3.2 Approximation using Wiener Chaos

Let O_3 denote a negligible quantity with respect to the $\left\|V_{\cdot}^{(T,T_i)}\right\|_{L^2}^3$, i.e.

$$O_3 = O\left(\left(\int_{s_1\leq s_2\leq s_3\leq T}\left\|V_{s_1}^{(T,T_i)}\right\|^2 ... \left\|V_{s_3}^{(T,T_i)}\right\|^2 ds_1...ds_3\right)^{1/2}\right)$$

We can prove the following lemma:

Lemma 4 *The expected value of the non linear stochastic expression* $\frac{A_{T_j}}{\sum_{i=1}^n B_{T_i}A_{T_i}}$ *can be given by a simple function of the correlation terms:*

$$\mathbb{E}_{Q_T}\left(\frac{A_{T_j}}{\sum_{i=1}^n B_{T_i}A_{T_i}}\right) = \frac{1}{K} - \frac{\sum_{i=1}^n B_{T_i}C(T_j,T_i)}{K^2} + \frac{\sum_{i,k=1}^n B_{T_i}B_{T_k}C(T_i,T_k)}{K^3} + \varepsilon$$

where the error term, ε, denotes a negligible quantity with respect to the $\left\|V_{\cdot}^{(T,T_i)}\right\|_{L^2}^3$, *i.e.* $\varepsilon = O_3$.

Proof: let us introduce some notations: $U_0 = 1$, $U_1 = \frac{\sum_{i=1}^n B_{T_i}I_1(V,T,T_i)}{K}$, $U_2 = \frac{\sum_{i=1}^n B_{T_i}I_2(V,T,T_i)}{S}$. By a Wiener Chaos expansion theorem 22, and result (E.2), we can expand the term A_{T_i} to get:

$$\sum_{i=1}^n B_{T_i}A_{T_i} = \sum_{i=1}^n B_{T_i} + \sum_{i=1}^n B_{T_i}I_1(V,T,T_i) + \sum_{i=1}^n B_{T_i}I_2(V,T,T_i) + \varepsilon_1$$

where the error term ε_1 is a negligible quantity with respect to the $\|V(T,T_i)\|_{L^2}^3$ ($\varepsilon_1 = O_3$). The simple Taylor expansion $\frac{1}{1+x} = 1 - x + x^2 + o(x^3)$ gives that the denominator can be written as a sum of terms:

$$\begin{aligned} &\frac{1}{\sum_{i=1}^n B_{T_i}A_{T_i}} \\ &= \frac{1}{K} - \frac{\sum_{i=1}^n B_{T_i}I_1(V,T,T_i)}{K^2} - \frac{\sum_{i=1}^n B_{T_i}I_2(V,T,T_i)}{K^2} + \frac{1}{K}\left(\frac{\sum_{i=1}^n B_{T_i}I_1(V,T,T_i)}{\sum_{i=1}^n B_{T_i}}\right)^2 + \varepsilon_2 \end{aligned} \tag{E.3}$$

where the error term ε_2 is a negligible quantity with respect to the $\left\|V_{\cdot}^{(T,T_i)}\right\|_{L^2}^3$ ($\varepsilon_2 = O_3$). In the expectation to calculate, $\mathbb{E}_{Q_T}\left(\frac{B_{T_j}A_{T_j}}{\sum_{i=1}^n B_{T_i}A_{T_i}}\right)$, the term A_{T_j} can be seen as a change of probability measure. We denote by Q^{T,T_j} the new probability measure defined by its Radon Nikodym derivative with respect to the forward neutral probability measure Q_T, and W_s^{T,T_j} the Q^{T,T_j} standard Brownian motion:

$$\begin{aligned}\frac{dQ^{T,T_i}}{dQ_T} &= e^{\int_0^T\left\langle V_s^{(T,T_j)},d\widetilde{W}_s\right\rangle-\frac{1}{2}\int_0^T\left\|V_s^{(T,T_j)}\right\|^2ds} \\ dW_s^{T,T_j} &= d\widetilde{W}_s - V_s^{(T,T_j)}ds\end{aligned}$$

Then the measure change eliminates the numerator term and simplifies the expectation to calculate as only a function of $\frac{1}{\sum_{i=1}^n B_{T_i}A_{T_i}}$ in a new probability measure Q^{T,T_j}. By linearity of the expectation operator and using the approximation (E.3) , we get

$$\begin{aligned}&\mathbb{E}_{Q^{T,T_j}}\left(\frac{1}{\sum_{i=1}^n B_{T_i}A_{T_i}}\right) \\ =\ & \frac{1}{K} - \mathbb{E}_{Q^{T,T_j}}\left(\frac{\sum_{i=1}^n B_{T_i}I_1(V,T,T_i)}{K^2}\right) - \mathbb{E}_{Q^{T,T_j}}\left(\frac{\sum_{i=1}^n B_{T_i}I_2(V,T,T_i)}{K^2}\right) \\ &+\frac{1}{K}\mathbb{E}_{Q^{T,T_j}}\left(\left(\frac{\sum_{i=1}^n B_{T_i}I_1(V,T,T_i)}{K}\right)^2\right) + \varepsilon_3\end{aligned}$$

where the error term ε_3 is a negligible quantity with respect to the $\left\|V_\cdot^{(T,T_i)}\right\|_{L^2}^3$ ($\varepsilon_3 = O_3$). One can derive the results after proving that:

$$\begin{aligned}\mathbb{E}_{Q^{T,T_j}}\left(I_1(V,T,T_i)\right) &= C(T_i,T_j) \\ \mathbb{E}_{Q^{TP,T_j}}I_2(V,T,T_i) &= O_3 \\ \mathbb{E}_{Q^{T,T_j}}\left(\left(\sum_{i=1}^n B_{T_i}I_1(V,T,T_i)\right)^2\right) &= \sum_{i,k=1}^n B_{T_i}B_{T_k}C(T_i,T_k) + O_3\end{aligned}$$

□

E.4 Results of the Quasi Monte Carlo simulation

This sub-section reports results of a Quasi Monte Carlo simulation for the four different models. The simulation was done using 10,000 draws. The convexity

Year	forward Swap Rates	CMS Swap	QMC price	convexity adjustment in basis point
0	4.163826	4.163826	4.163826	0
1	4.385075	4.43604	4.436145	5.57
2	4.600037	4.699187	4.699212	9.91
3	4.80722	5.951161	5.951101	14.39
5	5.13929	5.36107	5.36087	22.18
7	5.366385	5.649873	5.649921	28.35
10	5.586253	5.935744	5.935735	34.95

Table E.1: Convexity adjustment for Ho and Lee model. Result obtained with σ =1%

term was calculated on an interest rate curved dated September, 2, 1999. Interestingly, convexity adjustment are different depending on the model, but not that much different indeed.

Year	forward Swap Rates	CMS Swap	QMC price	convexity adjustment in basis point
0	4.163826	4.163826	4.163826	0
1	4.385075	4.400307	4.400318	1.52
2	4.600037	4.635506	4.635521	3.55
3	4.80722	4.868121	4.868136	6.09
5	5.13929	5.266523	5.266514	12.72
7	5.366385	5.579279	5.579263	21.29
10	5.586253	5.959299	5.959281	37.30

Table E.2: Convexity adjustment for Amin and Jarrow model. Results obtained with $\sigma_0 = 0.1\%$ and $\sigma_1 = 0.1\%$

Year	forward Swap Rates	CMS Swap	QMC price	convexity adjustment in basis point
0	4.163826	4.163826	4.163821	0.00
1	4.385075	4.441479	4.441467	5.64
2	4.600037	4.708704	4.708715	10.87
3	4.80722	4.963449	4.963459	15.62
5	5.13929	5.375376	5.375363	23.61
7	5.366385	5.662372	5.662368	29.60
10	5.586253	5.940745	5.940736	35.45

Table E.3: Convexity adjustment for Hull and White model. Results obtained with $\sigma = 1.1\%$ and $\lambda = 1\%$

Year	forward Swap Rates	CMS Swap	QMC price	convexity adjustment in basis point
0	4.163826	4.440826	4.440826	0.00
1	4.385075	4.440826	4.440812	5.58
2	4.600037	4.707352	4.707347	10.73
3	4.80722	4.961371	4.961368	15.42
5	5.13929	5.371831	5.371820	23.25
7	5.366385	5.657425	5.657414	29.10
10	5.586253	5.933928	5.933938	34.77

Table E.4: Convexity adjustment for Mercurio and Moraleda model. Results obtained with $\sigma = 0.9\% \quad \lambda = 1\% \gamma = 0.11\%$

Bibliography

Aase, K.K.: 1993, A Jump Diffusion Consumption based Capital Assert Pricing Model and the Equity Premium Puzzle, *Mathematical Finance* **3**, 65–84.

Abken P. A., Madan D. B. and Ramamurtie S.: 1996, Estimation of Risk-Neutral and Statistical Densities by Hermite Polynomial Approximation With an Application to Eurodollar Futures Options, *Federal Reserve Bank of Atlanta Working Paper, 96-5* .

Abramovitz M. and Stegun A.: 1970, *Handbook of Mathematical Function*, Dover.

Ahn, C.M. and Thompson, H.E. : 1988, Jump-diffusion processes and the term structure of interest rates, *The journal of finance* **43**, 155–174.

Alziary B., Decamps J-P, and Koehl P-F: 1997, A PDE Approach to Asian Options: Analytical and Numerical Evidence, *Journal of Banking and Finance* **21**, 613–640.

Amin K.I. and Jarrow R.A.: 1992, An Economic analysis of interest Rate Swaps, *Mathematical Finance* **2**, 217–232. Addison-Wesley.

Amin K.I. and V. Ng: 1993, Option Valuation with Systematic Stochastic Volatility, *Journal of Finance* **48**, 881–910.

Amin, K.J. : 1993, Jump diffusion option valuation in discrete time, *The journal of finance* **48**, 1833–1863.

Bachelier L.: 1900, Théorie de la spéculation, *Annales de l'Ecole Normale Supérieure*.

Bakshi G. and Madan D. J.: 1998, A Simplified Approach to the Valuation of Options, *University of Maryland, Working Paper*.

Bally V., Talay D.: 1995, The Euler scheme for Stochastic Differential Equations: Error analysis with Malliavin Calculus, *Mathematics and Computers in Simulation* **38**, 312–323.

Barucci E. and Mancino E.M. D.: 1997, WienerChaos and Hermite Polynomials Expansions for pricing and Hedging Contingent Claims, *Working paper, DIMADEFAS*.

Bates D.: 1996, Jumps and Stochastic Volatility: Exchange Rate Process Implicit in Deutsche Mark Options, *Review of Financial Studies* **9**(1), 69–107.

Bates, D.S. : 1991, The crash of 87: Was it expected? the evidence from options markets, *The journal of finance* **46**, 1009–1044.

Benhamou E.: 1999, A Martingale Result for the Convexity Adjustment in the Black Pricing Model, *London School of Economics, Working Paper*.

Benhamou E.: 2000a, A Generalisation of Malliavin Weighted Scheme for Fast Computation of the Greeks, *Financial Markets Group, London School of Economics, FMG Discussion Paper DP 350*.

Benhamou E.: 2000b, Pricing Convexity Adjustment with Wiener Chaos, *Financial Markets Group, London School of Economics, FMG Discussion Paper DP 351*.

Benhamou E.: 2000c, The Gamma Vega Parity and its Application, *London School of Economics, Working Paper*. April.

Bensaid, B., Lesne, J.P. Pagès, H. and Scheinkman, J.: 1992, Derivative asset pricing with transaction costs, *Mathematical Finance* **2**, 63–86.

Bermin Hans-Peter: 1998, Essays on Lookback and Barrier Options - A Malliavin Calculus Approach, *Ph.D. dissertation, Lund University* .

Bermin Hans-Peter: 1999, Hedging options : the Malliavin calculus approach versus the Delta hedging approach, *Preprint* .

Besseminder H.: 1991, Forward Contracts and Firm Value, *Journal of Financial and Qualitative Analysis* **26**, 519–572.

Bicksler J. and Chen A.H.: 1986, An Economic Analysis of Interest Swaps, *Journal of Finance* **41**, 645–656.

Black F.: 1976, The Pricing of Commodity Contracts, *Journal of Financial Economics* **3**, 167–179. March.

Black F., Derman E., Toy W.: 1990, A One Factor Model of Interest Rates and Its Application to Treasury Bond Options, *Financial Analysts Journal* pp. 33–39. Jan-Feb.

Black F. and Karinski P.: 1990, Bond and option Pricing When short Rates are lognormal, *Financial Analysts Journal* pp. 52–59. July-August.

Black F. and Scholes M.: 1973, The Pricing of Options and Corporate Liabilities, *Journal of Political Economy* **81**, 637–659.

Bouaziz L., Brys E. and Crouhy M.: 1998, The Pricing of Forward Starting Asian Options, *Journal of Banking and Finance* **18**, 823–829.

Bouchaud M., Cont R. and Potters M.: 1998, Financial Markets as Adaptive Systems, *Europhysics Letters* **41**(3).

Boyle P.P.: 1977, Options: A Monte Carlo Approach, *Journal of Financial Economics* **4**, 323–338.

Boyle P.P., Broadie M. and Glasserman P.: 1997, Monte Carlo Method for Security Pricing, *Journal of Economic Dynamic and Control* **21**(8-9), 1267–1321.

Boyle P.P., Joy C. and Tan K.S.B.: 1997, Quasi Monte Carlo Methods in Numerical Finance, *Working paper*. University of waterloo, Ontario N2L 3G1.

Boyle, P.P. and Vorst, T. : 1992, Option replication in discrete time with transaction costs, *Journal of finance* **47**, 271–293.

Brace A. and Musiela M.: 1995, Duration, Convexity and Wiener Chaos, *Working paper*. University of New South Wales, Australia.

Bratley P., Fox B.L. and Niederreiter H.: 1992, Implementation and Tests of Low Discrepancy Sequences, *ACM Transactions on Modeling and Computer Simulation* **2**(3), 195–213.

Brennan M.J., and E.S. Schwartz: 1978, Finite Difference Methods and Jump Processes Arising in the Pricing of Contingent Claims: A Synthesis, *Journal of Financial and Quantitative Analysis* **3**, 462–474.

Broadie M. and Glasserman P.: 1996, Estimating Security Price Derivatives using Simulation, *Management Science* **42**, 169–285.

Brotherton-Ratcliffe and Iben B.: 1993, Yield Curve Applications of Swap Products, *Advanced Strategies in financial Risk Management* pp. 400– 450 Chapter 15. R.Schwartz and C.Smith (eds.) New York: New York Institute of Finance.

Caflisch R.E., Mokoroff W. and Owen A.: 1997, Valuation of Mortgage Backed Securities Using Brownian Bridges to Reduce Effective Dimension, *Journal of Computational Finance* **1**(1). June.

Carr P. and Madan D. B.: 1999, Option Pricing and the Fast Fourier Transform, *Journal of Computational Finance* **3**(1). June.

Carverhill A. and Clewlow L.: 1992, Flexible Convolution, *From Black Scholes to Black Holes* pp. 165–171.

Chen, R.R. and Scott L.: 1992, Pricing Interest Rate Options in a Two Factor Cox-Ross-Ingersoll Model, *Review of Financial Studies* **5**, 613–636.

Cooper I.A. and Mello, A.S.: 1991, The default Risk of Swaps, *Journal of Finance* **46**, 597–620.

Courtadon G.: 1982, The pricing of options on default-free bonds, *Journal of financial and quantitative analysis* pp. 75–100.

Cox J. C., Ingersoll J.E. and Ross S.A.: 1981, The Relationship between Forward Prices and Futures Prices, *Journal of Financial Economics* **9**(December), 321–346.

Cox J. C. and Ross S.A.: 1976, The Valuation of Options for Alternative Stochastic Processes, *Journal of Financial Economics* **3**, 145–166.

Cox J. C., Ross S.A. and Rubinstein S.A.: 1979, Option Pricing: A Simplified Approach, *Journal of Financial Economics* **7**(December), 229–263.

Curran M.: 1994, Strata Gems, *RISK* pp. 70–71. March.

Curran M.: 1998, Greeks in Monte Carlo, *Monte Carlo Methodologies and Applications for Pricing and Risk Management* p. 79 Chapter 9. edited by Bruno dupire, Risk Publications.

Davis, M.H.A. and Panas, V. and Zariphopoulou, T.: 1993, European option pricing with transaction costs, *SIAM J. of Control and Optimization* **31**, 470–493.

Decreusefond L.: 1994, Méthodes de perturbation dans les réseaux de file d'attente - Applications aux réseaux haut débit, *Thèse de l'ENST Département Informatique et Réseaux* .

Derman E. and Kani A.: 1994, The volatility smile and its implied tree, *Preprint, December*.

Duffie D.: 1995, *Dynamic Asset Pricing Theory*, Princeton University Press.

Duffie D. and Huang M.: 1994, Swap rates and credit quality, *Graduate School of Business, Stanford University*.

Dumas B. and Fleming I. and Whaley R.: 1995, Implied volatility functions: empirical tests, *Preprint*.

Dupire B.: 1993, Pricing and hedging with smiles, *Preprint*.

El Karoui N., Geman H. and Rochet J.C.: 1995, Change of Numeraire, Changes of Probability Measure, and option Pricing, *Journal of Applied Probability* **32**, 443–458.

El Karoui, N. and Quenez, M-C.: 1995, Dynamic programming and pricing of contingent claims in an incomplete market, *SIAM J. control and Optim.* **33**, 29–66.

Elworthy K. D.: 1982, Stochastic Differential Equations on Manifolds, *Cambridge Univ. Press* pp. 383–405.

Elworthy K. D.: 1992, Stochastic Flows in Riemannian manifolds, *Diffusion Problems and Related Problems in Analytics* pp. 37–72. II eds : Pinsky M. A. and Vihstutz V., Birkhauser.

Fama E.F.: 1965, The Behavior of Stock Market Prices, *The Journal of Business* **38**, 34–105.

Forsyth K., V. and Zvan: 1998, Robust numerical methods for PDE models of Asian options, *Journal of Computational Finance* **1**. Winter.

Fournié E., Laszry J.M., Lebuchoux J. and Lions P.L.: 2000, Applications of Malliavin Calculus to Monte Carlo Methods in Finance. II., *Forthcoming in Finance and Stochastics* .

Fournié E., Laszry 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.

French K.: 1983, A Comparison of Futures and Forward Prices, *Journal of Financial Economics, 12, (November 83)* pp. 311–342.

Frey, R.: 1999, Superreplication in Stochastic Volatility Models and Optimal Stopping, *Finance and Stochastics* .

Galanti S. and Jung A.: 1997, Low discrepency sequences: Monte carlo simulation of option prices, *Journal of Derivatives* pp. 63–83.

Geman H. and Yor M.: 1993, Bessel processes, Asian Options and Perpetuities, *Mathematical Finance* **3**.

Glasserman P. and Yao D.D.: 1992, Some Guidelines and Guarantees for common random numbers, *Management Science, 38* pp. 884–908.

Glynn P.W.: 1989, Optimization of stochastic systems via simulation, *Proceedings of the 1989 Winter Simulation Conference, Society for Computer Simulation, San Diego, CA* pp. 90–105.

Gobet E.: 2000, Weak approximation of killed diffusion using Euler schemes, *To appear in Stochastic Processes and their Applications* .

Hagan P. and Woodward D.: 1998, Markov Interest Rate Models, *preprint* .

Harrison J.M. and Kreps D.M.: 1979, Martingales and Arbitrages in Multiperiod Securities Markets, *Journal of Economic Theory, 20* pp. 381–408.

Harrison J.M. and Pliska S.R.: 1981, Martingales and Stochastic Integrals in the Theory of Continuous Trading, *Stochastic Processes and Their Applications* pp. 55–61.

Hart Yan: 1997, Unifying Theory, *RISK, February* pp. 54–55.

He and Takahashi A.: 1996, A Variable Reduction Technique for Pricing Average-Rate-Options, *Japanese Journal of Financial Economics* **1**(2).

Heston S.L.: 1992, A closed form solution for options with stochastic volatility with applications to bond and currency options, *The review of Financial Studies* **6**, 327–343.

Ho T. S. Y. and Lee S.B.: 1986, Term Structure Movements and Pricing Interest Rate Contingent Claims, *Journal of Finance* **41**.

Hodges S.D. and Clewlow, L.J.: 1997, Optimal delta hedging under transactions costs, *Journal of Economic Dynamics and Control* **21**, 1353–1376.

Hodges S.D. and Neuberger, L.J.: 1989, Optimal Replication of Contingent Claims Under Transaction Costs, *The Review of Future Markets* **8**, 222–239.

Hull J.: 1990, Pricing Interest Rate Derivatives Securities, *Review of financial Studies* **3**, 573–592. Third Edition.

Hull J.: 1997, Options, Futures and Other Derivatives, *Prentice Hall* . Third Edition.

Hull J. and White, A.: 1987, The pricing of options on assets with stochastic volatilities, *The Journal of Finance* **17**, 281–300.

Jacques M.: 1996, On the Hedging Portfolio of Asian Options, *Astin Bulletin* **26**(2).

Jarrow R. A. and Rudd A.: 1982, Approximate option valuation for arbitrary stochastic processes, *Journal of financial Economics, 10* .

Jarrow, R.A. and Rosenfeld, E.R. : 1984, Jump risks and the intertemporal capital asset pricing model, *Journal of Business* **51**, 337–351.

Johanessma P.M. and Victor J.D.: 1986, Maximum-entropy approximations of stochastic nonlinear transductions: an extension of the Wiener theory, *Biological Cybernetics* **54**, 289–300.

Jorion P.: 1988, On Jump Processes in the Foreign Exchange and Stock market, *Reveiw of Financial Studies* **1**(4).

Jouini, E. and Kallal, H.: 1991, Arbitrage and equilibrium in securities markets with shortsale constraints, *Preprint. Laboratoire d'économètrie, Ecole Polytechnique.* .

Jouini E., Koehl P.F. and Touzi, N.: 1996, Incomplete markets, transaction costs and liquidity effects, *Preprint* .

Karatzas I. and Shreeve S.E.: 1988, Brownian Motion and Stochastic Calculus, *Springer Verlag* .

Kemma A. and Vorst, A.: 1990, A Pricing method for options based on average asset values, *Journal of Banking and Finance* **14**.

Kirikos G. and Novak D.: 1997, Convexity Conundrums, *RISK, March* .

Kohatsu A., F.: 1999, Rate of Convergence of a Particle Method to the Solution of the Mc Kean-Vlasov's Equation, *Department of Economics and Business, Universitat Pompeu Fabra Economics Working Papers* .

Kon T.S.: 1984, Model of Stock Returns: A Comparison, *Journal of Finance, 39* .

Lacoste V.: 1996, Wiener chaos: A New Approach to Option Hedging, *Mathematical Finance, vol 6, No.2* .

Lamberton D and Lapeyre B.: 1991, *Introduction au Calcul Stochastique Appliqué à la Finance*, Mathématiques et Applications, Ellipse, Paris.

L'Ecuyer P. and Perron G.: 1994, On the convergence rates of IPA and FDC derivative estimators, *Operation Research, 42* pp. 643–656.

Leland, R.E. : 1985, Option pricing with transaction costs, *The Journal of finance* **11**, 1283–1301.

Levy E. and Turnbull S.: 1992, Average Intelligence, *Risk Magazine* **5**.

Longstaff, F.A.: 1995, Hedging Interest Rate Risk with Options on Average Interest Rates, *The Journal of Fixed Income* . March.

Mandelbrot B.: 1963, The variation of certain speculative prices, *Journal of Business* **36**.

Melino, A. and S.M. Turnbull: 1990, Pricing Foreign Currency Options with Stochastic Volatility, *Journal of Econometrics* **45**.

Mercurio F. and Moraleda J.M.: 1996, A Family of Humped Volatility Structures, *Erasmus University, Working Paper* .

Merton, R.: 1976, Option pricing when underlying stock returns are discontinuous, *Journal of financial Economics* **3**, 125–144.

Milevsky M. A. and Posner S. E.: 1997, The Sum of Lognormals and the Reciprocal Gamma Distribution, *Schulich School of Business, Working Papers* pp. 16–97.

Millet A., Sanz Sol: 1997, A Stochastic wave equation in two space dimension : Smoothness of the law, *Universit Paris 6 Pierre et Marie Curie Working paper PMA-410* .

Morokoff, W.J.: 1997, Generating Quasi-Random Paths for Stochastic Processes, *Working Paper, UCLA and CATS Software Inc* .

Moskovitz B. and Caflisch R.: 1995, Smoothness and Dimension Reduction in Quasi-Monte-Carlo Methods, *Mathematical Computational Modelling* **23**(8/9), 63–65.

Musiela M. and Rutkowski M.: 1997, *Martingale Methods in Financial Modelling*, Springer Verlag.

Neuberger A.: 1999, Hedging Long-Term Exposures with Multiple Short-Term Futures Contracts, *Review of Financial Studies* **12**(3), 429–459. Fall 1999.

Nielsen J. A. and Sandmann K.: 1998, The Pricing of Asian Options under Stochastic Interest Rates, *University of Aarhus, Working Papers* **9**, 16–97. April.

Nualart D.: 1995, Malliavin Calculus and Related Topics, Probability and its Applications, *Springer Verlag* .

Øksendal B: 1997, An introduction to Malliavin Calculus with Applicaiton to Economics, *Working Paper Department of Mathematics University of Oslo May* .

Papageorgiou. A. and Traub J.: 1996, Beating Monte Carlo, *Risk* **9**(6), 63–50. June 1996.

Park, H. Y. and Chen A. H.: 1985, Differences between Futures and Forward Prices: A Further Investigation of Marking to Market Effects, *Journal of Future Markets* **5**, 77–88. February.

Paskov S.: 1994, Computing High Dimensional Integrals with Applications in Finance, *Technical Report CUCS, Columbia University 023-94* .

Paskov S. and Traub J.: 1995, Faster Valuation of Financial Derivatives, *Journal of Portfolio Management* pp. 113–120. Fall 1995.

Pradier L. and Lewicki P.: 1999, Interest Rate Option with Smile, *RISK* .

Press W., Teukowsky S., Vetterling W. and Flannery B. P.: 1992, Numerical Recipes in C (The art of Computing), *Cambridge University Press* .

Revuz D. and Yor M.: 1994, Continuous Martingales and Brownian Motion, *Second Edition, Springer Verlag* .

Rogers L. and Shi Z. S.: 1995, The Value of an Asian Option, *Journal of Financial Economics* **32**(9), 1077–1088.

Rozovskii, B.: 1997, Nonlinear Filtering, SPDE's, and Wiener Chaos, *SIAM* .

Rubinstein, M.: 1994, Implied binomial trees, *Journal of Finance* **49**, 771–818.

Schwartz ES: 1977, The Valuation of Warrants: Implementing a New Approach, *Journal of Financial Economics* **4**, 79–94.

Scott L.: 1997, Pricing Stock Options in a Jump-Diffusion Model with Stochastics Volatility and Interest Rates: Application of Fourier Inversion Methods, *Mathematical Finance* **7**, 413–426.

Serrat A.: 1996, A Dynamic Equilibrium Model of International Risk-Sharing Puzzles, *University of Chicago Graduate Working paper* .

Sharpe W.F.: 1978, *Investments*, Prentice-Hall, Englewoods cliff, New Jersey.

Stein E. and Stein, J.: 1991, Stock price distribution with stochastic volatility: an analysis approach, *Review of financial studies* **4**, 727–752.

Titman S.: 1992, Interest Rate Swaps and Corporate Financing Choices, *Journal of Finance* **47**, 1503–1516.

Turnbull S. M. and Wakeman, L.M.: 1991, Quick Algorithm for Pricing European average options, *Journal of Financial and Quantitative Analysis* **16**(3), 377–389.

Uwe F., Landre R. and Schott R.: 1998, Malliavin Calculus for Quantum Stochastic Processes, *University Ernst Moritz Arndt of Greifswald Institute of Mathematics and Computer Science, Working Papers* .

Vasicek, O. A.: 1977, An Equilibrium Characterization of the Term Structure, *Journal of Financial Economics* **5**, 177–188.

Viswanath, P. V.: 1989, Taxes and the Futures-Forward Price Difference in the 91-day T-Bill Market, *Journal of Money Credit and Banking* **21**(2), 190–205.

Vorst, T.C.F.: 1992, Prices and Hedge Ratios of Average Exchange Rate Options, *Journal of Financial and Quantitative Analysis* **26**(3), 337–389.

Wall L.D. and Pringle J.L.: 1989, Alternative Explanations of Interest Rate Swaps: A Theoretical and Empirical Analysis, *Financial Management* **18**, 59–73.

Wiener, N.: 1938, The Homogeneous Chaos, *American Journal Mathematics* **55**, 897–937.

Wiggins, J.: 1987, Option values under stochastic volatility, *Journal of financial economics* **19**, 351–372.

Williard, G.A.: 1997, Calculating Prices and Sensitivities for Path Independent Derivatives Securities in Multifactor Models, *Journal of Derivatives* pp. 45–61. Fall 1997.

Wilmott P., Dewyne J. and Howison S.: 1993, *Option Pricing: Mathematical Models and Computation*, Oxford Financial Press.

Zhang P.G. : 1995, Flexible Arithmetic Asian Options, *Journal of Derivatives* **2**, 53–63.

Zhang P.G. : 1998, *Exotic Options: A guide to Second Generation Options*, Word scientific.