POLITECNICO DI MILANO

Facoltà di Ingegneria dei Sistemi

Corso di Studi in INGEGNERIA MATEMATICA


Tesi di Laurea di Secondo Livello

# Stochastic Collocation and option pricing

Relatore: Prof. Fabio Nobile

Relatore esterno: Ing. Marzio Sala

Candidato:

Francesco Pizzi

matr.736757

Anno Accademico 2010-2011

## Abstract

This thesis deals with various applications to financial problems of the Stochastic Collocation method on Smolyak-type sparse grids.
In the first part, after the description of the method, we present this method and define both elliptic and parabolic Stochastic Partial Differential Equations (SPDEs) together with some of the well-posedness and convergence results known so far that will be then tested on simple cases.
In the second part we focus on financial models and apply the Stochastic Collocation to well known Black-Scholes equation under the hypothesis of stochastic parameters. We then analyze how this technique can be useful for Stochastic Differential Equations approximation considering both weak and strong errors. A comparison with Monte Carlo method in terms of computational cost is also performed. After deriving (in a heuristic way) a conditional Black and Scholes equation which fits to Basket Options or option pricing under stochastic volatility models we finally show the effectiveness of Stochastic Collocation in option pricing which is much more evident for Basket Options.

# Contents

# Chapter 1

# Introduction

This thesis deals with Stochastic Partial Differential Equations (SPDEs) that have gained a huge popularity in the last decade since one of the causes making simulations different from physical phenomena observable in nature is the use of uncertain data. In addition to coefficients, other sources of error are represented by forcing terms, boundary conditions and geometry of the problem. Often these data can not be estimated exactly because of lack of experimental data or intrinsic variability of the phenomenon and thus a deterministic approach would be unfulfilling.
After an appropriate modelization of the uncertainties, the simplest way to deal with them would be a random sampling (just like the Monte Carlo method). Nowadays however more efficient strategies have been proposed: the so called Stochastic Galerkin and Stochastic Collocation techniques. The key idea of this work is to apply tha latter technique and the related quadrature formulae to financial problems as proposed by the Quantitative Analysts from *UBS*. During a three months internship in the Zurich base we moved the first steps considering standard elliptic and parabolic Partial Differential Equations with random coefficients to evaluate the effectiveness of this method since these applications have been partially studied and convergence estimates were available. We then focused on the Black and Scholes equation with random parameters such as volatility and interest rates, treated as single random variables. Finally, we have focused on option pricing under Heston model of volatility where a Black and Scholes equation is written conditional to the volatility process. In this first stage we were able to obtain good results with the Black and Scholes equation and the Heston model under the hypothesis of non correlation between the Brownian motions driving the underlying and the volatility dynamics. In the case of non zero correlation we were not able to get positive results, in particular the Black and Scholes equation conditional to the volatility realization appeared "tricky" and difficult to approximate. For this reason our further efforts focused on the analyses of the approximation of stochastic differential equations with Stochastic Collocation. This investigation has led to substantial improvements with respect to the initial implementation.

In fact the main request from the UBS Quants was to find a different pricing method when several processes are involved: for example in case of stochastic volatility models with one underlying the pricing techniques consist in solving a bidimensional PDE or approximating the two SDEs and the same holds for Basket Options. Their idea was to approximate one of the SDEs with the Stochastic Collocation and then solving a conditional PDE. The results that we have obtained obtained show that for Basket Options this strategy can be competitive with those known so far, while in the Heston case the terms appearing in the conditional PDE are too complex to be approximated in an efficient way although the PDE derived seems to be the correct one.
We would like to point out that the PDE considered has an irregular drift and we can not provide well posedness results or to justify the use of Galerkin methods so this work is just one of the first attempts to implemente this new strategy. The numerical simulations however seem very encouraging and should represent a starting point for further analyses on this topic.

**Thesis structure**

In Chapter 2 we review the main ideas and concepts behind the Stochastic Collocation method and describe the construction of a sparse grid.
In Chapter 3 we will provide the mathematical formulation of an elliptic SPDE and its discrete version obtained with a Galerkin space discretization (tipically Finite Elements) and the Stochastic Galerkin or Stochastic Collocation probability approximation. We will not investigate further the first approach, while for the latter we will recall the convergence results available for some types of (sparse) grids, since this topic is still under investigation. In Section 3.2 we will then show on a couple of toy models the sharpness of these convergence rates. The same analysis will be done for parabolic SPDEs as well.
The main goal of our work however is to test the effectivness of the Stochastic Collocation method in financial applications. After recalling the Black-Scholes model and its extensions to forex contingent claim (Garman-Kohlhagen model), in Chapter 4 we apply the Stochastic Collocation technique assuming a uniform distribution for the volatility and/or the risk-free interest rate. Sometimes we will use the Stochastic Collocation just through the quadrature formulae involved (see Chapter 2) since closed-type pricing formulae are available; in more complex cases (like Barrier Options) we will solve parabolic SPDEs. Furthermore in Section 4.3 and 4.4 we will make a comparison with the Monte Carlo method.
We will then focus on Stochastic Differential Equations (SDEs) approximations (Chapter 5): in particular we will deal with some of the most famous SDEs that have been used to modelize mostly the volatility e.g. Geometric Brownian Motion, Ornstein-Uhlenbeck and Square Root process. We will show the efficacy of the Stochastic Collocation technique coupled with Euler and Milstein schemes by computing both weak errors (through the above-mentioned quadra-

ture formulae) and strong errors (through the interpolation procedure presented in Chapter 2). This represents a quite new approach and we will rely just on numerical simulations since no theoretical results have been produced on this topic so far. Afterwards in Section 5.6 we will compare the error-costs ratio involved in the probability approximation with Stochastic Collocation and Monte Carlo method. We will take into account also the theoretical rates of advanced methods as the Quasi-Monte Carlo or the Multi-level Monte Carlo, showing that Stochastic Collocation can be comparable and often superior to them. In Section 5.7 we will present a different discretization scheme for a Wiener path introduced by Caflisch [28] and adopted also by Griebel [31]: the so called Brownian Bridge discretization that will be used on anisotropic grids and in some cases provide really excellent convergence rate.
The last Chapter is dedicated to the pricing of contingent claims with the Heston volatility model and also European vanilla Basket options. We will first derive a conditional Black and Scholes PDE based on the arguments in [7] in a heuristhic way that will allow us to solve one SDE and one monodimensional PDE as a pricing technique. In fact, we will compare the prices of European vanilla Basket options written on two underlyings obtained by solving two SDEs or one SDE and our monodimensional PDE which in this case can be considered a non-standard SPDE with random drift and random boundary conditions. We will first analyze two pricing strategies based on Monte Carlo method and show that the conditional PDE approach leads to a mild variance reduction effect. We then explore the possibility of combining the conditional PDE with a Stochastic Collocation discretization of the PDE, relying on the results obtained in Chapter 5.
Then we will price vanilla and barrier options with the Heston volatility model under the assumption of uncorrelation between the Brownian motions driving the dynamics of the underlying and the volatility process: in this case the monodimensional PDE (which coincide with the Black-Scholes PDE) can be considered somehow a non-standard SPDE with random diffusion coefficient. With this assumption the Stochastic Collocation techique seems to be really efficient. On the other hand, when removing the hypothesis of zero correlation the conditional PDE strategy combined with Stochastic Collocation partially loses its effectiveness.
Finally a brief appendix recalls the Rannacher method that has been used to discretize in time parabolic PDEs.

# Chapter 2

# Stochastic Collocation

Let $\eta$ be a random variable taking values in $\Gamma$ with probability density function $\rho : \Gamma \to \mathbb{R}_+$. We define the space of square integrable functions

$$L^2_\rho(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} :< f, f >_\rho = \int_\Gamma f(\eta)^2 \rho(\eta) d\eta < \infty\} \tag{2.1}$$

which is a Hilbert space and consider a sequence of polynomials $\{Q_n\}_{n=0}^\infty$ of degree $n = 0, 1, \ldots$ orthogonal with respect to the measure $\rho(\eta)d\eta$, i.e. $< Q_n, Q_m >_\rho = 0$. Depending on the distribution of $\eta$ it is possible to introduce different basis for the space $L^2_\rho(\mathbb{R})$, in particular

- *Legendre polynomials* $\{Le_n\}_{n=0}^\infty$ are an orthogonal basis with respect to the uniform weight $\rho(\eta) = \frac{1}{2}$ on $[-1, 1]$ and they satisfy the following recurrence relation

$$Le_{n+1}(\eta) = \frac{2n+1}{n+1}\eta Le_n(\eta) - \frac{n}{n+1}Le_{n-1}(\eta) \tag{2.2}$$

  with $Le_0(\eta) = 1$ and $Le_1(\eta) = \eta$

- *Hermite polynomials* $\{H_n\}_{n=0}^\infty$ are an orthogonal basis with respect to the weight function $\rho(\eta) = \frac{1}{\sqrt{2\pi}} exp(-\frac{\eta^2}{2})$ on $(-\infty, +\infty)$ and they satisfy the following relation

$$H_{n+1}(\eta) = \eta H_n(\eta) - nH_{n-1}(\eta) \tag{2.3}$$

  with $H_0(\eta) = 1$ and $H_1(\eta) = \eta$

- *Laguerre polynomials* $\{La_n\}_{n=0}^\infty$ are an orthogonal basis with respect to the weight function $\rho(\eta) = e^{-\eta}$ on $[0, \infty)$ and they can be obtained through the following relation

$$La_{n+1}(\eta) = \frac{1}{n+1}[(2n+1-\eta)La_n(\eta) - nLa_{n-1}(\eta)] \tag{2.4}$$

  with $La_0(\eta) = 1$ and $La_1(\eta) = -\eta + 1$.

We aim at approximating the function $f$ in the polynomial subspace $\Sigma \subset L^2_\rho(\mathbb{R})$ that is a space of polynomials up to degree $w$ and we can do it in two different ways:

- $L^2_\rho$ projection on $\Sigma$: $f(\eta) \simeq f_w(\eta) = \sum_{i=0}^{w} f_i Q_i(\eta)$ where $f_i = \mathbb{E}[fQ_i]/\mathbb{E}[Q_i^2]$ and $Q_i$ are the orthogonal polynomials with respect to $\rho$ (e.g. Legendre, Laguerre and Hermite polynomials)
- interpolation on Gauss point: if $\eta_i$ $i = 0 \dots w$ are the roots of $Q_{w+1}$ and $l_i(\eta)$, $i = 0, \dots, w$ the corresponding Lagrange polynomials (which represent an alternative basis of $\Sigma$) then $f(\eta) \simeq f_w(\eta) = \sum_{i=0}^{w} f(\eta_i) l_i(\eta)$.

Given the coefficients of the second representation, it is possible to obtain those of the first one.

For both interpolation and $L^2$-projection, assuming $f$ analytic in $\Gamma$, we have the same convergence rate depending on the support of the density $\rho$: if bounded then

$$||f - f_w||_{L^2_\rho} \leq Ce^{-\tau w} \tag{2.5}$$

for a suitable $\tau > 0$. Otherwise, if $\Gamma$ is unbounded and $\rho(\eta) \sim e^{-\delta\eta^2}$ at infinity:

$$||f - f_w||_{L^2_\rho} \leq C\sqrt{w+1}e^{-\tau\delta\sqrt{w}} \tag{2.6}$$

for $f$ analytic such that $max_\eta(|f(\eta)|e^{-\alpha|\eta|}) < \infty$ for some $\alpha > 0$. We will discuss an analogous result in Section 3.1.3 .

Usually it is useful to evaluate first order moments, in particular the mean value, which can be calculated with the corrisponding Gaussian quadrature formula:

$$\mathbb{E}_\rho[f] = \int_\Gamma f(\eta)\rho(\eta)d\eta \simeq \sum_{i=0}^{w+1} w_i f(\eta_i) \tag{2.7}$$

with

$$w_i = \int_\Gamma l_i(\eta)^2 \rho(\eta) d\eta. \tag{2.8}$$

The monodimensional theory can be generalized to a function $f$ depending on a vector of independent random variables $\boldsymbol{\eta} = [\eta_1, \dots, \eta_N]$, each of them with corresponding image set $\Gamma_n = \eta_n(\Omega)$, where $\Omega$ is the set of all possible outcomes. If we assume that $\rho$ is the joint distribution of these N independent random variables then $\rho(\boldsymbol{\eta}) = \prod_{n=1}^{N} \rho_n(\eta_n)$ and we define $\Gamma = \Gamma_1 \times \dots \times \Gamma_N$. As in the mondimensional case we consider both a $L^2_\rho$-projection and an interpolation formula. Given a multi-index $\mathbf{k} = [k_1, \dots, k_N] \in \mathbb{N}^N$ we introduce a generic

basis function:

$$\mathbf{Q_k}(\boldsymbol{\eta}) = Q^1_{k_1}(\eta_1)\dots Q^N_{k_N}(\eta_N) \tag{2.9}$$

where each $Q_{k_i}$ is a monodimensional orthogonal polynomial (previously defined in the $L^2_\rho$ projection). Letting $\mathbf{k}$ vary in $\mathbb{N}^N$ we get an orthogonal basis for $L^2_\rho(\Gamma)$. Obviously we aim at approximating every function on a finite dimensional basis of $L^2_\rho(\Gamma)$ so a constraint on the set of the multi-indeces $\mathbf{k}$ has to be introduced. As a first step, we will take into account the set $\{\mathbf{k} : k_i \leq p_i\}$: this choice corresponds to the basis composed of all polynomials with maximum degree $p_i$ in each variable. To every vector of indeces $\mathbf{k} = [k_1, \dots, k_N]$ we associate the global index

$$k = k_1 + p_1(k_2 - 1) + p_1p_2(k_3 - 1)\dots \tag{2.10}$$

which allows to represent every function $f$ as:

$$f(\boldsymbol{\eta}) \simeq f_p(\boldsymbol{\eta}) = \sum_{k=1}^{N_p} f_k \mathbf{Q}_k(\boldsymbol{\eta}) \quad \text{where} \quad f_k = \mathbb{E}[f\mathbf{Q}_k]/\mathbb{E}[\mathbf{Q}_k^2] \tag{2.11}$$

and $\mathbf{p} = [p_1, \dots p_N]$ and $N_p = \prod_{n=1}^N (p_n + 1)$.

We will now focus on the approximation based on the Gauss interpolation. For each dimension $n = 1, \dots, N$ let $\eta_{n,k_n}, 1 \leq k_n \leq p_n + 1$ be the $p_n + 1$ roots of the one dimensional orthogonal polynomial $Q_{p_n+1}$ with respect to the weight $\rho_n$. We introduce the global index (2.10) and denote by $\boldsymbol{\eta}_k$ the point $[\eta_{1,k_1}, \eta_{2,k_2}, \dots \eta_{N,k_N}] \in \Gamma$. For each $n = 1, \dots, N$ we also introduce the Lagrange basis$\{l_{n,j}\}_{j=1}^{p_n+1}$ such that $l_{n,j}(\eta_{n,k}) = \delta_{j,k}$, $j, k = 1 \dots p_n + 1$ and we set $l_k(\boldsymbol{\eta}) = \prod_{n=1}^N l_{n,k_n}(\eta_n)$.
For any continuous function $f : \Gamma \to \mathbb{R}$ we have the following approximation:

$$f(\boldsymbol{\eta}) \simeq I_p f(\boldsymbol{\eta}) = \sum_{k=1}^{N_p} f(\boldsymbol{\eta}_k) l_k(\boldsymbol{\eta}) \tag{2.12}$$

where $N_p = \prod_{n=1}^N (p_n + 1)$.
In this framework we can also derive a Gauss quadrature formula for the mean value of the function $f$:

$$\mathbb{E}_\rho[f] = \int_\Gamma f(\boldsymbol{\eta})\rho(\boldsymbol{\eta})d\boldsymbol{\eta} \simeq \sum_{k=1}^{N_p} w_k f(\boldsymbol{\eta}_k) \tag{2.13}$$

where

$$w_k = \prod_{n=1}^N w_{k_n} \qquad w_{k_n} = \int_{\Gamma_n} l^2_{k_n}(\eta_n)\rho_n(\eta_n)d\eta_n. \tag{2.14}$$

This formula is exact if the integrand is a polynomial with degree at most $2p_n + 1$ with respect to each random variable $\eta_n$.

As in the multidimensional case an estimation of the rate of convergence of this polynomial approximation can be obtained, but we will discuss it in details later (see Section 3.1.3).
In the next chapter we will apply this theory to elliptic and parabolic partial differential equations.

## 2.1 Sparse grids approximation

The Collocation method presented so far consists in a tensor product approximation and interpolates a function depending on a set of $N$ random variables on a full tensor grid obtained by tensorization of one-dimensional Gaussian abscissas. This strategy becomes computationally expensive if the number of random variables increases significantly: this phenomenon is usually referred to as *curse of dimensionality* (see [4]).
Furthermore, the main consequence is a strong decrease in the convergence rate with respect to the number of collocation points. In fact let us consider a full tensor grid approximation involving $w+1$ points for each variable. The convergence result available (discussed in Section 3.1.3) provides an exponential decay of the error with respect to the level $w$ (the maximum polynomial degree). However the number of collocation points is equal to $N_p = (1+w)^N$ so we can recast the rate as $err \leq Cexp\{-\tau N_p^{\frac{1}{N}}\}$. If $N$ is sufficiently large than we can state that $N_p^{\frac{1}{N}} \simeq 1 + log(N_p)/N$ and the effective rate of convergence, that takes into account the number of points, is algebraic rather than exponential: in fact $err \leq CN_p^{-\tau/N}$.
In many cases a full tensor grid approximation is unnecessary and the number of evaluation points can drastically be reduced. This approach yields to the Sparse grids collocation method introduced by Smolyak ([23]), which is still based on a sequence of tensor grids, but allows to alleviate the curse of dimensionality.

To define a polynomial approximation on a sparse grid, first of all we consider the one dimensional interpolant operator $I_n^{m(i)}$ with respect each random variable $\eta_n$. The index $i \geq 1$ represents the level of approximation and $m$ the number of collocation points defining the interpolant at level $i$, under the assumptions that $m(1) = 1$, $m(i) < m(i+1)$ for $i \geq 1$ and $m(0) = 0$, which means that $I_n^{m(0)} = 0$. The evaluating points $\{\eta_{n,j}^{(i)}\}, j = 1 \dots m(i)$ for the one-dimensional operator $I_n^{m(i)}$ are the roots of the orthogonal polynomials $Q_{n,m(i)}$ with respect to the distribution of $\eta_n$. We then introduce the difference operators $\Delta_n^{m(i)} = I_n^{m(i)} - I_n^{m(i-1)}$. Following [4], given an integer $w$ and a multi-index $\mathbf{i} = (i_1, \dots, i_N), \mathbf{i} \geq \mathbf{1}$, we introduce a function $g : \mathbb{N}_+^{\mathbb{N}} \to \mathbb{N}$ strictly increasing in each argument and define a generalized

sparse grid approximation of a function $f$ as

$$S_w^{m,g}[f] = \sum_{\mathbf{i}\geq\mathbf{1}:g(\mathbf{i})\leq w} \bigotimes_{n=1}^{N} \Delta_n^{m(i_n)}(f). \tag{2.15}$$

Formula (2.15) can be recast in a more amenable way as

$$S_w^{m,g}[f] = \sum_{\mathbf{i}\geq\mathbf{1}:g(\mathbf{i})\leq w} c(\mathbf{i}) \bigotimes_{n=1}^{N} I_n^{m(i_n)}(f) \qquad \text{where} \quad c(\mathbf{i}) = \sum_{\mathbf{z}\in\{0,1\}^N:g(\mathbf{i}+\mathbf{z})\leq w} (-1)^{|\mathbf{z}|}. \tag{2.16}$$

Formula (2.16) shows that sparse grid approximation is still given by the sum of tensor product interpolations, but the constraint $g(\mathbf{i}) \leq w$ avoids using high order interpolants in each direction jointly. Furthermore the coefficients $c(\mathbf{i})$ is often equal to 0 and this implies that many tensor interpolations are not computed in practice. From formula (2.16) it follows that every function $f$ has to be evaluated in the set $H_w^{m,g} \subset \Gamma$, called the *sparse grid*, that consists of the union of all full tensor grids in (2.15) that are multiplied by a coefficient $c(\mathbf{i}) \neq 0$.
We have introduced the function $g$, which defines the set of admissible multi-indices $\mathbf{i}$. The requirement for $g$ to be strictly increasing guarantees that if $\mathbf{i}$ is an admissible multi-index than all $\mathbf{z} \leq \mathbf{i}$ are admissible as well.

We now want to define the polynomial space spanned by a sparse grid approximation of level $w$. First of all, since $m$ might not be surjective in $\mathbb{N}_+$ we introduce the left inverse:

$$m^{-1}(k) = min\{i \in \mathbb{N}_+ : m(i) \geq k\} \tag{2.17}$$

$m^{-1}$ is therefore a nondecreasing function. Let $\mathbf{m}(\mathbf{i}) = (m(i_1), \ldots, m(i_N))$ and consider the set of polynomial multidegrees

$$\Lambda^{m,g}(w) = \{\mathbf{p} \in \mathbb{N}^N : \ g(\mathbf{m}^{-1}(\mathbf{p}+\mathbf{1})) \leq w\}. \tag{2.18}$$

We denote by $\mathbb{P}_{\Lambda^{m,g}(w)}(\Gamma)$ the corresponding polynomial space spanned by the monomials with multidegree in $\Lambda^{m,g}(w)$

$$\mathbb{P}_{\Lambda^{m,g}(w)}(\Gamma) = span\Big\{ \prod_{n=1}^{N} \eta_n^{p_n} : \mathbf{p} \in \Lambda^{m,g}(w)\Big\}. \tag{2.19}$$

As stated in [4], if $f \in \mathbb{C}^0(\Gamma)$ then $S_w^{m,g}[f] \in \mathbb{P}_{\Lambda^{m,g}(w)}(\Gamma)$.
Different definitions of the functions $m$ and $g$ lead to different approximating spaces: we will consider in the following the choices Tensor Product (considered so far)(TP), Total Degree (TD), Smolyak (SM) and Hyperbolic Cross (HC), summarized in Table 2.1.

| **Approx. space** | **Sparse grid:** $m, g$ | **Polynomial space** $\Lambda(w)$ |
|---|---|---|
| Tensor Product | $m(i) = i, \quad g(\mathbf{i}) = max_n(i_n - 1) \leq w$ | $\{\mathbf{p} \in \mathbb{N}^N : max_n p_n \leq w\}$ |
| Total Degree | $m(i) = i, \quad g(\mathbf{i}) = \sum_n (i_n - 1) \leq w$ | $\{\mathbf{p} \in \mathbb{N}^N : \sum_n p_n \leq w\}$ |
| Hyperbolic Cross | $m(i) = i, \quad g(\mathbf{i}) = \prod_n (i_n) \leq w + 1$ | $\{\mathbf{p} \in \mathbb{N}^N : \prod_n (p_n + 1) \leq w + 1\}$ |
| Smolyak | $m(i) = \begin{cases} 1, & i = 1 \\ 2^{i-1} + 1, & i > 1 \end{cases}$ $g(\mathbf{i}) = \sum_n (i_n - 1) \leq w$ | $\{\mathbf{p} \in \mathbb{N}^N : \sum_n f(p_n) \leq w\}$ $f(p) = \begin{cases} 0, & p = 0 \\ 1, & p = 1 \\ \lceil log_2(p) \rceil, & p \geq 2 \end{cases}$ |

Table 2.1: Sparse approximation formulas and corresponding polynomial space

The next figures (Fig. 2.1, 2.2, 2.3, 2.4) show the grids obtained with the different rules and a level $w = 5$ with respect to two random variables with uniform distribution over $[-1, 1]$. It should be noticed however that the Smolyak grid (Fig. 2.3) involves a greater number of points since the level $w = 5$ corresponds to a maximum polynomial degree equal to 32 in this case.

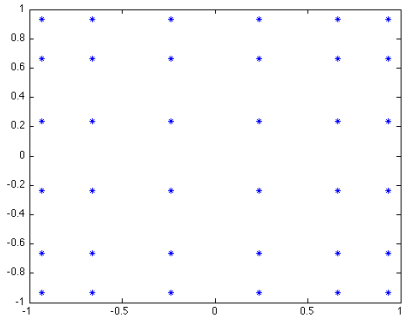

Figure 2.1: TP grid, N=2 w=5

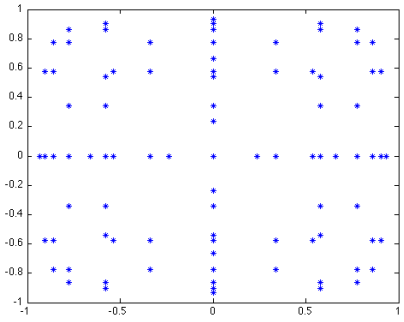

Figure 2.2: TD grid, N=2 w=5

Just like for the tensor product approximation, it is possible to compute the first moment of a function $f$ on a sparse grid as:

$$\mathbb{E}_\rho[f] \simeq \sum_{\mathbf{i} \geq \mathbf{1} : g(\mathbf{i}) \leq w, c(\mathbf{i}) \neq 0} c(\mathbf{i}) \sum_{j_1=1}^{m(i_1)} \cdots \sum_{j_N=1}^{m(i_N)} \boldsymbol{w}_{\boldsymbol{j}}^{\boldsymbol{i}} f(\boldsymbol{\eta}_{\boldsymbol{j}}^{\boldsymbol{i}}) \tag{2.20}$$

where $\boldsymbol{w}_{\boldsymbol{j}}^{\boldsymbol{i}} = [w_{1,j_1}^{(i_1)}, \ldots, w_{N,j_N}^{(i_N)}]$, $\boldsymbol{\eta}_{\boldsymbol{j}}^{\boldsymbol{i}} = [\eta_{1,j_1}^{(i_1)}, \ldots, \eta_{N,j_N}^{(i_N)}]$ and $c(\mathbf{i})$ defined as in 2.16.
In [1] is also provided a useful convergence estimation for the isotropic Smolyak sparse grid: if $u$ is a continuous function on $\Gamma$ taking values in a function space $V$ and analytic with respect

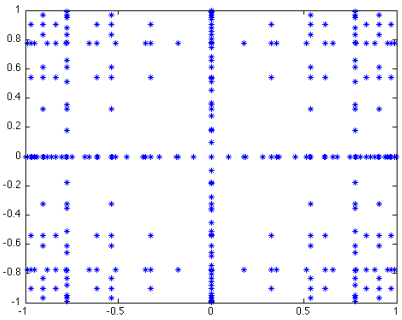

Figure 2.3: SM grid, N=2 w=5

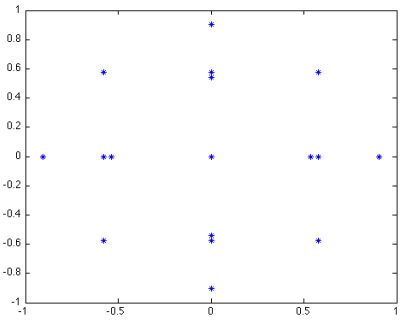

Figure 2.4: HC grid, N=2 w=5

to all random variables $\eta_n$ then the Smolyak formula based on Gaussian abscissas satisfies

$$||u - S_w^{m,g}[u]||_{L^2_\rho(\Gamma;V)} \leq C(N)N_p^{-\mu} \tag{2.21}$$

where $\mu \simeq \frac{C_1}{C_2+\log(N)}$ and $N_p$ still represents the total number of collocation points. As previously stated, if $N \gg \log(N_p)$, then the isotropic full tensor product interpolation converges as $C(N)N_p^{-\tau/N}$. It is quite evident that the Smolyak grid is able to reduce the curse of dimensionality when $N$ is large, since a faster rate of convergence with respect to the number of collocation points has been proved.
Furthermore, if the function to be approximated shows a high dependence on some of the random variables, it is possible to define anisotropic sparse grids. We first introduce the weights $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)$ and we then modify the construction of the sparse grid in order to take into accounts the weights $g(\mathbf{i}; \boldsymbol{\alpha})$ (for example in the Smolyak case we can consider $g(\mathbf{i}; \boldsymbol{\alpha}) = \sum_n \frac{\alpha_n}{\alpha_{min}}(i_n - 1)$). Our anisotropic sparse grid approximation is given by formula (2.15) where the sum is computed over all the multi-indeces $\mathbf{i}$ such that $g(\mathbf{i}; \boldsymbol{\alpha}) \leq w$ (for further details see [1]).

As a final remark that will be useful in our further discussions, we underline that if we consider a function $f$ depending on an $N$-dimensional random vector, given a certain realization of it ($\bar{\boldsymbol{\eta}}$), we can approximate the value $f(\bar{\boldsymbol{\eta}})$ by formula (2.16). It will be sufficient to sum up the values of the interpolant functions related to all the admissible multi-indices $\mathbf{i}$ such that $c(\mathbf{i}) \neq 0$. This will be advantageous whenever the direct evaluation of $f(\boldsymbol{\eta})$ is very expensive (e.g. in the case where evaluating $f(\boldsymbol{\eta})$ implies the solution of a PDE, see next chapter) whereas the evaluation of $S_w^{m,g} f(\boldsymbol{\eta})$ is in general quite cheap once the sparse grid approximation has been formed[1].

[1]In our codes for the HC rules we have considered the rule $\prod_n(i_n) \leq w$, so the HC levels actually starts from $w = 1$

# Chapter 3

# Elliptic and Parabolic PDEs with random coefficients

In this chapter we want to investigate the rate of convergence of the various grids presented in Chapter 2 when used to approximate the solution of PDE with random coefficients, considering both elliptic and a parabolic SPDEs. In this chapter we will not deal with financial problems but this analysis will be useful for applications to the Black and Scholes equation in Chapter 4.

## 3.1 Elliptic SPDEs

We consider a physical domain $D$ and a probability space given by $(\Omega, \mathcal{F}, P)$ where $\Omega$ is the set of all possible outcomes $\omega$, $\mathcal{F}$ is the $\sigma$-algebra of events and $P : \mathcal{F} \to [0, 1]$ is a probability measure. An elliptic stochastic partial differential equation is given by the following problem: find $u \in \bar{D} \times \Omega \to \mathbb{R}$ such that almost surely the following equation holds:

$$\begin{cases} -div(a(\mathbf{x}, \omega)\nabla u(\mathbf{x}, \omega)) = f(\mathbf{x}, \omega) & \mathbf{x} \in D \\ u(\mathbf{x}, \omega) = 0 & \mathbf{x} \in \partial D \times \Omega \end{cases} \tag{3.1}$$

where we are considering deterministic boundary conditions, though the theory can be extended to the case of random boundary conditions as well. We assume that there exist $a_{min} > 0$ and $a_{max} < \infty$ such that:

$$P(a_{min} \leq a(\mathbf{x}, \omega) \leq a_{max} \quad \forall \mathbf{x} \in \bar{D}) = 1 \tag{3.2}$$

and $f$ is square integrable with respect to $P$

$$\int_D \mathbb{E}[f^2] d\mathbf{x} < \infty. \tag{3.3}$$

This excludes the case of a "white noise" forcing term. Moreover, we consider a diffusion coefficient $a(\cdot,\omega)$ and a forcing term $f(\cdot,\omega)$ that can be parametrized by $N$ random variables $\eta_1,\dots,\eta_N$ (usually independent) and are analytic functions of $\eta_1,\dots,\eta_N$. It follows that $u$ depends on $\omega\in\Omega$ only through the value taken by the random vector $\boldsymbol{\eta}$ and we can therefore replace the probability space $(\Omega,\mathcal{F},P)$ with $(\Gamma,\mathcal{B}(\Gamma),\rho(\boldsymbol{\eta})d\boldsymbol{\eta})$ where $\mathcal{B}(\Gamma)$ is the Borel $\sigma$-algebra on $\Gamma$.

We introduce the space $H^1(D)$ of the square integrable functions with square integrable derivatives, the subspace $H^1_0(D)$ of functions with vanishing trace at the boundary and the space $L^2_\rho(\Gamma)$ of square integrable functions on $\Gamma$ with respect to the weight $\rho$. The weak formulation of (3.1) can be written as: find $u\in H^1_0(D)\otimes L^2_\rho(\Gamma)$ such that for every $v\in H^1_0(D)\otimes L^2_\rho(\Gamma)$

$$\mathbb{E}\big[\int_D a(\mathbf{x},\boldsymbol{\eta})\nabla u(\mathbf{x},\boldsymbol{\eta})\nabla v(\mathbf{x},\boldsymbol{\eta})d\mathbf{x}\big]=\mathbb{E}\big[\int_D f(\mathbf{x},\boldsymbol{\eta})v(\mathbf{x},\boldsymbol{\eta})d\mathbf{x}\big]. \tag{3.4}$$

Under the assumption (3.2), from Lax-Milgram theorem we know that there exists a unique solution of the weak problem. This solution then satisfies the stability condition

$$||\nabla u||_{L^2(D)\otimes L^2_\rho(\Gamma)}\leq\frac{C_p}{a_{min}}\big(\int_D\mathbb{E}[f^2]d\mathbf{x}\big)^{\frac{1}{2}} \tag{3.5}$$

where $C_p$ is the Poincaré constant. We recall here a regularity result given in [1]. Assume that condition (3.2) is satisfied, $a(\mathbf{x},\boldsymbol{\eta})$ is analytic in $\Gamma\quad\forall\mathbf{x}\in\bar{D}$ and there exists a N-dimensional vector $\boldsymbol{\gamma}:\ \gamma_n\leq+\infty\ \forall n$ independent of $\boldsymbol{\eta}$ such that

$$\begin{aligned}&||\frac{\partial^k_{\eta_n}a(\mathbf{x},\boldsymbol{\eta})}{a(\mathbf{x},\boldsymbol{\eta})}||_{L^\infty(D)}\leq\gamma_n^k k!\quad\forall n=1,\dots,N\\&\frac{||\partial^k_{\eta_n}f(\boldsymbol{\eta})||_{L^2(D)}}{1+||f(\boldsymbol{\eta})||_{L^2(D)}}\leq\gamma_n^k k!\quad\forall n=1\dots N\end{aligned} \tag{3.6}$$

We will further make some assumptions on $f$ and $\rho$. We introduce a weight $\sigma(\boldsymbol{\eta})=\prod_{n=1}^N\sigma_n(\eta_n)\leq 1$ where

$$\sigma_n(\eta_n)=\begin{cases}1 & \text{if } \Gamma_n \text{ is bounded}\\ e^{-\alpha_n|\eta_n|}\quad\text{for some } \alpha_n>0 & \text{if } \Gamma_n \text{ is unbounded}\end{cases} \tag{3.7}$$

and the functional space

$$\mathbb{C}^0_\sigma(\Gamma;V)=\big\{v:\Gamma\to V,\ v\text{ continuous in }\boldsymbol{\eta}, max_{\boldsymbol{\eta}\in\Gamma}\ ||\sigma(\boldsymbol{\eta})v(\boldsymbol{\eta})||_V<\infty\big\} \tag{3.8}$$

where V is a Banach space of functions on $D$. In what follows we will consider functions

$f \in \mathbb{C}^0_\sigma(\Gamma; L^2(D))$ and probability density functions $\rho$ such that

$$\rho(\boldsymbol{\eta}) \leq C_\rho e^{-\sum_{n=1}^{N}(\delta_n \eta_n)^2} \tag{3.9}$$

for some constant $C_\rho > 0$ and $\delta_n$ strictly positive if $\Gamma_n$ is unbounded and 0 otherwise. So if $\Gamma$ is unbounded, we are focusing on continuous forcing term with at most exponential growth at infinity and Gaussian-type probability distribution functions.
Under the latter assumptions and (3.6) it is proved in [1] that the solution $u(\mathbf{x}, \eta_n)$ as function of $\eta_n$ admits an analytic extension $u(x, z), z \in \mathbb{C}$ in the region of the complex plane

$$\Xi(\Gamma_n; \tau_n) = \{z \in \mathbb{C}, \quad dist(z, \Gamma_n) \leq \tau_n\} \tag{3.10}$$

whith $0 < \tau_n < 1/(2\gamma_n)$.
This result justifies a polynomial interpolation in the probability space.

### 3.1.1 Semi-discrete formulation

We consider a triangulation $\mathcal{T}_h$ in the physical domain $D$ and a finite elements space $V_h(D) \subset H^1_0(D)$, with dimension $N_h$. The semi-discrete problem (3.1) becomes: find $u_h \in V_h(D) \otimes L^2_\rho(\Gamma)$ such that for all $v_h \in V_h(D)$

$$\int_D a(\mathbf{x}, \boldsymbol{\eta}) \nabla u_h(\mathbf{x}, \boldsymbol{\eta}) \nabla v_h(\mathbf{x}) d\mathbf{x} = \int_D f(x, \boldsymbol{\eta}) v_h(\mathbf{x}) d\mathbf{x} \qquad \rho \quad \text{a.e. in } \Gamma. \tag{3.11}$$

For every $\boldsymbol{\eta} \in \Gamma$ the latter problem admits a unique solution, which satisfies the same regularity and stability results that hold for the solution of the continuous problem (3.1).
Let $\{\phi_i(\mathbf{x})\}_{i=1}^{N_h}$ be a Lagrangian basis of $V_h(D)$; we can represent the solution $u_h(\mathbf{x}, \boldsymbol{\eta}) = \sum_{i=1}^{N_h} u_i(\boldsymbol{\eta}) \phi_i(\mathbf{x})$ and then recast the semi-discrete problem (3.11) into its algebraic formulation

$$K(\boldsymbol{\eta}) U(\boldsymbol{\eta}) = F(\boldsymbol{\eta}) \tag{3.12}$$

where $U(\boldsymbol{\eta}) = [u_1(\boldsymbol{\eta}), \ldots, u_{N_h}(\boldsymbol{\eta})]$ is the random vector of nodal values, $K_{ij} = \int_D a(\mathbf{x}, \boldsymbol{\eta}) \nabla \phi_j(\mathbf{x}) \nabla \phi_i(\mathbf{x}) d\mathbf{x}$ is the random stiffness matrix and $F_i = \int_D f(\mathbf{x}, \boldsymbol{\eta}) \phi_i(\mathbf{x}) d\mathbf{x}$, $i = 1 \ldots N_h$ is the random forcing term. Obviously problem (3.12) is derived from (3.11) setting $v_h(\mathbf{x}) = \phi_i(\mathbf{x})$.

### 3.1.2 Fully discrete approximation

If we consider the basis of orthogonal polynomials $\Sigma = span\{\mathbf{Q}_1, \mathbf{Q}_2, \ldots\}$ and we want to approximate the random vector of nodal values $U(\boldsymbol{\eta})$ then $u_i(\boldsymbol{\eta}) \simeq u_i^w(\boldsymbol{\eta}) = \sum_{j=1}^{N_p} u_{ij} \mathbf{Q}_j(\boldsymbol{\eta})$ and the discrete approximation of $u(\mathbf{x}, \boldsymbol{\eta})$ is then given by $u_h^w = \sum_{i=1}^{N_h} \sum_{j=1}^{N_p} u_{ij} \phi_i(\mathbf{x}) \mathbf{Q}_j(\boldsymbol{\eta})$.
The fully discrete problem will then be: find $u_h^w \in V_h(D) \otimes \Sigma$ such that for every $v_h \in V_h(D)$

and every $\mathbf{Z} \in \Sigma$:

$$\mathbb{E}\Big[\int_D a(\mathbf{x},\boldsymbol{\eta})\nabla u_h^w(\mathbf{x},\boldsymbol{\eta})\nabla v_h(\mathbf{x})\mathbf{Z}(\boldsymbol{\eta})\Big] = \mathbb{E}\Big[\int_D f(\mathbf{x},\boldsymbol{\eta})v_h(\mathbf{x})\mathbf{Z}(\boldsymbol{\eta})\Big]. \tag{3.13}$$

Setting the $v_h(\mathbf{x}) = \phi_i(\mathbf{x})$, $\mathbf{Z}(\boldsymbol{\eta}) = \mathbf{Q}_l(\boldsymbol{\eta})$ and considering the last expansion for $u_h^w$ we can derive an algebraic formulation which will lead to the Stochastic Galerkin method (see [3],[2],[18]). This method entails the solution of a certain number of coupled deterministic problems. We will not get into further details beacuse we aim at investigating another efficient technique for solving the SPDE.

In the semi-discrete representation of the solution $u_h(\mathbf{x},\boldsymbol{\eta}) = \sum_{i=1}^{N_h} u_i(\boldsymbol{\eta})\phi_i(\mathbf{x})$ we replace each nodal value with its interpolant as in (2.12): $u_i(\boldsymbol{\eta}) \simeq \sum_{j=1}^{N_p} u_i(\boldsymbol{\eta}_j)l_j(\boldsymbol{\eta})$. We are considering here the TP rule in Table 2.1. Then the fully discrete solution has the following expansion $u_h^w(\mathbf{x},\boldsymbol{\eta}) = \sum_{i=1}^{N_h}\sum_{j=1}^{N_p} u_i(\boldsymbol{\eta}_j)\phi_i(\mathbf{x})l_j(\boldsymbol{\eta}) = \sum_{j=1}^{N_p} u_h(\mathbf{x},\boldsymbol{\eta}_j)l_j(\boldsymbol{\eta})$.
Therefore to obtain the fully discretized solution we have to evaluate the semi-discrete solution $u_h(\mathbf{x},\boldsymbol{\eta})$ in the $N_p$ collocation points $\boldsymbol{\eta}_j$. It follows that we have to solve $N_p$ uncoupled deterministic problems: $\forall j = 1 \dots N_p$ find $u_h(\mathbf{x},\boldsymbol{\eta}_j) \in V_h(D)$ such that:

$$\int_D a(\mathbf{x},\boldsymbol{\eta}_j)\nabla u_h(\mathbf{x},\boldsymbol{\eta}_j)\nabla v_h(\mathbf{x})d\mathbf{x} = \int_D f(\mathbf{x},\boldsymbol{\eta}_j)v_h(\mathbf{x})d\mathbf{x} \qquad \forall v_h(\mathbf{x}) \in V_h(D). \tag{3.14}$$

This technique is named Stochastic Collocation method ([1],[2]) since we are collocating a certain SPDE in the probability knots used to build the interpolant function.

### 3.1.3 Rates of convergence

The convergence result for $||u - u_h^w||_{L^2_\rho(\Gamma)\otimes H^1_0(D)}$ is based on the triangular inequality since $u - u_h^w = (u - u_h) + (u_h - u_h^w)$, so we can split the total error in two errors: one involving the physical space approximation the other the probability approximation. It can be noticed that $u_h$ is the orthogonal projection of $u$ onto the subspace $V_h \otimes L^2_\rho(\Gamma)$ with respect to the inner product $\int_{\Gamma\times D} \rho a|\nabla \cdot|^2$, so from Galerkin optimality it follows

$$\begin{aligned} ||u - u_h||_{L^2_\rho(\Gamma)\otimes H^1_0(D)} \leq & \frac{1}{\sqrt{a_{min}}}\left(\int_{\Gamma\times D} \rho a\nabla|u - u_h|^2\right)^{1/2} \\ & \leq \frac{1}{\sqrt{a_{min}}} \inf_{v\in V_h\otimes L^2_\rho(\Gamma)} \left(\int_{\Gamma\times D} \rho a\nabla|u - v|^2\right)^{1/2}. \end{aligned} \tag{3.15}$$

The term $||u_h - u_h^w||_{L^2_\rho(\Gamma)\otimes H^1_0(D)}$ represents the approximation error in probability, which can be estimated as well. In particular we quote the following result presented in ([1]) related to the Tensor Product rule in Table 2.1.

**Teorema 3.1.1.**
*If $f \in \mathbb{C}^0_\sigma(\Gamma; L^2(D))$ and $a$ satisfies the assumptions (3.6) and (3.9) then there exist positive*

*constants* $r_n, n = 1 \dots N$ *independent of* $h$ *and* $p_n$*, such that*

$$\begin{aligned} ||u - u_h^w||_{L^2_\rho(\Gamma)\otimes H^1_0} \leq & \frac{1}{\sqrt{a_{min}}} \inf_{v \in V_h \otimes L^2_\rho(\Gamma)} \left( \int_{\Gamma \times D} \rho a \nabla |u - v|^2 \right)^{1/2} \\ & + C \sum_{n=1}^{N} \beta_n(p_n) exp\{-r_n p_n^{\theta_n}\} \end{aligned} \tag{3.16}$$

*where, if* $\Gamma_n$ *is bounded*

$$\begin{cases} \theta_n = \beta_n = 1 \\ r_n = log\big[\frac{2\tau_n}{|\Gamma_n|}\big(1 + \sqrt{1 + \frac{|\Gamma_n|^2}{4\tau_n}}^{\,2}\big)\big] \end{cases} \tag{3.17}$$

*else*

$$\begin{cases} \theta_n = \frac{1}{2}, \quad \beta_n = O(\sqrt{p_n}) \\ r_n = \tau_n \delta_n \end{cases} \tag{3.18}$$

*where* $\tau_n$ *as in (3.10) and* $\delta_n$ *as in (3.9).*

In next section we will see how to apply (3.16). It has to be noticed that from (3.16) it is possible to derive a convergence result for the error in the mean value measured in $L^2(D)$ or $H^1(D)$ norms and for the error in the second moment measured in the $L^1(D)$ norm, since:

$$\begin{aligned} ||E[u - u_h^w]||_{V(D)} &\leq ||u - u_h^w||_{L^2_\rho(\Gamma)\otimes V(D)} \\ ||E[u^2 - (u_h^w)^2]||_{L^1(D)} &\leq C||u - u_h^w||_{L^2_\rho(\Gamma)\otimes L^2(D)} \end{aligned} \tag{3.19}$$

where $V(D) = L^2(D)$ or $V(D) = H^1(D)$.

### 3.1.4 Sparse grids and elliptic PDE

The results presented so far in this chapter are related to a full Tensor Product space but the same considerations can be repeated for Sparse Grid approximations. In the latter case, using the typical error splitting result, it has been shown in [5] that the probability approximation error for the isotropic Smolyak grid is the one presented in the previous chapter (formula 2.21).

## 3.2 Examples of Elliptic SPDEs

We examine an elliptic problem in one physical dimension:

$$\begin{cases} -(a(x,\boldsymbol{\eta})u(x,\boldsymbol{\eta})')' = f(x) \quad x \in D = (0,1), \qquad \eta \in \Gamma \\ u(0,\boldsymbol{\eta}) = u(1,\boldsymbol{\eta}) = 0 \qquad \boldsymbol{\eta} \in \Gamma \end{cases} \tag{3.20}$$

where $a(x,\boldsymbol{\eta})$ is a function of space $x$ and random variables $\boldsymbol{\eta}$.

If we consider a finite elements approximation with degree $r$ in the physical space, we can estimate the term $\inf_{v \in V_h \otimes L^2_\rho} \left( \int_{\Gamma \times D} \rho a \nabla |u-v|^2 \right)^{1/2}$ in (3.16) with $C_1 h^r \mathbb{E}[||u||^2_{H^{r+1}}]^{1/2}$ under the assumption (3.2). So if we are considering a vector of random variables with bounded support and a Tensor Product approximation in probability our convergence result is

$$\mathbb{E}[||u - u_h^w||^2_{H^1_0(D)}]^{\frac{1}{2}} \leq C_1 h^r \mathbb{E}[||u||^2_{H^{r+1}}]^{\frac{1}{2}} + C_2 e^{-C_3 w}. \tag{3.21}$$

Similarly, since the error splitting holds, we can derive a convergence rate in the norm $||\cdot||_{L^2(D)\otimes L^2_\rho(\Gamma)}$:

$$\mathbb{E}[||u - u_h^w||^2_{L^2(D)}]^{\frac{1}{2}} \leq C_1 h^{r+1} \mathbb{E}[||u||^2_{H^{r+1}}]^{\frac{1}{2}} + C_2 e^{-C_3 w}. \tag{3.22}$$

Moreover, with respect to the number of collocation points we can derive an algebraic convergence for the Smolyak grid and a sub exponential convergence for the Tensor Product grid (as discussed in Section 2.1), coupled with the usual rate of convergence of the finite elements method.

### 3.2.1 Numerical results

We consider problem (3.20) where $a(x,\boldsymbol{\eta}) = (1+0.1\eta_1+0.5\eta_2)$ and $\eta_1$,$\eta_2$ are random variables with uniform distribution over [-1,1]. If $f(x) = 1$ then the exact solution of the problem is given by $u(x,\boldsymbol{\eta}) = \frac{x(1-x)}{2a(x,\boldsymbol{\eta})}$.

Our objective is to estimate $||\mathbb{E}[u] - \mathbb{E}[u^w]||_{L^2(D)}$; $\mathbb{E}[u]$ is known in every knot of the physical space and is equal to $\frac{x(1-x)}{2}\mathbb{E}[\frac{1}{a(x,\boldsymbol{\eta})}]$. We are also able to evaluate the right term ($\mathbb{E}[u^w]$): given a level $w$ and the corresponding probability grid, we can compute the expected value of the solution in every physical knot. Considering a grid step equal to $\frac{1}{10i}$ $i = 4,..,8$ and various levels $w$ with the different rules we get the following graphs representing the $L^2$ error (3.22)

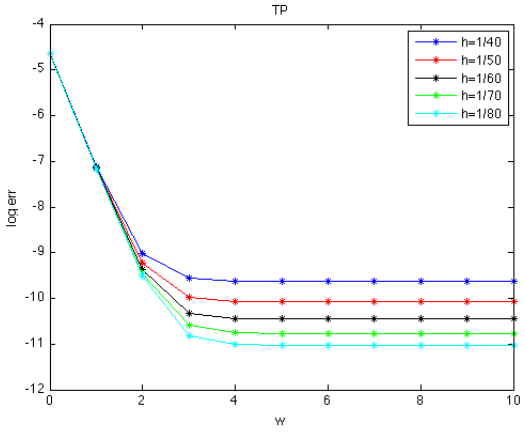


Figure 3.1: First moment convergence TP grid: $||\mathbb{E}[u] - \mathbb{E}[u^w]||_{L^2(D)}$

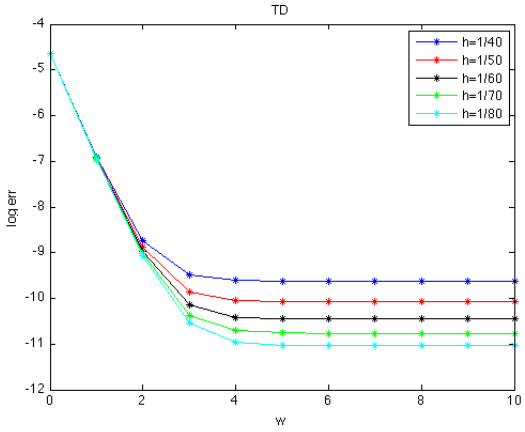


Figure 3.2: First moment convergence TD grid: $||\mathbb{E}[u] - \mathbb{E}[u^w]||_{L^2(D)}$

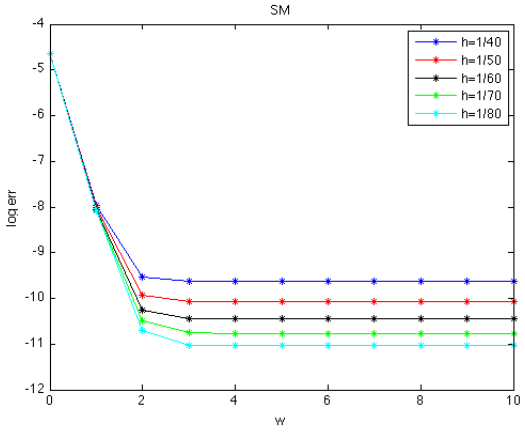


Figure 3.3: First moment convergence SM grid: $||\mathbb{E}[u] - \mathbb{E}[u^w]||_{L^2(D)}$

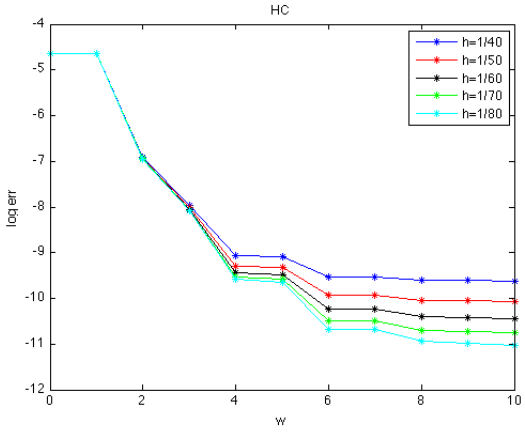


Figure 3.4: First moment convergence HC grid: $||\mathbb{E}[u] - \mathbb{E}[u^w]||_{L^2(D)}$

These graphs show that the probability convergence for the different grids flattens when the space approximation error dominates. In order to eliminate the latter, since we are using linear finite elements, it is possibile to consider a linear interpolation for the exact solution to estimate $\mathbb{E}[||u_h - u_h^w||^2_{L^2(D)}]^{\frac{1}{2}}$. The next two graphs represent this error plotted against the number of collocation points involved in the computation (Fig. 3.5) and the maximum polinomial degree (Fig. 3.6). It should be noticed that for TP, TD and HC rules the maximum polynomial degree of the polynomial approximation coincides with the level $w$. In case of SM grid the maximum polynomial degree is equal to $2^w$. A linear convergence (on a semi-logarithmic scale) is observable just for the TP, TD and HC rules (almost linear in the last case). In the Smolyak case the convergence seems to be algebric rather than exponential. Fig. 3.5 shows a more than algebraic convergence with respect to the number of evaluation

points for all the methods, which is faster for TP and SM methods. This is not surprising since the sparse grid methods become effective if the number of variables is large.

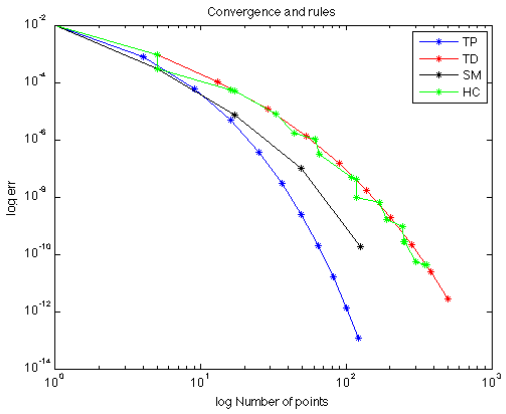


Figure 3.5: First moment convergence vs number of points: $||\mathbb{E}[u] - \mathbb{E}[u^w]||_{L^2(D)}$

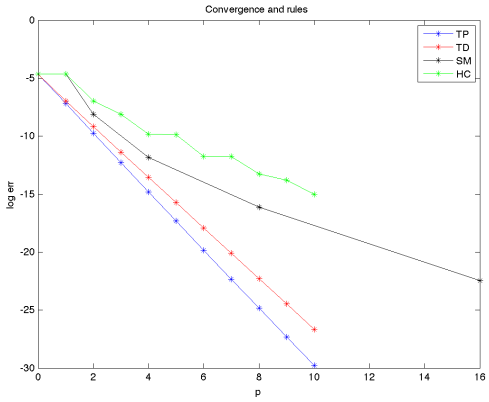


Figure 3.6: First moment convergence vs polynomial degree: $||\mathbb{E}[u] - \mathbb{E}[u^w]||_{L^2(D)}$

### 3.2.2 Strong convergence

Let us consider the problem (3.20) where $a(x, \boldsymbol{\eta}) = e^x(1 + 0.1\eta_1 + 0.5\eta_2)$ and $\eta_1, \eta_2$ are two random variables with uniform distribution over [-1,1]. If we define $f(x) = e^x(x + 0.5)$, the solution is given by $u(x, \eta) = \frac{x(1-x)}{2a(x,\boldsymbol{\eta})}$.
With the aim of obtaining an estimation for $\mathbb{E}[||u - u_h^w||^2_{L^2(D)}]^{\frac{1}{2}}$, having set a level $w$, we solve numerically the equation for every physical knot and for every quadrature knot. We then create $M = 2000$ random samples with the same distribution of the two random variables and evaluate the solution in every physical knot using an interpolation formula in the probability space, because the solution is represented on an interpolatory basis as discussed in the previous chapter. Since we know the analytical solution for every generated random sample, we are able to get an estimation of the strong error:

$$\mathbb{E}[||u - u_h^w||^2_{L^2(D)}]^{\frac{1}{2}} \simeq \big[\frac{1}{M}\sum_{i=1}^{M} ||u_i - u_{hi}^w||^2\big]^{\frac{1}{2}}. \tag{3.23}$$

We get the following plots considering different grid sizes $(\frac{1}{10}, \frac{1}{20}, \frac{1}{40}, \frac{1}{80})$:

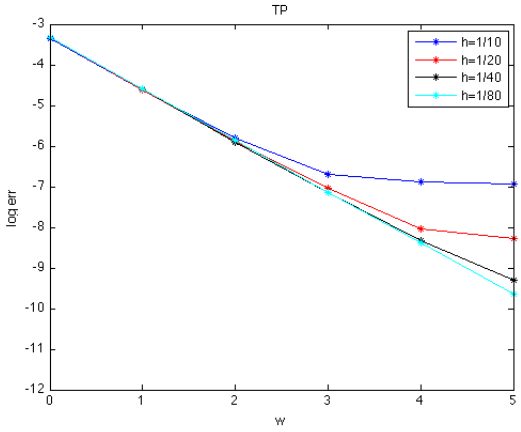


Figure 3.7: Strong error convergence TP grid: estimation of $\mathbb{E}[||u - u_h^w||^2_{L^2(D)}]^{\frac{1}{2}}$

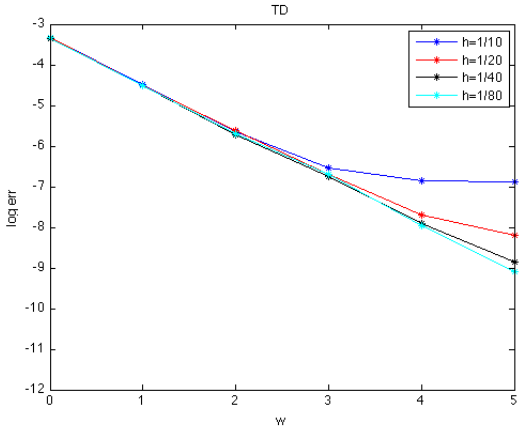


Figure 3.8: Strong error convergence TD grid: estimation of $\mathbb{E}[||u - u_h^w||^2_{L^2(D)}]^{\frac{1}{2}}$

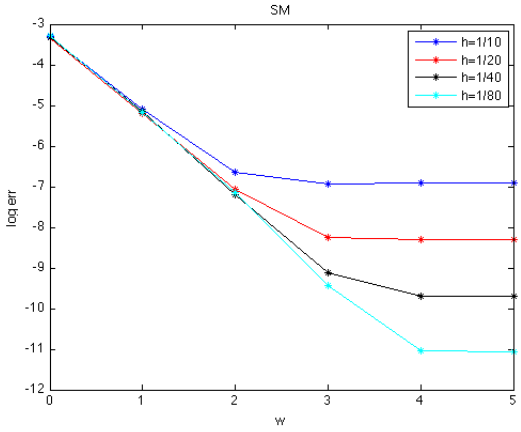


Figure 3.9: Strong error convergence SM grid: estimation of $\mathbb{E}[||u - u_h^w||^2_{L^2(D)}]^{\frac{1}{2}}$

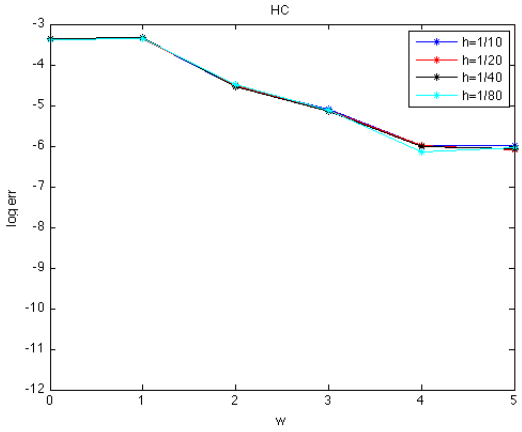


Figure 3.10: Strong error convergence HC grid: estimation of $\mathbb{E}[||u - u_h^w||^2_{L^2(D)}]^{\frac{1}{2}}$

Considering different values of level, we obtain an estimation of the convergence rate with respect to the number of collocation points

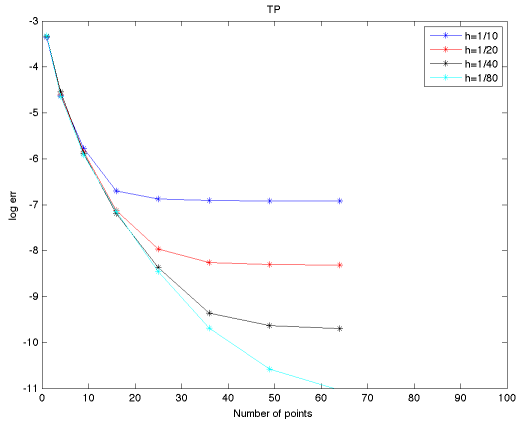


Figure 3.11: Strong error convergence vs number of points TP grid: estimation of $\mathbb{E}[||u-u_h^w||^2_{L^2(D)}]^{\frac{1}{2}}$

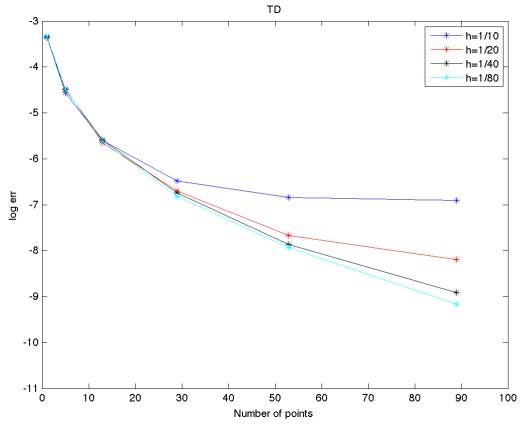


Figure 3.12: Strong error convergence vs number of points TD grid: estimation of $\mathbb{E}[||u-u_h^w||^2_{L^2(D)}]^{\frac{1}{2}}$

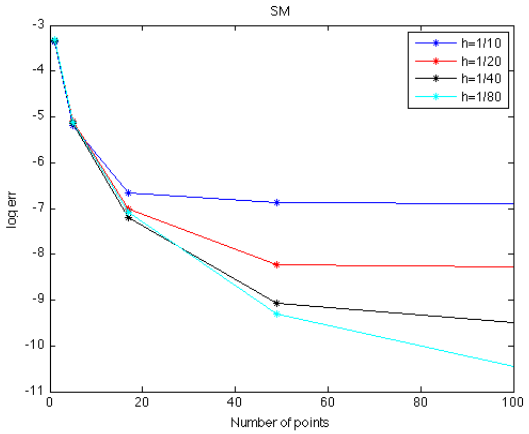


Figure 3.13: Strong error convergence vs number of points SM grid: estimation of $\mathbb{E}[||u-u_h^w||^2_{L^2(D)}]^{\frac{1}{2}}$

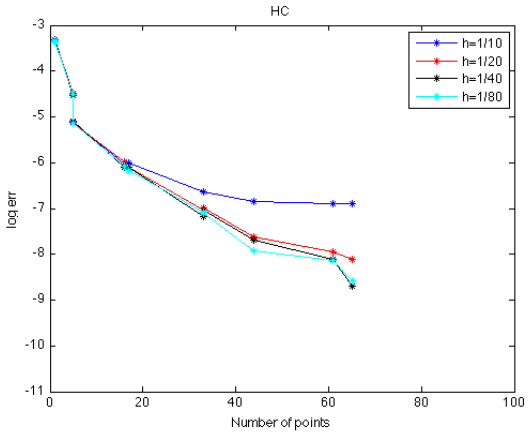


Figure 3.14: Strong error convergence vs number of points HC grid: estimation of $\mathbb{E}[||u-u_h^w||^2_{L^2(D)}]^{\frac{1}{2}}$

### 3.2.3 Finite Elements error

In a similar way, if we are interested in evaluating the space discretization error we set the level $w = 1 \dots 5$ and vary the spacial step getting the trends shown in Fig. 3.15, Fig. 3.16, Fig. 3.17 and Fig. 3.18.

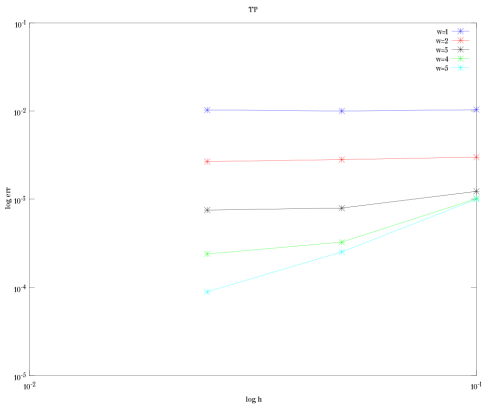


Figure 3.15: Finite Elements error TP grid: estimation of $\mathbb{E}[||u - u_h^w||^2_{L^2(D)}]^{\frac{1}{2}}$

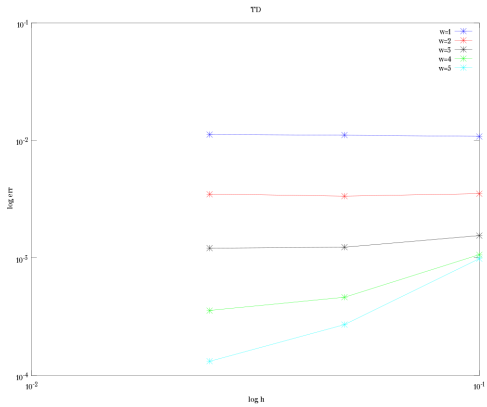


Figure 3.16: Finite Elements error TD grid: estimation of $\mathbb{E}[||u - u_h^w||^2_{L^2(D)}]^{\frac{1}{2}}$

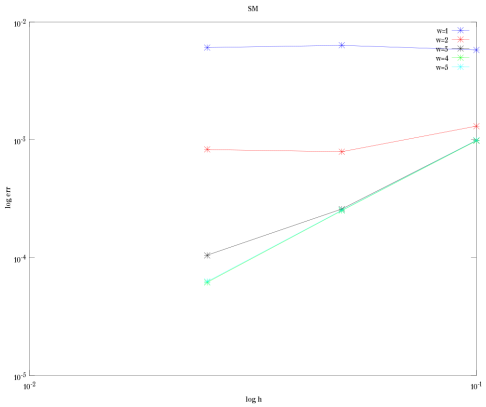


Figure 3.17: Finite Elements error SM grid: estimation of $\mathbb{E}[||u - u_h^w||^2_{L^2(D)}]^{\frac{1}{2}}$

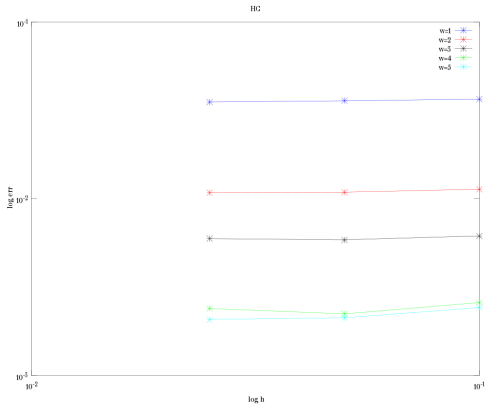


Figure 3.18: Finite Elements error HC grid: estimation of $\mathbb{E}[||u - u_h^w||^2_{L^2(D)}]^{\frac{1}{2}}$

We observe exactly the theoretical convergence rate related to the finite elements approximation until the probability error becomes dominant and the space error is constant.

After these analyses we can conclude that the TP rule represents the best choice since the number of variables is very small. In particular from Fig. 3.5 and 3.7 the best convergence rate with respect to the number of points is observable for the TP considering both the first moment $||\mathbb{E}[u] - \mathbb{E}[u^w]||_{L^2(D)}$ and the strong error $\mathbb{E}[||u - u_h^w||^2_{L^2(D)}]^{\frac{1}{2}}$. We just recall that the sparse grids show their effectiveness in alleviating the curse of dimensionality, therefore for high dimensional problems they are supposed to outperform the full tensioral grids.

## 3.3 Parabolic SPDE

The well-posedness for parabolic SPDEs can be studied similarly to the elliptic case. Given a physical domain $D$ and a probability space $(\Omega, \mathcal{F}, P)$, a stochastic parabolic partial differential equation is defined as: find $u \in \bar{D} \times [0,T] \times \Omega \to \mathbb{R}$ such that almost surely the following equation holds:

$$\begin{cases} \partial_t u(\mathbf{x},t,\omega) - div(a(\mathbf{x},\omega)\nabla u(\mathbf{x},t,\omega)) = f(x,t,\omega) \quad \text{in} \quad D \times (0,T] \times \Omega \\ u(\mathbf{x},t,\omega) = 0 \quad \text{on} \quad \partial D \times [0,T] \times \Omega \\ u(\mathbf{x},0,\omega) = u_0 \quad \text{on} \quad \partial D \times \Omega \end{cases} \tag{3.24}$$

We make the assumption (3.2) on $a$ and we also assume that $f$ is square integrable:

$$\int_\Omega \int_{D\times[0,T]} f^2(\mathbf{x},t,\omega) d\mathbf{x} dt dP < \infty. \tag{3.25}$$

Again this assumption rules out the case of space-time white noise. As in the previous section, if $a$ and $f$ depend on a finite number of random variables, we can replace the probability space $(\Omega, \mathcal{F}, P)$ with $(\Gamma, \mathcal{B}(\Gamma), \rho(\boldsymbol{\eta})d\boldsymbol{\eta})$. The weak formulation of problem (3.24) becomes: find $u \in L^2([0,T]; H_0^1(D) \otimes L^2_\rho(\Omega))$ such that $\partial_t u \in L^2([0,T]; H^{-1}(D) \otimes L^2_\rho(\Omega))$, $u = u_0$ for $t = 0$ and $\forall v \in H_0^1(D) \otimes L^2_\rho(\Omega)$ a.e. in $(0,T]$:

$$\mathbb{E}\big[\int_D \partial_t u v d\mathbf{x}\big] + \mathbb{E}\big[\int_D a\nabla u \nabla v d\mathbf{x}\big] = \mathbb{E}\big[\int_D f v d\mathbf{x}\big]. \tag{3.26}$$

From assumptions (3.2) and (3.25) follows that there exists a unique solution $u \in L^2([0,T]; H_0^1(D) \times L^2_\rho(\Omega))$ which satisfies the energy estimation:

$$||u(T)||^2_{L^2(D)\otimes L^2_\rho(\Gamma)} + a_{min}||u||^2_{L^2([0,T];H_0^1(D))\otimes L^2_\rho(\Omega)} \leq \frac{C_p^2}{a_{min}}||f||^2_{L^2([0,T];L^2(D))\otimes L^2_\rho(\Omega)} + ||u_0||^2_{L^2(D)} \tag{3.27}$$

where $C_p$ is the Poincaré constant.
Just like in section (3.1), if we make the same assumptions on the diffusion coefficient $a$, it is possible to prove that $u(\mathbf{x},\cdot)$ admits an analytic extension in the complex plane $u(z,\cdot)$, $z \in \mathbb{C}$ with respect to each random variable (see [2]).

### 3.3.1 Stochastic Galerkin-FEM

If we consider a finite dimensional approximating space $V_h^w \subset H_0^1(D) \otimes L^2_\rho(\Gamma)$ we can define a semi-discrete formulation of (3.26): find $u_h^w \in V_h^w$ such that for every $t \in [0,T]$ and every $v_h^w \in V_h^w$

$$\mathbb{E}\big[\int_D \partial_t u_h^w(t) v_h^w d\mathbf{x}\big] + \mathbb{E}\big[\int_D a\nabla u_h^w(t) \nabla v_h^w d\mathbf{x}\big] = \mathbb{E}\big[\int_D f(t) v_h^w d\mathbf{x}\big] \tag{3.28}$$

with initial condition $u_h^w(0) = u_{0,h}^w$ where $u_{0,h}$ is the projection of $u_0$ on $V_h^w$. As a finite dimensional space $V_h^w$ we will consider a finite elements space in the physical domain and a polynomial space ($\Sigma$) given by the span of orthogonal polynomials $\Sigma = span\{\mathbf{Q}_i\}$ for certain multi-indeces $i \in \mathbb{N}^N$. Let $\{\phi_j\}_{j=1}^{N_h}$ and $\{\mathbf{Q}_k\}$ be the basis functions of the finite dimensional space $V_h^w$ then every function $u$ can be represented as $u_h^w(\mathbf{x}, t, \boldsymbol{\eta}) = \sum_{j=1}^{N_h} \sum_{k=1}^{N_p} u_{jk}(t)\phi_j(\mathbf{x})\mathbf{Q}_k(\boldsymbol{\eta})$. If we substitute the latter expansion in (3.28) and the basis functions of $V_h^w$ as test functions in the semi-discrete formulation, we get that the $u_{jk}$ satisfy a sistem of ODEs.
We will not get into further details since this choice will lead to the Stochastic Galerkin method while we will solve SPDEs with the Stochastic Collocation technique, instead.

### 3.3.2 Stochastic Collocation-FEM

In this framework our approximating space $V_h^w \subset H_0^1(D) \otimes L_\rho^2(\Gamma)$ consists in the finite element space plus a polynomial space given by the interpolating functions $l_j,\ j = 1 \dots N_p$ defined in Chapter 2. Therefore the semi-discrete expansion for the solution $u$ (since we have not discretized the time dependence) is given by $u_h^w(\mathbf{x}, t, \boldsymbol{\eta}) = \sum_{k=1}^{N_h} \sum_{j=1}^{N_p} u_k(t, \boldsymbol{\eta}_j)\phi_i(\mathbf{x})l_j(\boldsymbol{\eta}) = \sum_{j=1}^{N_p} u_h(\mathbf{x}, t, \boldsymbol{\eta}_j)l_j(\boldsymbol{\eta})$ (we are considering the TP rule with level $w$).
As explained for elliptic SPDEs, Stochastic Collocation method entails the solution of a certain number of deterministic uncoupled problems so we collocate the equation on each collocation knot. Thus for every $\boldsymbol{\eta}_j$ we get a deterministic solution that we will call $u_h(\mathbf{x}, t, \boldsymbol{\eta}_j) = \sum_{k=1}^{N_h} u_k(t, \boldsymbol{\eta}_j)\phi_k(\mathbf{x}),\ j = 1 \dots N_p$, where all the coefficients $u_k(t, \boldsymbol{\eta}_j)$ satisfy a system of ODEs:

$$\sum_{k=1}^{N_h} \dot{u}_k^j(t) \int_D \phi_i \phi_k d\mathbf{x} + \sum_{k=1}^{N_h} u_k^j(t) a^j(\phi_i, \phi_k) = F(\phi_i) \qquad i = 1 \dots N_h \tag{3.29}$$

where $F(\phi_i) = \int_D f(\mathbf{x}, t, \boldsymbol{\eta}_j)\phi_i(\mathbf{x})d\mathbf{x}$ and $a^j(\phi_i, \phi_k) = \int_D a(\mathbf{x}, \boldsymbol{\eta}_j)\phi_k(\mathbf{x})\phi_i(\mathbf{x})d\mathbf{x}$ or in a matricial form

$$M\dot{\mathbf{u}}^j(t) + K^j\mathbf{u}^j(t) = \mathbf{f}^j(t) \tag{3.30}$$

where $\mathbf{u}$ is the vector of degrees of freedom, $M$ is the mass matrix, $K^j$ the stiffness matrix computed in the collocation point $\boldsymbol{\eta}_j$ and $\mathbf{f}$ is the forcing term. Having discretized the time interval $[0, T]$ into $L$ subintervals $[t_z, t_{z+1}],\ z = 0 \dots L-1$ of length $\Delta t$, we can approximate the time derivative in (3.30) with an incremental ratio depending on a parameter $0 \leq \theta \leq 1$ and obtain the $\theta$-method

$$M\frac{\mathbf{u}^j(t_{z+1}) - \mathbf{u}^j(t_z)}{\Delta t} + K^j[\theta\mathbf{u}^j(t_{z+1}) + (1-\theta)\mathbf{u}^j(t_z)] = \theta\mathbf{f}^{j+1}(t_{z+1}) + (1-\theta)\mathbf{f}^j(t_z). \tag{3.31}$$

Letting $\theta$ vary we get the usual methods of Backward Euler, Forward Euler and Crank-Nicolson. In our simulations we will use the Crank-Nicolson scheme or Rannacher scheme (see Appendix for further details) to get a fully-discrete approximation.

### 3.3.3 Rates of convergence

Also in this case, the convergence rate estimation is based on an error splitting result. If we call $u$ the exact solution of (3.24), $u^w_{h,\Delta t}$ the fully-discrete approximation and $u_{h,\Delta t}$ the semi-discrete one involving a physical space and time approximation then $||u - u^w_{h,\Delta t}|| \leq ||u - u^w_{h,\Delta t}|| + ||u_{h,\Delta t} - u^w_{h,\Delta t}||$. In this section we present an estimation for $||u_{h,\Delta t} - u^w_{h,\Delta t}||$ (for the sake of the simplicity we will denote $u^* = u^w_{h,\Delta t}$). In particular as discussed in [2], if $\Gamma$ is bounded then the following probability approximation error with respect to $w$ (the maximum degree of the polynomial space) holds:

$$\mathbb{E}[||(u_{h,\Delta t} - u^*)||^2_{L^2(D)}(t_L)] + a_{min}\mathbb{E}[\sum_{z=1}^{L} \Delta t ||u_{h,\Delta t}(t_z) - u^*(t_z)||^2_{H^1_0(D)}] \leq C(\tau_{min}) \sum_{n=1}^{N} e^{-\tau_n w} \tag{3.32}$$

where $\tau_n$ is still related to the distance of the analyticity region of $u$ in the complex plane to the nearest singularity.

### 3.3.4 Sparse grids and parabolic SPDEs

As for the elliptic SPDEs we can repeat the same construction also for sparse grids. In this case some results have been derived for the Smolyak rule in [2])

$$\mathbb{E}[||(u_{h,\Delta t} - S^{m,g}_w[u_{h,\Delta t}])(t_L)||^2_{L^2(D)}] + a_{min}\mathbb{E}[\sum_{z=1}^{L} \Delta t ||u_{h,\Delta t} - S^{m,g}_w[u_{h,\Delta t}]||^2_{H^1_0(D)}(t_z)] \leq$$
$$C(\tau_{min}, a_{min}, a_{max}, f, C_p, N) N_p^{\frac{\tilde{C}}{1+\log(2N)}} \tag{3.33}$$

with $S^{m,g}_w$ defined as in (2.16).

## 3.4 Example of Parabolic SPDE

We consider the following parabolic equation in one physical dimension

$$\begin{cases} u_t - (a(x,\boldsymbol{\eta})u(x,t,\boldsymbol{\eta})')' = f(x,t,\boldsymbol{\eta}) \quad x \in D = (0,1), \qquad \eta \in \Gamma \qquad t \in [0,T] \\ u(0,t,\boldsymbol{\eta}) = u(1,t,\boldsymbol{\eta}) = 0 \qquad \boldsymbol{\eta} \in \Gamma \\ u(x,0,\boldsymbol{\eta}) = \frac{sin(\pi x)}{(1+0.1\eta_1+0.5\eta_2)^2} \end{cases} \tag{3.34}$$

where $a(x,\boldsymbol{\eta}) = 2(1+0.1\eta_1+0.5\eta_2)$ and $\eta_1$, $\eta_2$ are random variables with uniform distribution over [-1,1]. If we set $f(x,t,\boldsymbol{\eta}) = 2sin(\pi x)e^{-2t}\big[\frac{\pi^2}{(1+0.1\eta_1+0.5\eta_2)} - \frac{1}{(1+0.1\eta_1+0.5\eta_2)^2}\big]$, then the solution is given by $u(x,t,\boldsymbol{\eta}) = \frac{sin(x)e^{-2t}}{(1+0.1\eta_1+0.5\eta_2)^2}$.
The theoretical convergence estimate for a parabolic SPDEs is very similar to that presented for elliptic ones and is based on an error splitting result. Having discretized the time interval

$[0,T]$ into $L$ intervals $[t_z, t_{z+1}]$, $\quad z = 0 \dots L-1$ of length $\Delta t$, we solve every deterministic problem related to each collocation point using a $\theta-method$ scheme (3.31). If $u$ is the solution of problem (3.24) and $u^w_{h,\Delta t}$ the solution of the fully discrete problem (3.30) then from (3.32) and from the classical results involving finite elements and $\theta$-method (see[19]) follows:

$$\mathbb{E}[||u(T) - u^w_{h,\Delta t}(t_L)||^2_{L^2(D)}] + a_{min}\Delta t\mathbb{E}[\sum_{z=1}^{L} ||u(t_z) - u^w_{h,\Delta t}(t_z)||^2_{H^1(D)}] \leq$$
$$C_1(h^{2r} + \Delta t^q) + C_2 e^{-\tau w} \tag{3.35}$$

where $q$ depends upon the time discretization scheme ($q = 2$ for Backward Euler, $q = 4$ for Crank Nicolson). Formula (3.35) holds for a Tensor Product grid where $w$ is the maximum polynomial degree.

### 3.4.1 Numerical Results

To get an estimation of the left term of (3.35), having set a level $w$, we solve numerically the equation in every knot of the physical and temporal grid and every collocation knot. After generating $M = 1000$ random samples with the same distribution of $\eta_1$ and $\eta_2$, we evaluate the solution in every physical and temporal knot using an interpolation in probability. Since we know the exact solution for every random sample, we are able to get an estimation for the convergence rate by replacing the mean values in (3.35) with sample means based on the M samples. To perform our computation, we use linear finite elements with spatial steps equal to $\frac{1}{10}, \frac{1}{20}, \frac{1}{40}$ and the Crank Nicolson scheme with temporal steps $\frac{1}{10}, \frac{1}{15}, \frac{1}{20}$. With the different rules we get the following results:

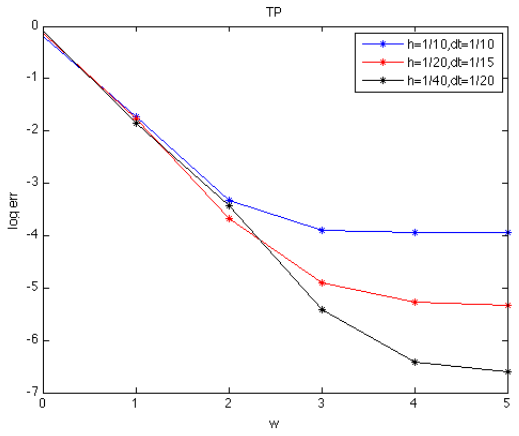


Figure 3.19: Strong error convergence TP grid: estimation of $\mathbb{E}[||u(T) - u^w_{h,\Delta t}(t_L)||^2_{L^2(D)}] + a_{min}\Delta t\mathbb{E}[\sum_{k=1}^{L} ||u(t_k) - u^w_{h,\Delta t}(t_k)||^2_{H^1(D)}]$

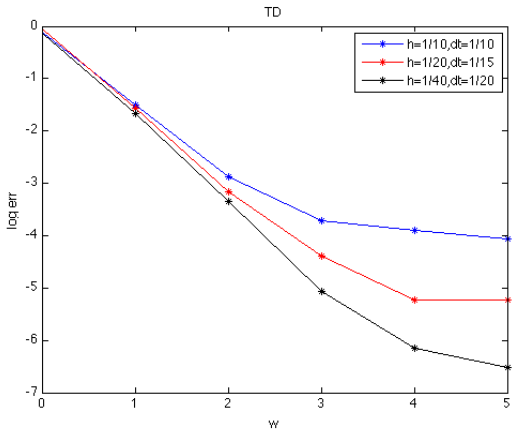


Figure 3.20: Strong error convergence TD grid: estimation of $\mathbb{E}[||u(T) - u^w_{h,\Delta t}(t_L)||^2_{L^2(D)}] + a_{min}\Delta t\mathbb{E}[\sum_{k=1}^{L} ||u(t_k) - u^w_{h,\Delta t}(t_k)||^2_{H^1(D)}]$

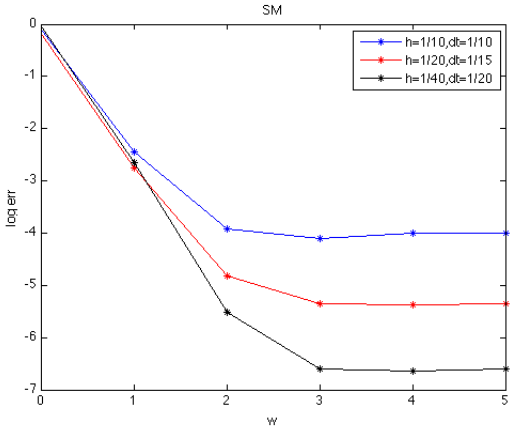


Figure 3.21: Strong error convergence SM grid: estimation of $\mathbb{E}[||u(T) - u^w_{h,\Delta t}(t_L)||^2_{L^2(D)}] + a_{min}\Delta t\mathbb{E}[\sum_{k=1}^{L} ||u(t_k) - u^w_{h,\Delta t}(t_k)||^2_{H^1(D)}]$

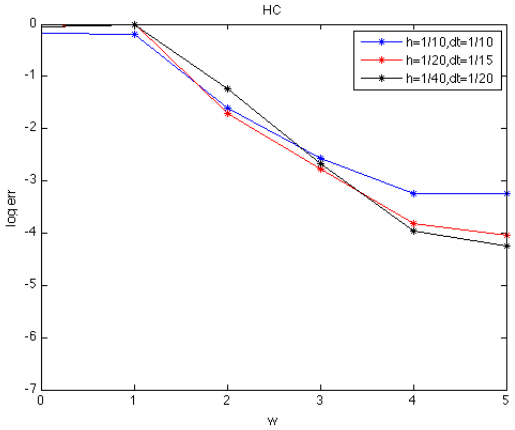


Figure 3.22: Strong error convergence HC grid: estimation of $\mathbb{E}[||u(T) - u^w_{h,\Delta t}(t_L)||^2_{L^2(D)}] + a_{min}\Delta t\mathbb{E}[\sum_{k=1}^{L} ||u(t_k) - u^w_{h,\Delta t}(t_k)||^2_{H^1(D)}]$

As for elliptic SPDEs, an almost exponential convergence with respect to $w$ is observable until the time/physical space approximation error becomes dominant.

# Chapter 4

# Black and Scholes equation with uncertain parameters

In this Chapter we will apply the results presented in the previous one to the parabolic Black and Scholes PDE under the assumption of random parameters. This represents a classical approach while the innovative part of this thesis consists in Chapter 5 and 6.
We will first present the Black and Scholes equation and its extension to forex contingent claims (Section 4.1) and then analyze the price sensitivity to the parameters using the Stochastic Collocation technique (Section 4.2). In Section 4.3 and 4.4 we will make a comparison with Monte Carlo method.

## 4.1 Black and Scholes model

In the classical Black and Scholes model the market consists of two assets with dynamics (under the objective probability $P$) given by:

$$\begin{aligned} dB(t) &= rB(t)dt \\ dS(t) &= \mu S(t)dt + \sigma S(t)d\bar{W}(t) \end{aligned} \tag{4.1}$$

where $B$ is the price of a risk free asset, $S$ is the price of an underlying asset and $\bar{W}(t)$ denotes a scalar Wiener process. $r$ is the risk free interest rate, $\mu$ is the local mean rate of return of $S$ and $\sigma$ is the volatility; in the original theory all these parameters are constant.
We now define a new probability measure $Q$ (risk neutral probability measure) such that under $Q$ the dynamics of S is given by:

$$dS(t) = rS(t)dt + \sigma S(t)dW(t) \tag{4.2}$$

where $W(t)$ is a brownian motion under $Q$. The price of a contingent claim $(F(t, S(t)))$ with exercise date $T$ of the form $\Phi(S(T))$, for example a European vanilla call option where $\Phi(S(T)) = max(S(T)-K, 0)$ can be obtained by the following formula, under the assumption that the market is free of arbitrage possibilities:

$$F(t, S(t)) = e^{-r(T-t)}\mathbb{E}^Q[\Phi(S(T))|\mathcal{F}_t] \tag{4.3}$$

where $\mathcal{F}_t$ is the filtration at time $t$. By Feynman-Kac formula we derive that the price of the contingent claim is the solution of the backward parabolic PDE:

$$\begin{cases} \frac{\partial F}{\partial t} + rS\frac{\partial F}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 F}{\partial S^2} - rF = 0 \quad \text{in} \quad \mathbb{R} \times [0, T) \\ F(T, S(T)) = \phi(S(T)) \end{cases} \tag{4.4}$$

In case of an European call option for example, a closed type formula has been derived:

$$C(t, S(t)) = S(t)N[d_1] - Ke^{-r(T-t)}N[d_2] \tag{4.5}$$

where $N(x)$ is the cumulated probability distribution of a standard normal variable and

$$\begin{aligned} d_1 &= \frac{1}{\sigma\sqrt{T-t}}\big(log\frac{S(t)}{K} + (r + \frac{\sigma^2}{2})(T-t)\big) \\ d_2 &= d_1 - \sigma\sqrt{T-t} \end{aligned} \tag{4.6}$$

Under the previous arbitrage assumption, the same results can be obtained defining a self financing portfolio consisting of one unit of the contingent claim and a certain amount of the underlying.

### 4.1.1 Garman-Kohlhagen extension

We now study a model involving not only the domestic market, but also a market for the exchange rate between the domestic currency and a fixed foreign currency, as well as a foreign market. We are interested in derivatives written directly on the exchange rate, which we still call $S$ for the sake of simplicity.
We take as given the following dynamics (under the objective probability measure $P$):

$$\begin{aligned} dB_d(t) &= r_d B_d(t)dt \\ dB_f(t) &= r_f B_f(t)dt \\ dS(t) &= \mu S(t)dt + \sigma S(t)d\bar{W}(t) \end{aligned} \tag{4.7}$$

where $\bar{W}(t)$ denotes a scalar Wiener process, $\mu$, $\sigma$, $r_d$ and $r_f$ (domestic and foreign short rate respectively) are constants.

In this framework a usual risk neutral valuation formula holds:

$$F(t, S(t)) = e^{-r_d(T-t)}\mathbb{E}^Q[\Phi(S(T))|\mathcal{F}_t] \tag{4.8}$$

where $Q$ is characterized by the property that every domestic asset has the short rate $r_d$ as local rate of return under $Q$.
First of all we will assume that all holdings of the foreign currency are invested in the foreign riskless asset, so they will evolve according to the second equation of (4.7).
The possibility of buying foreign currency and investing it at the foreign short rate of interest, is equivalent to the possibility of investing in a domestic asset with price process $\tilde{B_f}$, where

$$\tilde{B_f} = B_f(t)S(t). \tag{4.9}$$

The dynamics of $\tilde{B_f}$ are given by:

$$d\tilde{B_f} = \tilde{B_f}(r_f + \mu)dt + \tilde{B_f}\sigma d\bar{W}(t) \tag{4.10}$$

From previous statements, it follows that the $Q$-dynamics of $\tilde{B_f}$ are:

$$d\tilde{B_f} = \tilde{B_f}r_d dt + \tilde{B_f}\sigma dW(t). \tag{4.11}$$

Since by definition:

$$S(t) = \frac{\tilde{B_f}}{B_f} \tag{4.12}$$

applying Ito's formula we get the $Q$-dynamics of $S$:

$$dS(t) = S(t)(r_d - r_f)dt + S(t)\sigma dW \tag{4.13}$$

So from equation (4.8), we obtain that the price of the contingent claim is now the solution of the PDE:

$$\begin{cases} \frac{\partial F}{\partial t} + (r_d - r_f)S\frac{\partial F}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 F}{\partial S^2} - r_d F = 0 \\ F(T, S_T) = \phi(S(T)) \end{cases} \tag{4.14}$$

Also in this framework is possible to derive a closed type formula for the price of an European vanilla call:

$$C(t, S(t)) = S(t)e^{-r_f(T-t)}N[d_1] - Ke^{-r_d(T-t)}N[d_2] \tag{4.15}$$

where $N(x)$ is the cumulated probability distribution of a standard normal variable and

$$\begin{aligned} d_1 &= \frac{1}{\sigma\sqrt{T-t}}\big(log\frac{S(t)}{K} + (r_d - r_f + \frac{\sigma^2}{2})(T-t)\big) \\ d_2 &= d_1 - \sigma\sqrt{T-t} \end{aligned} \tag{4.16}$$

## 4.2 Black and Scholes with random variables

In this section our objective is to investigate the use of Stochastic Collocation technique applied to forex option pricing under the hypothesis that volatility and interest rates are described by uniform random variables. For the sake of simplicity we will consider only the BS equation with constant parameters, but the results can be easily extended to time dependent parameters, as used in real life. In particular we have focused on European Vanilla options (Section 4.2.1) and European Barrier options (Section 4.2.2). We finally underline that the option value or price is given by the sum of the intrinsic and time value (IV and TV respectively). The intrinsic value is the value of exercising the option now: e.g. for a European vanilla Call Option is given by $max(S(0) - K, 0)$. The time value of an option instead is based on its potential to increase in value before expiring. Therefore for out-of-the money and at-the-money options we will refer to the option value as the time value (TV) since they coincide.

### 4.2.1 European Vanilla Call

As a first example we consider the price of an European vanilla forex call with the following parameters: $S_0 = 100$, $K = 100$, $\sigma = 0.2$, $r_d = 0.05$, $r_f = 0.03$, $T = 1$ (year). Using Garman and Kohlhagen formula price is 8.6525.
We now suppose that the values of $\sigma$, $r_d$ and $r_f$ are not certain but have a uniform distribution over [0.1,0.3], [0.005,0.095], [0.005,0.055] respectively. The variation of the prices with respect to $\sigma$, $r_d$ and $r_f$ is given by the next graphs (Fig. 4.1,4.2,4.3):

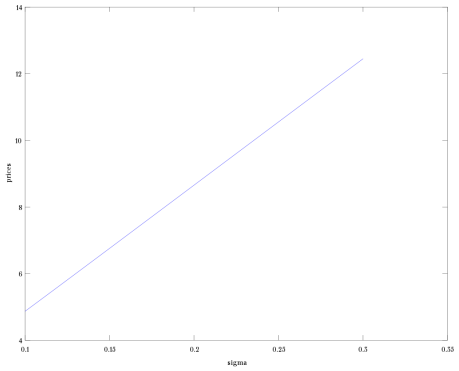


Figure 4.1: TV vs $\sigma$

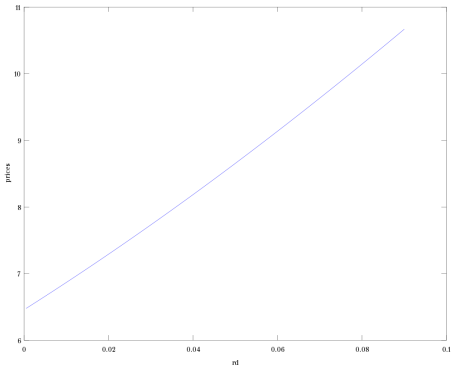


Figure 4.2: TV vs $r_d$

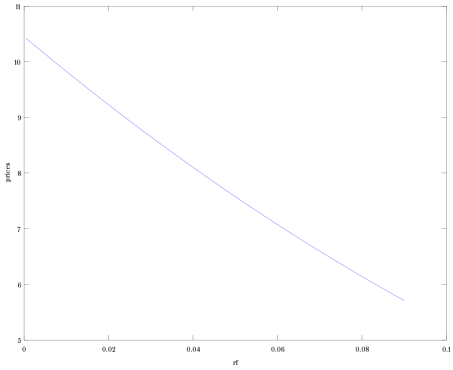


Figure 4.3: TV vs $r_f$

The previous plots show that the dependence of TV on the parameters $\sigma, r_f, r_d$ is almost linear so the mean of the prices with respect to the uniform distribution is almost equal to the price computed with the mean value of the parameters:

$$\mathbb{E}[TV(\theta)] \simeq TV(\mathbb{E}[\theta]) \tag{4.17}$$

for $\theta = \sigma, r_d, r_f$.

### 4.2.2 European Barrier Call

As a second example we consider the price of an European barrier call with the following parameters: $S_0 = 100$, $K = 100$, $\sigma = 0.2$, $r_d = 0.05$, $r_f = 0.03$, $T = 1$, $L = 80$ (low barrier), $U = 120$ (up barrier). Using a PDE solver we get 1.0504.
We again suppose that the value of $\sigma$ is not certain but has a uniform distribution over [0.1,0.3]. The distribution of the prices with respect to $\sigma$ is given by next graph (Fig. 4.4).

In this case the mean of the prices is not equal to the price evaluated with the mean of $\sigma$. In

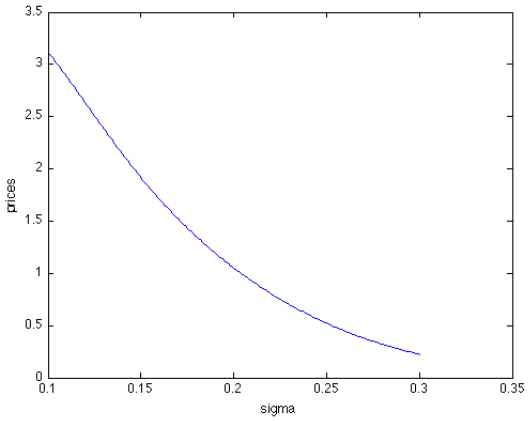


Figure 4.4: TV bar vs $\sigma$

fact considering a polynomial interpolation on Gauss-Legendre points with respect to $\sigma$ with different degree $w$ we get the following mean prices:

| w | Mean | Perc. of Variation |
|---|---|---|
| 1 | 1.0504 | 0% |
| 2 | 1.2764 | 21.52% |
| 3 | 1.2738 | 21.27% |
| 4 | 1.2709 | 20.99% |
| 5 | 1.2707 | 20.97% |
| 6 | 1.2707 | 20.97% |
| 7 | 1.2707 | 20.97% |

Table 4.1: Mean of TV (middle column) and percentage of variation from relation (4.17) (right column) with random volatility $\sigma \sim \mathcal{U}([0.1, 0.3])$

In order to get informations on the distribution of the prices we consider the approximation related to level 5 and we sample the polynomial interpolant in 10000 random i.i.d. samples of $\sigma$ uniformly distributed over $[0.1, 0.3]$. The next figure shows the histogram corresponding to the probability of the prices using 100 bins (Fig. 4.5).

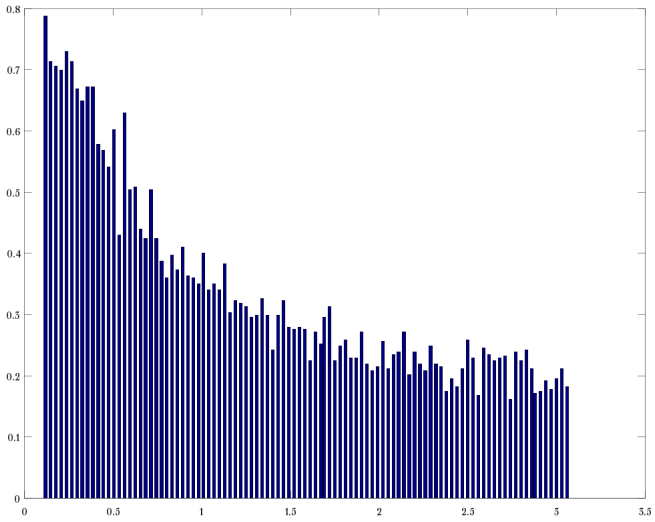


Figure 4.5: Prices histogram

The graph suggests that the probability density might be a truncated exponential distribution, so we use a logarithmic regression in Excel in order to estimate it. As shown in the following graph, the trend seems to be well represented by this kind of regression (Fig. 4.6).

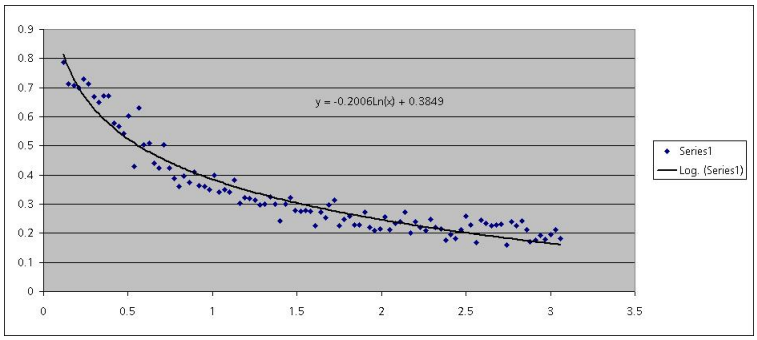


Figure 4.6: Probability distribution of TV as a function of $\sigma \sim \mathcal{U}([0.1, 0.3])$. Regression with a truncated exponential distribution.

Conversely, prices have a linear (or almost linear) trend with respect to $r_d$ and $r_f$ even with barrier options (Fig. 4.7,4.8), therefore any further analysis using uniform distribution for these input parameters would be useless, because formula (4.17) is very accurate.

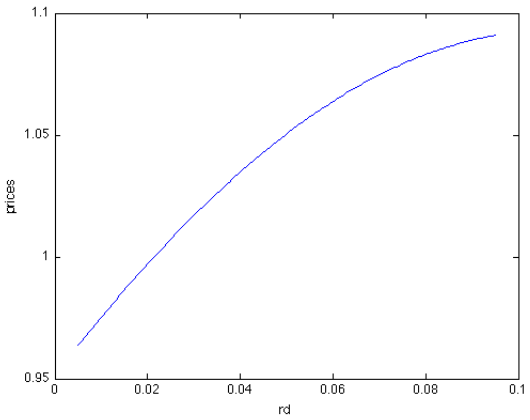


Figure 4.7: TV bar vs $r_d$

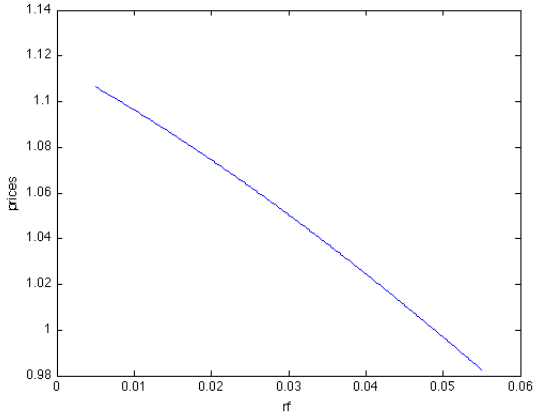


Figure 4.8: TV bar vs $r_f$

Actually, option prices seem to slightly diverge from a linear trend in (Fig. 4.7), in fact using the Stochastic Collocation tecnique with different values of level $w$ we notice just a small bias in relation (4.17)

| w | Mean | Perc. of Variation |
|---|---|---|
| 1 | 1.0504 | 0% |
| 2 | 1.0426 | 0.74% |
| 3 | 1.0427 | 0.74% |
| 4 | 1.0427 | 0.74% |
| 5 | 1.0427 | 0.74% |
| 6 | 1.0427 | 0.74% |
| 7 | 1.0427 | 0.74% |

Table 4.2: Mean of TV (middle column) and percentage of variation from relation (4.17) (right column) with random domestic interest rate $r_d \sim \mathcal{U}([0.005, 0.095])$

It follows that, in case of uniform distribution assumption on the parameters and considering European Vanilla and Barrier Options, the Stochastic Collocation method is useful just in the case of the Barrier option, when (4.17) does not hold for $\theta = \sigma$. Obviously, assuming different and more complex distributions, the use of Stochastic Collocation technique would be fully justified.

## 4.3 Black-Scholes equation with random volatility: Collocation vs Monte Carlo

Given an underlying asset whose dynamics is described by the usual GBM (Geometric Brownian Motion) as in (4.2) we consider the Black-Scholes equation for the pricing of an European Call.

In this further example we assume that the volatility is constant over the time interval but is a random variable with uniform distribution; in particular we want to compare the results obtained with the Stochastic Collocation method with two types of Monte Carlo simulations. The same assumption on volatility has been made by Pulch and von Emmerich for the evaluation of an Asian Option (see [26]). Conversely to our work, they exploit a Polynomial Chaos expansion which results in the above-mentioned Stochastic Galerkin method (Chap. 3).

In this framework $\eta$ is a random variable with uniform distribution over the interval $[0.1, 0.5]$ which represents the underlying volatility. Instead, $\eta_i$ are the volatility values with respect to which we calculate the option price ($f(\eta_i)$) and they correspond to the collocation points. All the prices can be computed solving the Black-Scholes PDE or using the Black-Scholes formula (4.5), while the mean value will be given by (2.7). We now remark how to define collocation points on the interval $[-1, 1]$ related to the Legendre polynomials and associated to a random variable with uniform distribution on $[-1, 1]$ (see Chapter 2). We can apply a linear transformation mapping the interval $[-1, 1]$ into $[0.1, 0.5]$ and recast the problem to one depending on a uniform random variable in $[0.1, 0.5]$. Indeed if $\eta$ is a random variable with uniform distribution on $[-1, 1]$ then $V = 0.3 + 0.2\eta$ has a uniform distribution on $[0.1, 0.5]$. It follows:

$$\mathbb{E}[f] = \int_{0.1}^{0.5} f(v)\frac{1}{0.4}dv = \int_{-1}^{1} f(0.3 + 0.2\eta)\frac{1}{2}d\eta. \tag{4.18}$$

By the same linear transformation we can obtain the collocation points on $[0.1, 0.5]$ from those on $[-1, 1]$. We can both use the Gauss-Legendre formula (2.7) with weights given by (2.8) for the integral defined on $[-1, 1]$ or approximate the integral on $[0.1, 0.5]$ with appropriate weights.

We can also use the Monte Carlo method in two ways. In the first one we generate M values of volatility with uniform distribution on $[0.1, 0.5]$ and then we apply formula (4.5) M times. The option price is simply approximated by the sample mean of these M prices. In the second one, after simulating M samples of volatility, we simulate the underlying process at time T by:

$$S_i(T) = S(0)exp((r - 0.5\sigma_i^2)T + \sigma_i\sqrt{T}Z_i) \tag{4.19}$$

where $Z_i$ $i = 1..M$ are values generated from a standard normal distribution. The option value is the mean of the discounted payoffs.

### 4.3.1 Numerical results

We decide to evaluate the option price by using the Matlab financial toolbox function *blsprice* instead of solving the Black-Scholes PDE to reduce the computational time. We will call *MC bls* the price given by the Monte Carlo method coupled with the blsprice function,

*MC payoffs* the price obtained by formula (4.19) and *Collocation* the one computed with the Collocation method. In the following table, M represents the number of collocation points/Monte Carlo samples. The other parameters are $S(0) = 100, K = 90, r = 0.03, T = 1$.

| M | Collocation | MC bls | MC payoffs |
|---|---|---|---|
| 5 | 18.7854668346 | 21.864798 | 7.502145 |
| 10 | 18.7854307168 | 18.710908 | 6.298455 |
| 15 | 18.7854290264 | 19.826537 | 41.926484 |
| 50 | 18.7854439741 | 18.936011 | 17.488151 |
| 100 | 18.7854439741 | 18.411561 | 21.62262 |
| 1000 | | 18.746321 | 18.87440 |
| 10000 | | 18.745300 | 18.774770 |

Table 4.3: Prices and Number of Points/MC Samples

We decide not to evalute the price for high numbers of collocation points because already M=100 gives the price up to machine precision. In Fig. 4.9 and 4.10, we plot the value reported in the previous table as well as the 95% confidence interval given by the Monte Carlo method. Notice the different vertical scale in the two plots and the much larger variance estimates.

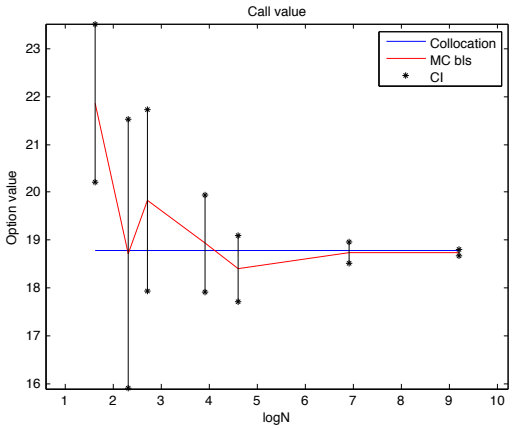


Figure 4.9: Prices with Collocation and MC bls methods with confidence interval as functions of $M$

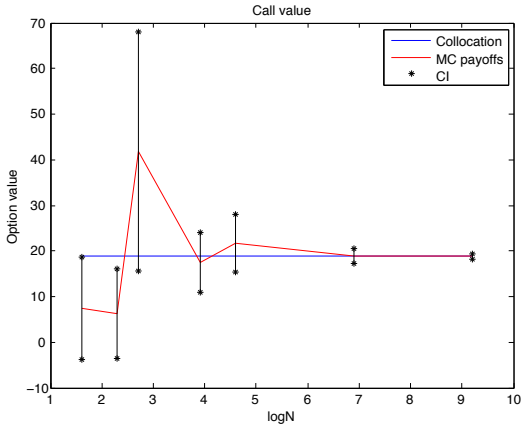


Figure 4.10: Prices with Collocation and MC payoff methods with confidence interval as functions of $M$

We can assume one of the collocation value for M=100 as highly accurate approximation of the real option value and plot the errors of these 3 methods versus the number of points M (Fig. 4.11).

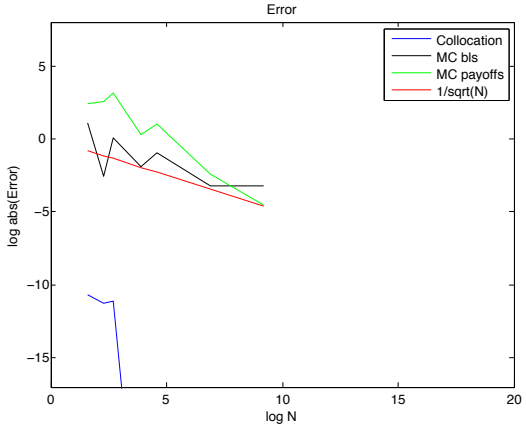


Figure 4.11: Log of the error vs log of evaluation points when $\sigma$ is a uniform random variable over $[0.1, 0.5]$

### 4.3.2 Greeks computation

As previously stated we use a polynomial approximation, that is

$$f(\eta) \simeq f_w(\eta) = \sum_{i=1}^{w+1} f(\eta_i) l_i(\eta) \tag{4.20}$$

where $l_i(\eta) = \prod_{j \neq i} \frac{\eta - \eta_j}{\eta_i - \eta_j}$ is $i$-th Lagrange polynomial and $\eta$ is the stochastic parameter value. In particular $\eta_i$ are the Stochastic Collocation knots, that are the volatility values in which we

have already computed the option prices ($f(\eta_i)$). We seek a price approximation with respect to all volatility value $\eta$ as well as the price derivative with respect to volatility evaluated in $\eta$, known as vega. In fact to compute vega it is enough to differentiate (4.20) and we obtain

$$f'(\eta) \simeq f'_w(\eta) = \sum_{i=1}^{w+1} f(\eta_i) l'_i(x).$$

The last formula shows that we can compute vega (or any other derivative with respect to a stochastic parameter) by a sum of polynomial derivatives.

We consider 5 or 10 Stochastic Collocation points and build the polynomial interpolant in order to approximate the option prices and the vega on a uniform grid on [0.1,0.5] consisting in 20 knots. We can also compare the vega values computed with the exact ones given by the following formula: $Vega(S(0),0) = S(0)\sqrt{T}\phi(d_1)$ where $\phi$ is the density function of the standard normal distribution and $d_1 = \frac{1}{\sigma\sqrt{T}}\{ln(\frac{S(0)}{K}) + (r + 0.5\sigma^2)T\}$.

Fig. 4.12 and 4.13 show the results obtained for the option price and the vega respectively. It follows that the price curves are identical for 5 and 10 collocation points, while vega is better approximated with 10 points.

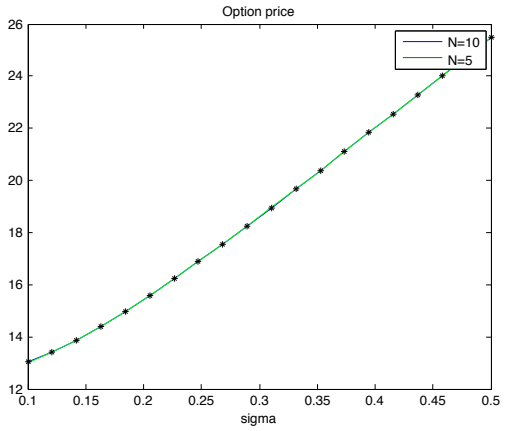


Figure 4.12: Price Curve with M=5,10 (the black knots represent the exact values)

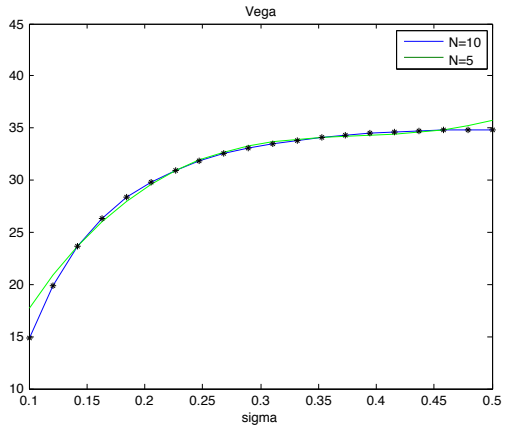


Figure 4.13: Vega with M=5,10 (the black knots represent the exact values)

## 4.4 Black-Scholes equation with random volatility and random risk-free interest rate: Collocation vs Monte Carlo

In addition to the assumptions of the previous model, we consider a risk-free interest rate with a uniform distribution over $[0.01, 0.11]$.

As far as stochastic collocation is concerned we use Legendre knots in dimension 2, obtained with the hyperbolic cross rule and different values of level $w$; we will therefore use (2.20) to compute mean values.

For the Monte Carlo method we generate $M$ couples $(\sigma_i, r_i)$ of volatility and risk-free interest rate with the above-mentioned distributions and use the *blsprice* function (MC bls).

Another alternative is to simulate the path for each couple $(\sigma_i, r_i)$

$$S_i(T) = S(0)exp((r_i - 0.5\sigma_i^2)T + \sigma_i\sqrt{T}Z_i) \tag{4.21}$$

where $Z_i, i = 1, \ldots, M$ are values generated from a normal standard distribution. The option value (OV) will be given by the mean value of the payoffs (MC payoffs)

$$\mathbb{E}[OV] \simeq exp(-r_iT)\frac{1}{M}\sum_{i=1}^{M} max(S_i(T) - K, 0). \tag{4.22}$$

The next tables summarize the results obtained (Table 4.4 and 4.5).

| w | M | Prices |
|---|---|---|
| 4 | 16 | 20.520477461011048 |
| 6 | 33 | 20.51913579695262 |
| 8 | 44 | 20.518882783165765 |
| 9 | 61 | 20.518898786999689 |

Table 4.4: Mean Prices (right column), levels (left column) and corresponding Number of Points (middle column) with random volatility $\sigma \sim \mathcal{U}([0.1, 0.5])$ and random risk-free interest rate $r \sim \mathcal{U}([0.01, 0.11])$, Collocation method

| M | Prices - MC bls | Prices - MC payoff |
|---|---|---|
| 5 | 18.90631268079 | 13.902687861092456 |
| 10 | 20.31652644405 | 34.188704500741416 |
| 15 | 19.61017709424 | 32.080885656803979 |
| 50 | 20.68906181798 | 19.221960125179837 |
| 100 | 20.31956232255 | 18.797934791542374 |
| 1000 | 20.57911553664 | 21.676892634061694 |
| 10000 | 20.51741068877 | 20.357372989470225 |

Table 4.5: Mean Prices and Number of Samples with random volatility $\sigma \sim \mathcal{U}([0.1, 0.5])$ and random risk-free interest rate $r \sim \mathcal{U}([0.01, 0.11])$, Monte Carlo methods

The next graphs show the mean prices and the 95% confidence intervals related to Table 4.5 and the Collocation mean prices from Table 4.4 (Fig. 4.14 and 4.15).

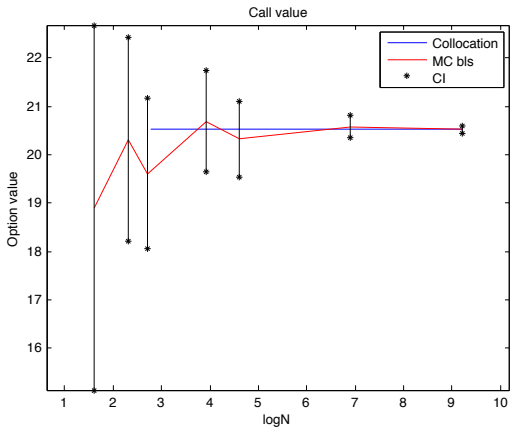


Figure 4.14: Prices MC bls method (middle column of Table 4.5) with related 95% CI and Collocation mean values (right column of Table 4.4) vs log Number of Points/MC Samples, $\sigma \sim \mathcal{U}([0.1, 0.5])$ and $r \sim \mathcal{U}([0.01, 0.11])$

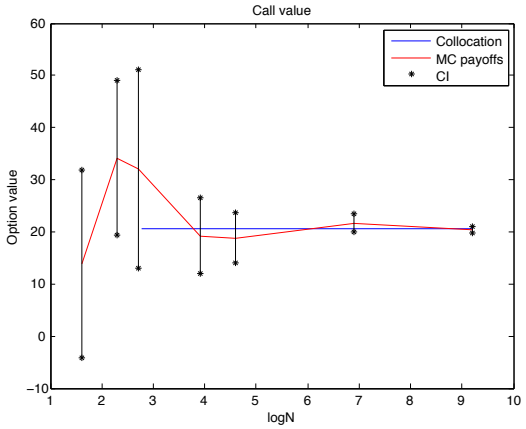


Figure 4.15: Prices MC payoff method (right column of Table 4.5) with related 95% CI and Collocation mean values (right column of Table 4.4) vs log Number of Points/MC Samples, $\sigma \sim \mathcal{U}([0.1, 0.5])$ and $r \sim \mathcal{U}([0.01, 0.11])$

The last graph (Fig. 4.16) instead shows the error vs number of points after assuming the mean price obtained with level 9 in Table 4.4 as the exact one.

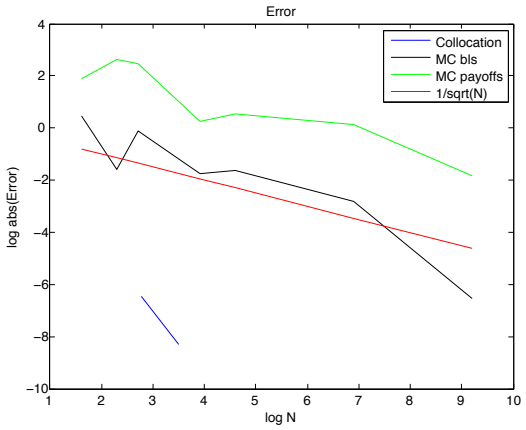


Figure 4.16: Log Error as function of number of log samples when $\sigma \sim \mathcal{U}([0.1, 0.5])$ and $r \sim \mathcal{U}([0.01, 0.11])$

In this framework we can conclude that the Collocation technique seems to be the best one and we can fairly state the MC bls method provides more accurate results than the MC payoff one.

# Chapter 5

# Approximation of stochastic processes

In the Section 5.1 we will justify the assumption of stochastic volatility and introduce some of the best-known models. After presenting the Euler and Milstein scheme in Section 5.2, we will investigate the effectiveness of the Stochastic Collocation coupled with these schemes when applied to stochastic volatility models. Then we will make a comparison between Stochastic Collocation and Monte Carlo method in terms of computational cost (Section 5.6) and discuss a different discretization method to approximate Wiener paths called Brownian Bridge construction (Section 5.7).

## 5.1 Stochastic volatility models

The Black and Scholes model has shown significant bias with respect to historical data. First of all the logarithm of the price of the underlying does not show a normal distribution. Moreover the assumption of constant parameters can be easily rejected.
We suppose to observe from the market the price of a given call option $c$ with value of strike $K$, today's value of underlying $S(t)$ and risk free interest rate $r$. Applying Black and Scholes formula we can find the value of $\sigma$ (called implied volatility) such that

$$c = C(t, S(t), K, T, r, \sigma) \tag{5.1}$$

where $C$ denotes the Black-Scholes formula; we repeat the same computation for different values of $K$. If we plot implied volatility as a function of the exercise price we should obtain an horizontal straight line. Conversely it is often observed that options far out of the money and deep into the money are traded at higher implied volatility than options at the money. This volatility curve is usually termed the volatility smile.
To overcome these discrepancies, new models have been proposed; among them we will focus

on Heston model belonging to the set of stochastic volatility models [6].

Refusing the assumption of constant volatility it is possible to introduce a process for the volatility $\sigma$ (or the variance $v = \sigma^2$) such that:

- $v(t)$ is positive at every time
- $v(t)$ is mean reverting: if the process diverges from a certain mean value $\bar{v}$ then its dynamics tends to move it closer to the mean value

We suppose that the stock price $S$ and its variance $v$ satisfy the following SDEs under the objective probability measure

$$\begin{cases} dS(t) = \mu S(t)dt + \sqrt{v(t)}S(t)dW_1(t) \\ dv(t) = \alpha(S, v, t)dt + \beta(S, v, t)\sqrt{v(t)}dW_2(t) \end{cases} \tag{5.2}$$

where $dW_1(t)dW_2(t) = \rho dt$. As presented in [24], in this framework, it is possible to derive a pricing formula for a contingent claim defining a risk-free portfolio consisting in the contingent claim, a certain amount of the underlying and another amount of an asset whose value depends on volatility. Our pricing equation is given by:

$$\frac{\partial F}{\partial t} + rS\frac{\partial F}{\partial S} + \frac{1}{2}\big(S^2 v\frac{\partial^2 F}{\partial S^2} + 2\rho\beta^2 S\frac{\partial^2 F}{\partial S \partial v} + \beta^2 v\frac{\partial^2 F}{\partial v^2}\big) - rF = -(\alpha - \phi\beta)\frac{\partial F}{\partial v} \tag{5.3}$$

where $\phi = \phi(S, v, t)$ is the market price of volatility risk that we will consider equal to 0. While equation (4.4) can be recast in the classical heat equation using a change of variable, equation (5.3) degenerates at the boundary: for $v = 0$ all the coefficients of the second order derivatives vanish and the operator becomes hyperbolic. In this case however any change of variables prevents the degeneration of the elliptic operator. A special theory is necessary to examine the well posedness of this particular class of equations: parabolic equations degenerating at the boundary of the domain. As a further consequence, one source of uncertainty has been added ($W_2$) but not another underlying asset traded in the market. This implies that the risk neutral probability measure is not unique anymore, but it is still possible to define a measure $Q$ such that the $Q$-dynamics of $S$ are given by:

$$dS(t) = rS(t)dt + \sqrt{v(t)}S(t)dW_1(t). \tag{5.4}$$

### 5.1.1 Some examples

- ***Hull-White model***

  The first model we will analyze is the Hull-White model for volatility (see [20]):

$$d\sigma = \sigma(\bar{\xi}dt + \bar{\epsilon}dW_2) \tag{5.5}$$

or applying Ito's Lemma

$$dv = v(\xi dt + \epsilon dW_2) \tag{5.6}$$

where $\xi = 2\bar{\xi} + \bar{\epsilon}^2$ and $\epsilon = 2\bar{\epsilon}$. In his paper, Hull and White presented this Geometric Brownian Motion model of volatility which is positive at every time but not mean reverting, under the assumption of zero correlation between the brownian motions driving the underlying and the volatility processes.

- ***Scott model***
  Another possible model is due to Scott [21] where $y = \log v$

$$dy = (\bar{\theta} - \chi y)dt + \epsilon dW_2 \tag{5.7}$$

or

$$dv = v(\theta - \chi \log v)dt + \epsilon v dW_2 \tag{5.8}$$

where $\theta = \bar{\theta} + \frac{\epsilon^2}{2}$. In this case, the log-variance is a classical Ornstein Uhlenbeck process, thus the variance process is mean reverting and positive at every time. This model was at first introduced by Scott without correlation in [21], but has been extended to a correlated one by Scott and Chesney in [22].

- ***Heston model***
  The most famous model in this class is due to Heston [6]

$$dv(t) = \chi(\theta - v(t))dt + \epsilon\sqrt{v(t)}dW_2(t). \tag{5.9}$$

This is a Square Root Model where $\theta$ is the long term volatility, $\chi$ is the rate of mean reversion and $\epsilon$ is the volatility of volatility.
If the condition $2\chi\theta \geq \epsilon^2$ is satisfied, then the process is also strictly positive, otherwise the origin is accessible and strongly reflecting (see [35]). Among the stochastic volatilty models, it is also the most famous since Heston provided an analytical pricing formula in [6].

## 5.2 Euler and Milstein discretization schemes

In the rest of chapter we will analyze time discrete approximation of the processes introduced in Section 5.1, relying on weak and strong convergence results. In particular we shall focus on the Euler and Milstein methods.

Let $X = \{X_t, 0 \leq t \leq T\}$ be an Ito process satisfying the scalar stochastic differential equation:

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t \tag{5.10}$$

with initial value $X(0) = X_0$. For a given discretization $0 = t_0 < t_1 < \ldots < t_n \ldots < t_N = T$ of the time interval, an Euler approximation is a continuous time stochastic process $Y = \{Y(t), t_0 \leq t \leq T\}$ satisfying the iterative scheme

$$Y_{n+1} = Y_n + a(t_n, Y_n)(t_{n+1} - t_n) + b(t_n, Y_n)(W_{t_{n+1}} - W_{t_n}) \tag{5.11}$$

for $n = 0, 1, \ldots N-1$ with initial value $Y_0 = X_0$ where $Y_n = Y(t_n)$. $\{W(t), t \geq 0\}$ is a Wiener process and from theory we know that each increment $\Delta W_n = (W_{t_{n+1}} - W_{t_n})$ is independent from all the others and is distribuited as a Gaussian random variable with mean 0 and variance $t_{n+1} - t_n$. In the same way we can define the Milstein approximation as a continuous time stochastic process $Y = \{Y(t), t_0 \leq t \leq T\}$ satisfying the iterative scheme

$$Y_{n+1} = Y_n + a(t_n, Y_n)(t_{n+1} - t_n) + b(t_n, Y_n)\Delta W_{n+1} + \frac{1}{2} b(t_n, Y_n) b'(t_n, Y_n)\{(\Delta W_n)^2 - (t_{n+1} - t_n)\} \tag{5.12}$$

for $n = 0, 1, \ldots N-1$ with initial value $Y_0 = X_0$.
We define $\delta = max_n \Delta_n = max_n(t_{n+1} - t_n)$ and introduce two kinds of convergence for the approximated process $(Y^\delta)$ with respect to the maximum step size $(\delta)$.
We shall say that a time discrete approximation $Y^\delta$ converges weakly with order $\beta$ to $X$ at time $T$ as $\delta \to 0$ if for each $g \in \mathcal{C}^{2(\beta+1)}$ there exists a positive constant $C$, which does not depend on $\delta$, and a finite $\delta_0 > 0$ such that

$$|\mathbb{E}[g(X(T))] - \mathbb{E}[g(Y^\delta(T))]| \leq C\delta^\beta \tag{5.13}$$

for each $\delta \in (0, \delta_0)$. This kind of convergence requires an approximation of the probability distribution of $X(T)$. It can be proved (see [10]) that under regularity conditions on $a$ and $b$ both Euler and Milstein scheme converge weakly with order $\beta = 1$.

We shall say that a time discrete approximation $Y^\delta$ converges strongly with order $\gamma > 0$ at time $T$ if there exists a positive constant $C$, which does not depend on $\delta$, and a $\delta_0 > 0$ such that

$$\mathbb{E}[|X(T) - Y^\delta(T)|] \leq C\delta^\gamma \tag{5.14}$$

for each $\delta \in (0, \delta_0)$. Conversely to the weak convergence, formula (5.14) is the expectation of the absolute value of the difference between the Ito process and the approximation at time $T$: it requires a pathwise approximation of the Ito process.
Also in this case the convergence orders are provided in [10]. In particular

**Teorema 5.2.1.**

*Suppose the initial condition $X_0$ to be deterministic and*

$$\begin{aligned} &|a(t,x)-a(t,y)|+|b(t,x)-b(t,y)| \le K_1|x-y| \\ &|a(t,x)|+|b(t,x)| \le K_2(1+|x|) \\ &|a(s,x)-a(t,x)|+|b(s,x)-b(t,x)| \le K_3(1+|x|)|s-t|^{1/2} \end{aligned} \tag{5.15}$$

*for all $s,t \in [0,T]$ and $x,y \in \mathbb{R}$, where the constants $K_1, K_2, K_3$ do not depend on $\delta$. Then for the Euler approximation $Y^\delta$ the estimate*

$$\mathbb{E}[|X(T)-Y^\delta(T)|] \le K_4\delta^{1/2} \tag{5.16}$$

*holds where the constant $K_4$ doed not depend on $\delta$.*

So if $a$ and $b$ satisfy a linear growth and a Lipschitz condition then the Euler approximation has strong convergence order 0.5.
Similarly for Milstein scheme if we define $\underline{a} = a - \frac{1}{2}bb'$ then the following result holds:

**Teorema 5.2.2.**
*Suppose the initial condition $X_0$ to be deterministic and*

$$\begin{aligned} &|\underline{a}(t,x)-\underline{a}(t,y)| \le K_1|x-y| \\ &|b(t,x)-b(t,y)| \le K_1|x-y| \\ &|b'(t,x)-b'(t,y)| \le K_1|x-y| \\ &|\underline{a}(t,x)|+|\underline{a}'(t,x)| \le K_2(1+|x|) \\ &|b(t,x)|+|b'(t,x)| \le K_2(1+|x|) \\ &|\underline{a}(s,x)-\underline{a}(t,x)| \le K_3(1+|x|)|s-t|^{1/2} \\ &|b(s,x)-b(t,x)| \le K_3(1+|x|)|s-t|^{1/2} \\ &|b'(s,x)-b'(t,x)| \le K_3(1+|x|)|s-t|^{1/2} \end{aligned} \tag{5.17}$$

*for all $s,t \in [0,T]$ and $x,y \in \mathbb{R}$, where the constants $K_1, K_2, K_3$ do not depend on $\delta$. Then for the Milstein approximation $Y^\delta$ the estimate*

$$\mathbb{E}[|X(T)-Y^\delta(T)|] \le K_4\delta \tag{5.18}$$

*holds where the constant $K_4$ doed not depend on $\delta$.*

As in [10] (see Theorem 10.6.3) we can recast the strong convergence orders for both our

schemes with respect to the following norms

$$\mathbb{E}[sup_{0\leq t\leq T}|X(t)-Y^{\delta}(t)|] \qquad \text{or} \qquad \left(\mathbb{E}[sup_{0\leq t\leq T}|X(t)-Y^{\delta}(t)|^{p}]\right)^{1/p} \tag{5.19}$$

for any $p \leq 2$. Recently Higham, Mao and Stuart (see [17]) have proved a convergence result for the Euler scheme in the strong mean square sense ($p = 2$). In particular they have shown that the Euler scheme has strong convergence order 0.5 even if the coefficient $a$ and $b$ are locally Lipschitz and $r$-th moment of the exact and numerical solution are bounded for some $r > 2$.

In the following sections, we aim at approximating common volatility processes and derive a convergence rate for the stochastic differential equations (5.5), (5.7) and (5.9).
The key idea is that, at final time, approximated processes (just like in (5.11) and (5.12)) can be thought of as functions of $N$ independent random variables $\eta_1 \dots \eta_N$ representing the increments of the brownian motion (scaled by the time step $\Delta t$). In a general framework we could use the Monte-Carlo method possibly coupled with usual variance reduction techniques (e.g. antithetic variables or control variables). From our point of view we shall focus on the Stochastic Collocation technique to test the convergence result. In particular our random knots will be given by the collocation points defined in Chapter 2, so the objective is to evaluate the weak and strong error and obtain a convergence rate with respect to $\Delta t$ and the level $w$, which determines the number of evaluation knots. This estimation is necessary to choose optimally the two discretization parameters $\Delta t$ and $w$ in order to balance the two error contributions.

## 5.3 Hull-White model

We will now focus on equation (5.5) with the following parameters $\sigma_0 = 0.2$, $\bar{\xi} = 0.2$ e $\bar{\epsilon} = 0.1$. We solve the equation with the Euler scheme:

$$\sigma_{n+1} = \sigma_n[1 + \bar{\xi}\Delta t + \bar{\epsilon}\sqrt{\Delta t}\eta_{n+1}] \tag{5.20}$$

where $\eta_{n+1} \sim N(0,1),\ n = 1 \dots N-1$ or with the Milstein scheme

$$\sigma_{n+1} = \sigma_n[1 + \bar{\xi}\Delta t + \bar{\epsilon}\sqrt{\Delta t}\eta_{n+1} + \frac{1}{2}\bar{\epsilon}^2\Delta t(\eta_{n+1}^2 - 1)]. \tag{5.21}$$

We want to evaluate the weak error (5.13) with $g(x) = x^2$. In particular, having set a level $w$ and letting $\Delta t$ vary, we compute $|\mathbb{E}[\sigma(T)^2] - \mathbb{E}[(\sigma_{\Delta t}^w(T))^2]|$, which means that we compare the exact mean $\mathbb{E}[\sigma(T)^2] = \sigma_0^2 \exp\{2\bar{\xi}T + \bar{\epsilon}^2 T\}$ with that obtained solving the SDE (5.5) on a sparse grid of level $w$ and evaluating the mean value at the final time step through (2.20).

For all of our simulations we shall rely on the hyperbolic cross rule. With schemes (5.20) and (5.21) we get  It is evident that when the time discretization error is dominant ($w \geq 4$) the

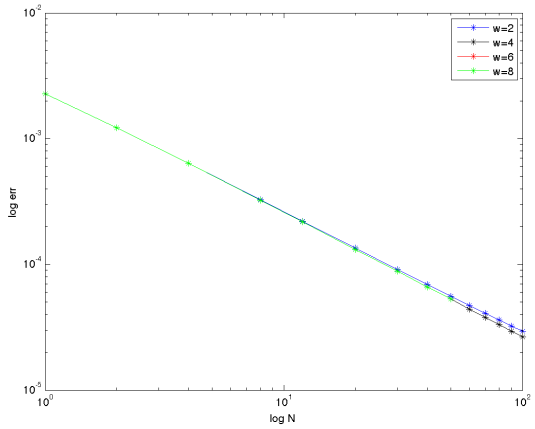


Figure 5.1: Hull-White, Weak error for Stochastic Collocation method coupled with Euler scheme: the error is given by $|\mathbb{E}[\sigma(T)^2] - \mathbb{E}[(\sigma^w_{\Delta t}(T))^2]|$

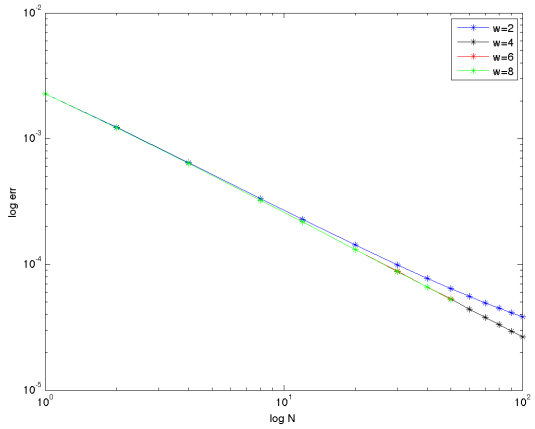


Figure 5.2: Hull-White, Weak error for Stochastic Collocation method coupled with Milstein scheme: the error is given by $|\mathbb{E}[\sigma(T)^2] - \mathbb{E}[(\sigma^w_{\Delta t}(T))^2]|$

usual linear convergence with respect to the time step is observable.

For the strong convergence we evaluate $\left(E\left[(\sigma(T)-\sigma^w_{\Delta t}(T))^2\right]\right)^{\frac{1}{2}}$ by generating $M = 20000$ samples with standard normal distribution, that is we generate $M$ Wiener paths using different time steps ($\Delta t$) or equivalently a different number of random variables $N$. For each $N$-dimensional sample we sum all the components, obtaining therefore a Wiener path $v_j$, $j = 1 \dots M$; we then evaluate the related volatility trajectory by $\sigma(T;\omega_j) = \sigma_0 e^{(\bar{\xi}-\bar{\epsilon}^2/2)T+\bar{\epsilon}v_j\sqrt{\Delta t}}$. At the same time, for each $N$, we generate a sparse grid of level $w$ with $N$ variables and approximate the process over the collocation knots with the Euler/Milstein scheme. So we are able to interpolate the $M$ samples over the sparse grid and obtain the related values at the final time $T$. To estimate the strong error we use the following formula

$$\left(E\left[(\sigma(T)-\sigma^w_{\Delta t}(T))^2\right]\right)^{\frac{1}{2}} \simeq \sqrt{\sum_{j=1}^{M}\frac{1}{M}(\sigma(T;\omega_j)-\sigma^w_{\Delta t}(T;\omega_j))^2}. \tag{5.22}$$

The next figures show the results obtained for Euler and Milstein scheme (Fig. 5.3 and 5.4).

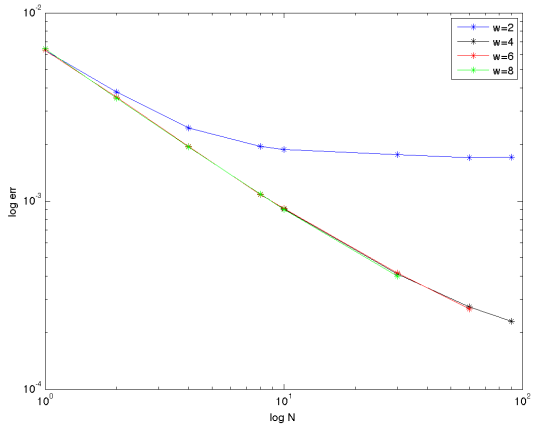


Figure 5.3: Hull-White, Strong error for Stochastic Collocation method coupled with Euler scheme: the error is given by $\left(E\left[(\sigma(T)-\sigma^w_{\Delta t}(T))^2\right]\right)^{\frac{1}{2}} \simeq \sqrt{\sum_{j=1}^{M}\frac{1}{M}(\sigma(T;\omega_j)-\sigma^w_{\Delta t}(T;\omega_j))^2}$

From the figures, it follows that if the level is sufficiently high ($w = 8$), the usual convergence rates (0.5 for Euler and 1 for Milstein) are observable.

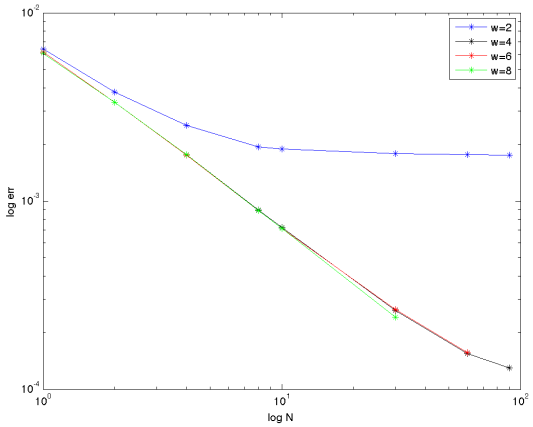


Figure 5.4: Hull-White, Strong error for Stochastic Collocation method coupled with Milstein scheme: the error is given by $\left(E\left[(\sigma(T)-\sigma^w_{\Delta t}(T))^2\right]\right)^{\frac{1}{2}} \simeq \sqrt{\sum_{j=1}^{M} \frac{1}{M}(\sigma(T;\omega_j)-\sigma^w_{\Delta t}(T;\omega_j))^2}$

From Fig. 5.1 and 5.2, it emerges that the Stochastic Collocation method with Hyperbolic Cross rule perferms well using a level $w = 4$. In fact in this case both for Euler and Milstein schemes the usual rates of convergence related to the weak error are observable and this implies that the time discretization error is dominant. Moreover with our choice of parameters $\mathbb{E}[\sigma(T)^2] \simeq 6.0210^{-2}$ and using a value $N = 20$ the weak error is close to 1%. It has to be noticed although that all the bias between level 2 and 4 have maximum magnitude of order $10^{-5}$, so also the level $w = 2$ has to be taken into account for further simulations.
From Fig. 5.3 and 5.4 it follows that the level $w = 4$ is still the best choice since higher levels do not decrease the approximation error (which is dominated by the time discretization error).

In order to compute an estimation of the probability approximation error, we consider the exact mean of the discrete process at the $N$-th time step, which is given by $\mathbb{E}[\sigma^2_{\Delta t}(T)] = \sigma_0^2(1+\bar{\xi}^2\Delta t^2+\bar{\epsilon}^2\Delta t+2\bar{\xi}\Delta t)^N$ for Euler scheme. We compare this value with the mean of the square values computed by solving the SDE over the sparse grid of level $w$. We thus get a new type of weak convergence error shown in the next graphs. We remark that conversely for Milstein scheme the discrete mean is given by $\mathbb{E}[\sigma^2_{\Delta t}(T)] = \sigma_0^2(1+\bar{\xi}^2\Delta t^2+\bar{\epsilon}^2\Delta t+2\bar{\xi}\Delta t+\frac{1}{2}\bar{\epsilon}^4\Delta t^2)^N$.

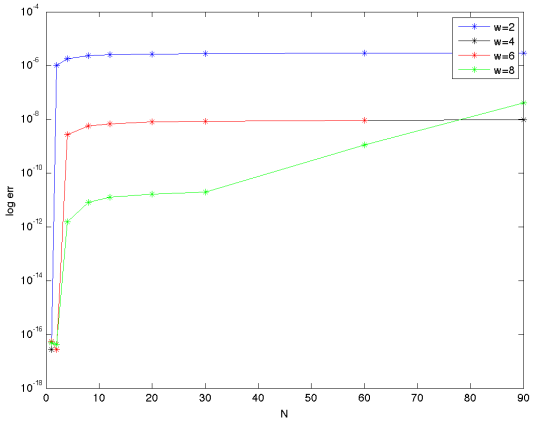


Figure 5.5: Hull-White, Weak probability error for Stochastic Collocation method coupled with Euler scheme with respect to the exact discrete mean: the error is given by $|\mathbb{E}[\sigma^2_{\Delta t}(T)] - \mathbb{E}[(\sigma^w_{\Delta t}(T))^2]|$

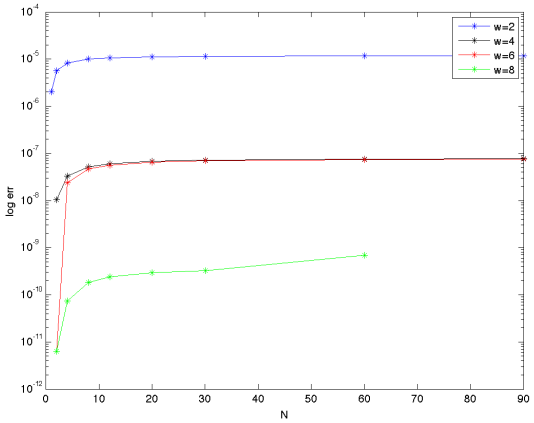


Figure 5.6: Hull-White, Weak probability error for Stochastic Collocation method coupled with Milstein scheme with respect to the exact discrete mean: the error is given by $|\mathbb{E}[\sigma^2_{\Delta t}(T)] - \mathbb{E}[(\sigma^w_{\Delta t}(T))^2]|$

It is also possible to evaluate a different kind of strong convergence. In fact we generate $M$ samples and solve (5.5) with a time discrete scheme. At the same time we evaluate the discrete process over a grid of level $w$ and then interpolate the $M$ samples over it. Applying formula (5.22) we get an estimation of strong error for both schemes shown in the next figures.

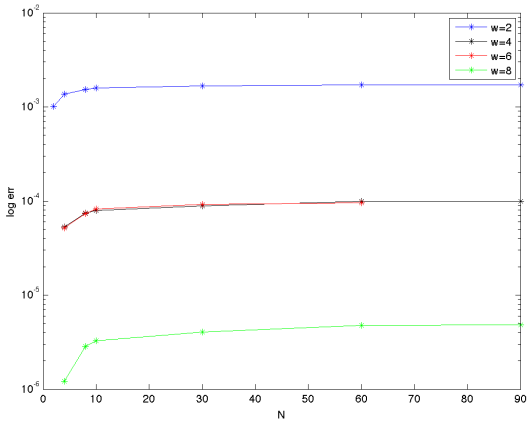


Figure 5.7: Hull-White, Strong probability error for Stochastic Collocation method coupled with Euler scheme: the error is given by $\left(E\left[(\sigma_{\Delta t}(T) - \sigma^w_{\Delta t}(T))^2\right]\right)^{\frac{1}{2}} \simeq \sqrt{\sum_{j=1}^{M} \frac{1}{M}(\sigma_{\Delta t}(T)(\omega_j) - \sigma^w_{\Delta t}(T;\omega_j))^2}$

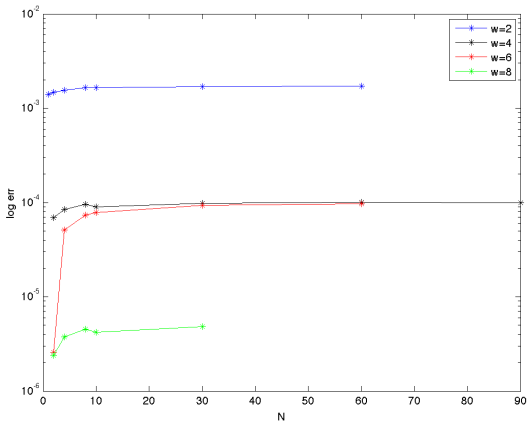


Figure 5.8: Hull-White, Strong probability error for Stochastic Collocation method coupled with Milstein scheme: the error is given by $\left(E\left[(\sigma_{\Delta t}(T) - \sigma^w_{\Delta t}(T))^2\right]\right)^{\frac{1}{2}} \simeq \sqrt{\sum_{j=1}^{M} \frac{1}{M}(\sigma_{\Delta t}(T)(\omega_j) - \sigma^w_{\Delta t}(T;\omega_j))^2}$

From Fig. 5.5, 5.6, 5.7 and 5.8 it follows that the probability error increases with the number of variables involved, therefore as the latter grows, a higher value of level $w$ is required to keep the error low. The best choice however is still $w = 4$ since the level $w = 6$ is basically equivalent while the level $w = 8$ involves a much greater number of points and is thus computationally more expensive: for example for $N = 10$ the sparse grid with level $w = 8$ is composed of 5951 points while the level $w = 4$ only 511.

## 5.4 Scott model

We now consider equation (5.7) with the following parameters: $y_0 = -2.4$, $\epsilon = 0.4$, $\xi = 0.1$ e $\bar{\theta} = -0.12$. Firstly we observe that in this case, since the diffusion coefficient doesn't depend on the process that we aim at approximating, Euler and Milstein scheme coincide and are equal to

$$y_{n+1} = y_n + (\bar{\theta} - \chi y_n)\Delta t + \epsilon\sqrt{\Delta t}\eta_{n+1} \tag{5.23}$$

where $\eta_{n+1} \sim N(0,1),\ n = 0 \dots N-1$.
We estimate the weak convergence (5.13) with $g(x) = e^x$. The exact mean is given by $\mathbb{E}[e^{y(T)}] = exp\{e^{-\xi T}y_0 + \frac{\bar{\theta}}{\xi}(1 - e^{-\xi T}) + \frac{1}{4\xi}\epsilon^2(1 - e^{-2\xi T})\}$ and we compare it with the mean of the exponential values obtained by solving (5.23) over a sparse grid of level $w$. The next figure shows these results for different values of $N$ and $w$: for high values of $w$ a linear convergence is observable.

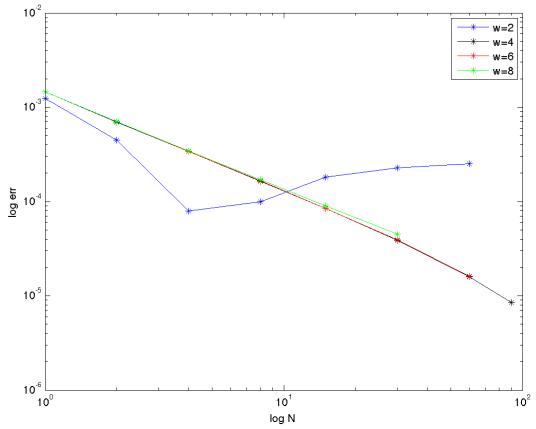


Figure 5.9: Scott, Weak error for Stochastic Collocation method: the error is given by $|\mathbb{E}[\exp(y(T))] - \mathbb{E}[\exp(y^w_{\Delta t}(T)]|$

In order to estimate the strong error, we know that the solution at time $T$ of (5.7) is given by $y(T) = e^{-\xi T}y_0 + \frac{\bar{\theta}}{\xi}(1 - e^{-\xi T}) + \epsilon\int_0^T e^{-\xi(T-s)}dW(s)$. The first problem consists in computing the Ito integral $\int_0^T e^{-\xi(T-s)}dW(s)$ accurately. For this purpose we divide the interval $[0,T]$ into $N' = 1000$ sub-interval of equal length $\Delta t'$ and generate $M$ $N'$-dimensional random samples $(\mathbf{Z}_j,\ j = 1 \dots M)$ with standard normal distribution. For each $N'$-dimensional knot we are able to approximate the integral in the Ito way as $\sum_{k=0}^{N'-1} e^{-\xi(T-t'_k)}\sqrt{\Delta t'}\eta_{k+1}$ and obtain $y(T;\omega_j),\ j = 1 \dots M$. At the same time we generate a sparse grid with level $w$ with respect to $N \ll N'$ random variables and apply scheme (5.23) over the collocation knots. For each $N$, let us define $n = N'/N$ and a $N$-dimensional vector $\mathbf{v}^i$ with components defined as $v_i^j = \frac{1}{\sqrt{n}}\sum_{k=1}^n Z^j_{in+k},\ i = 0 \dots N-1, j = 1 \dots M$. It is now possible to obtain the values $y^w_{\Delta t}(T;\omega_j)$ evaluating the vectors $\mathbf{v}^i$ over the interpolating functions related to the sparse

grid. The strong error is the one defined in (5.22). This figure shows a linear convergence

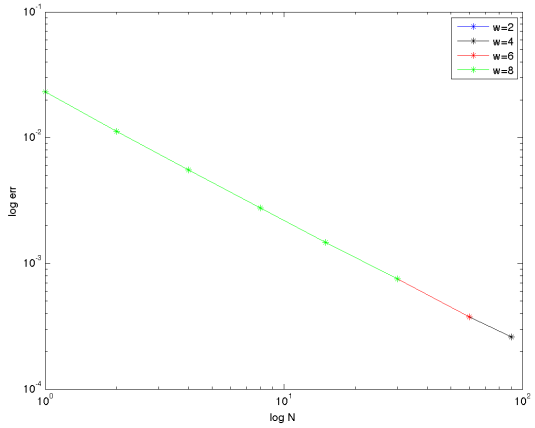


Figure 5.10: Scott, Strong error for Stochastic Collocation method: the error is given by $\left(E\left[(y(T) - y^w_{\Delta t}(T))^2\right]\right)^{\frac{1}{2}} \simeq \sqrt{\sum_{j=1}^{M} \frac{1}{M}(y(T;\omega_j) - y^w_{\Delta t}(T;\omega_j))^2}$ and is based on an approximation for the stochastic integral appearing in $y(T)$

for the strong error for every value of $w \geq 2$, since from (5.23) it follows that the process $y = y(T, \eta_1, \ldots, \eta_N)$ is a linear function of the random vector $\boldsymbol{\eta} = [\eta_1, \ldots, \eta_N]$ and therefore it is interpolated exactly by a sparse grid of level $w \geq 1$.
Also in this case we can compute an estimation of the probability approximation error, considering the exact mean of the discrete process at the $N$-th time step, which is given by $\mathbb{E}[e^{y_{\Delta t}(T)}] = e^{m+\frac{d}{2}}$, where $m = \sum_{n=1}^{N}(1-\Delta t)^{N-n}\bar{\theta}\Delta t + (1-\Delta t)^N y_0$ and $d = \epsilon^2 \Delta t \sum_{n=1}^{N}(1-\Delta t)^{2(N-n)}$. We compare this value with the mean of the square values computed by solving the SDE over the sparse grid of level $w$. We thus get a new type of weak convergence error shown in the next graph (Fig. 5.11) which is analogous to Fig. 5.5, 5.6.

Also for this model the level $w = 4$ provides a good trade-off between accurancy and computational cost: the level $w = 8$ seems to perform better in Fig. 5.11 but involves too many evaluation knots, the level $w = 2$ is equivalent for the strong error in Fig. 5.10 since we are approximating a multi-linear function, but Fig. 5.9 represents a sure sign that the level $w = 4$ is the best choice. In particular for our choice of parameters the mean value $\mathbb{E}[e^{y(T)}]$ is equal to 0.1093: in order to approximate this mean value with a precision close to 1% only ten variables/time intervals are necessary ($N = 10$).

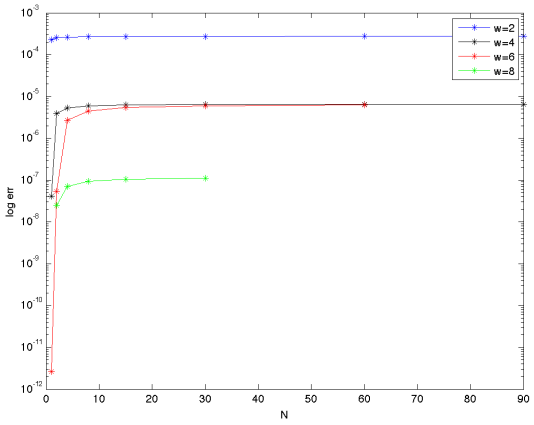


Figure 5.11: Scott, Weak probability error for Stochastic Collocation method with exact discrete mean: the error is given by $|\mathbb{E}[\exp(y_{\Delta t}(T))] - \mathbb{E}[(\exp(y^w_{\Delta t}(T))]|$

## 5.5 Heston model

In this section we shall discuss the approximation techniques for a square root process (5.9). As previously stated the condition $2\chi\theta > \epsilon^2$ guarantees that 0 is an unattainable barrier for the continuous time process. However if we discretize it with an Euler scheme we get

$$v_{n+1} = v_n + \chi(\theta - v_n)\Delta t + \epsilon\sqrt{v_n}\sqrt{\Delta t}\eta_{n+1} \tag{5.24}$$

for $i = 0 \dots N-1$ where each $\eta_{n+1} \sim N(0,1)$, while with the Milstein scheme

$$v_{n+1} = v_n + \chi(\theta - v_n)\Delta t + \epsilon\sqrt{v_n}\sqrt{\Delta t}\eta_{n+1} + \frac{1}{4}\epsilon^2\Delta t(\eta^2_{n+1} - 1) \tag{5.25}$$

where each $\eta_{n+1} \sim N(0,1)$. This means that the approximated process can become negative, so we will consider modified Euler and Milstein schemes. We shall focus on Euler scheme first.

Intuitively the easiest way to fix the approximated scheme consists in setting the process equal to zero whenever it attains a negative value or reflecting it in the origin: these strategies are named absorption or reflection. We will present all the fixes proposed as well as some related convergence results. First of all we can recast all these schemes as:

$$v_{n+1} = f_1(v_n) + \chi(\theta - f_2(v_n))\Delta t + \epsilon\sqrt{f_3(v_n)}\sqrt{\Delta t}\eta_{n+1} \tag{5.26}$$

where the functions $f_i,\ i = 1,2,3$ have to satisfy:

- $f_i(x) = x$ for $x \geq 0$ and $i = 1,2,3$

- $f_3(x) \geq 0$ for $x \in \mathbb{R}$

These conditions guarantees that the scheme does not introduce bias: if the variance is positive we don't introduce any correction, otherwise we have to correct the square root term ($f_3(x) \geq 0$).
In the absorption fix we consider $f_1(x) = f_2(x) = f_3(x) = x^+$ where $x^+ = max(x, 0)$. There are no papers discussing this method so the expected rate of convergence is still unknown. In the reflection fix we set $f_1(x) = f_2(x) = f_3(x) = |x|$: the main results are due to Bossy and Diop (see [13]) and Berkaoui (see[12]). In particular Bossi and Diop have proven that an Euler discretization with the reflection fix converges weakly in time step with order given by $min\{1, \chi\theta\epsilon^{-2}\}$ while Berkaoui et al. (see[12]) have proven a strong converge in the following sense. Since in (5.9) the drift coefficient is globally Lipschitz with constant $\chi$ then if $\Delta t \leq min\{v_0, 1/(2\chi)\}$ and

$$\frac{\epsilon^2}{8}\big(\frac{2\chi\theta}{\epsilon^2} - 1\big) \geq \mathcal{K}(4p) \tag{5.27}$$

where $\mathcal{K}(4p) = max\{K(4p-1), 2\epsilon(2p-1)^2\}$ then the exact process $v(t)$ and its approximation $v^\delta(t)$ satisfy

$$\big(\mathbb{E}[sup_{0\leq t\leq T}|v(t) - v^\delta(t)|^{2p}]\big)^{1/2p} \leq C(T, p)\sqrt{\Delta t}. \tag{5.28}$$

Hingham and Mao in [11] have considered the following scheme $f_1(x) = f_2(x) = x$ and $f_3(x) = |x|$ proving strong convergence of order 0.5 for their discretization without restriction on the parameters. The same result holds for the partial truncation scheme by Deelstra and Delbaen ([14]) obtained by setting $f_1(x) = f_2(x) = x$ and $f_3(x) = x^+$. In 2006 Lord, Koekkoek and van Dijk (see [9]) proposed a new scheme called full truncation method, which consists in choosing $f_1(x) = x$ and $f_2(x) = f_3(x) = x^+$. In this framework they proved a strong convergence with order 0.5 in $L^1$ sense and 0.25 in $L^2$ sense:

$$\begin{aligned} &sup_{t\in[0,T]}\mathbb{E}\left[|v(t) - v^\delta(t)|\right] \leq C_1\sqrt{\Delta t} \\ &\mathbb{E}\left[sup_{t\in[0,T]}|v(t) - v^\delta(t)|^2\right]^{1/2} \leq C_2(\Delta t)^{1/4}. \end{aligned} \tag{5.29}$$

Though no proof has been provided, the authors state that probably the true order of $L^2$ convergence is higher.
As a final remark, in [9] is shown that if we want to simulate the underlying process as well (with Euler or Milstein method) in order to price a European Call option, the absorption fix seems to present the highest bias. The partial truncation and full truncation might be the best approximation, the latter in particular seems to be the best one. The reflection fix and Hingham-Mao method also show a quite high bias.

The Milstein scheme has positive probability of generating negative paths as well. We will

rely on Milstein-type schemes, for example the implicit Milstein scheme:

$$v_{n+1} = \frac{v_n + \chi\theta\Delta_n + \epsilon\sqrt{v_n}\eta_{n+1}\sqrt{\Delta_n} + 0.25\epsilon^2\Delta_n(\eta_{n+1}^2 - 1)}{1 + \chi\Delta_n}. \tag{5.30}$$

This scheme generates strictly positive paths if $4\chi\theta > \epsilon^2$ (see[15]). Equation (5.30) can be recast as a Balanced Milstein method presented in [16]:

$$v_{n+1} = v_n + \chi(\theta - v_n)\Delta_n + \epsilon\sqrt{v_n}\eta_n\sqrt{\Delta_n} + 0.25\epsilon^2\Delta_n(\eta_{n+1}^2 - 1) + (v_n - v_{n+1})D_n(v_n) \tag{5.31}$$

where $D_n(v_n) = d_0(v_n)\Delta_n + d_1(v_n)\Delta_n(\eta_{n+1}^2 - 1)$. In order to obtain (5.30) it is sufficient to set $d_0(v_n) = \chi$ and $d_1(v_n) = 0$. Milstein schemes have first order weak and strong convergence, under the regularity conditions (5.17). From numerical tests however it emerges that these schemes perform fairly well in certain parameter regimes but are typically less robust than Euler scheme: in fact due to the presence of a square root in (5.9) the conditions (5.17) are violated.
Kahl in [16] has shown a linear strong convergence in the $L^2$ sense $\left(\mathbb{E}[|X(t) - Y^\delta(t)|^2]\right)^{1/2}$ for this choice of parameters: $\chi = \theta = 1, \epsilon = 1.4$.
Other important schemes are provided by Alfonsi in [25].

### 5.5.1 Further considerations

From standard theory (see [27]) we know that defining

$$d = \frac{4\chi\theta}{\epsilon^2}; \qquad n(t,T) = \frac{4\chi e^{-\chi(T-t)}}{\epsilon^2(1 - e^{-\chi(T-t)})} \tag{5.32}$$

$v(T)$, conditional on $v(t)$, is distribuited as $e^{-\chi(T-t)}/n(t,T)$ times a non-central chi-square distribution with $d$ degrees of freedom and non-centrality parameter $v(t)n(t,T)$. In particular the first two moments of this distribution are:

$$\begin{aligned} &\mathbb{E}[v(T)|v(t)] = \theta + (v(t) - \theta)e^{-\chi(T-t)} \\ &Var[v(T)|v(t)] = \frac{v(t)\epsilon^2 e^{-\chi(T-t)}}{\chi}(1 - e^{-\chi(T-t)}) + \frac{\theta\epsilon^2}{2\chi}(1 - e^{-\chi(T-t)})^2. \end{aligned} \tag{5.33}$$

Therefore $v_{n+1}$ is proportional to a non-central chi-square distribution with non-centrality parameter $v_n n(t_n, t_{n+1})$ where $n$ is independent of $v_n$. In all Euler-type schemes (5.26) we are approximating the distribution of $v_{n+1}$ conditional to $v_n$ with a Gaussian variable. These methods then provide good results if $v(t)$ is sufficiently large since the non-central chi-square distribution approaches a Gaussian distribution if the non-centrality parameter tends to $\infty$ (or also if $\Delta_n \to 0$ and $n(t_n, t_{n+1}) \to \infty$). Then the approximation of $v_{n+1}$ with a Gaussian variable is not accurate when $v_n$ is very small, tipically when the positivity condition is not

satisfied.

In order to overcome this bias, more efficient schemes have been proposed (Truncated Gaussian or Quadratic Exponential scheme, see [27]) but they are not consistent with the Stochastic Collocation technique.

### 5.5.2 Numerical results

In this section we will consider scheme (5.30) and the reflection fix for Euler approximations. Since an analytical solution for (5.9) does not exist for both, so we can only compute weak convergence and strong probability errors (formula 5.22).
In the first case we evaluate $|\mathbb{E}[v(T)] - \mathbb{E}[(v^w_{\Delta t}(T))]|$, which means that we compare the exact mean $\mathbb{E}[v(T)] = \theta + (v_0 - \theta)e^{-\chi T}$ with that obtained solving the SDE (5.5) on a sparse grid of level $w$ and evaluating the mean value at the final time step through (2.20). If we choose $\chi = 0.3$, $\theta = 0.4$, $\epsilon = 0.3$ (the positivity condition is satisfied) we get the results in Fig. 5.12 and 5.13. For the Implicit Milstein scheme because we notice that the conditional mean value

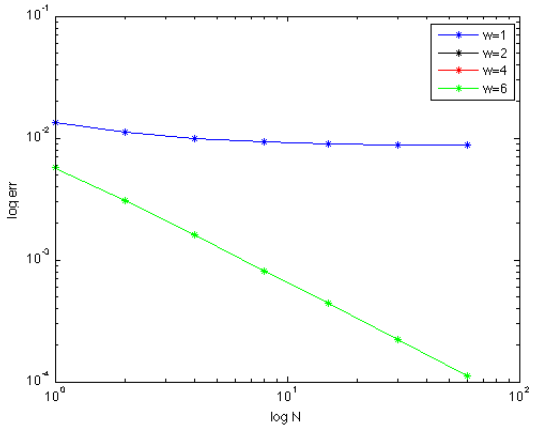


Figure 5.12: Heston, Weak error for Stochastic Collocation method coupled with Implicit Milstein scheme: the error is given by $|\mathbb{E}[v(T)] - \mathbb{E}[(v^w_{\Delta t}(T))]|$. Positivity condition $2\chi\theta > \epsilon^2$ is satisfied

is computed exactly with every level $w \geq 1$ and $w \geq 2$ for the Euler scheme. The convergence rate then is linear with the first scheme and (maybe) more than linear with the latter. With our choice of parameters we get $\mathbb{E}[v(T)] = 0.251836$, therefore a value of $N$ at least equal to 10 guarantees a relative error equal to 1%.
If we increase $\epsilon$ from 0.3 to 0.5 and the positivity condition is slightly violated we get exactly the same results.

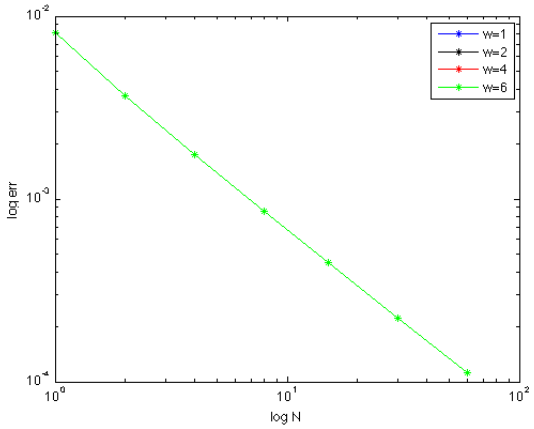


Figure 5.13: Heston, Weak error for Stochastic Collocation method coupled with Euler scheme and reflection fix: the error is given by $|\mathbb{E}[v(T)] - \mathbb{E}[(v^w_{\Delta t}(T))]|$. Positivity condition $2\chi\theta > \epsilon^2$ is satisfied

As previously stated, we will also focus on the strong probability errors (formula 5.22). In particular if we choose $\chi = 0.3$, $\theta = 0.4$, $\epsilon = 0.3$ (the positivity condition is satisfied) we get the results shown in Fig. 5.14 and 5.15.

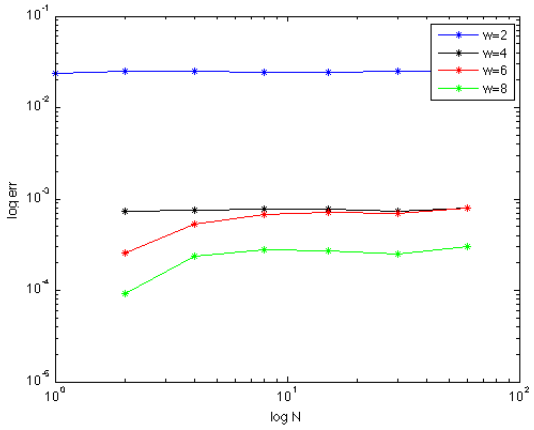


Figure 5.14: Heston, Strong probability error for Stochastic Collocation method coupled with Implicit Milstein scheme: the error is given by $\left(E\left[(v_{\Delta t}(T) - v^w_{\Delta t}(T))^2\right]\right)^{\frac{1}{2}} \simeq \sqrt{\sum_{j=1}^{M} \frac{1}{M}(v_{\Delta t}(T;\omega_j) - v^w_{\Delta t}(T;\omega_j))^2}$. Positivity condition $2\chi\theta > \epsilon^2$ is satisfied

The Implicit Milstein scheme provides results similar to other stochastic processes. In the Euler scheme with reflection fix we are introducing a non-smooth dependency of the process with respect to the random variables which is due to the absolute value function. This may cause an error rate decrease when increasing the number of the random variables involved in the approximation. Also for this process, we can conclude that for the strong probability error the best choice would be $w = 4$.

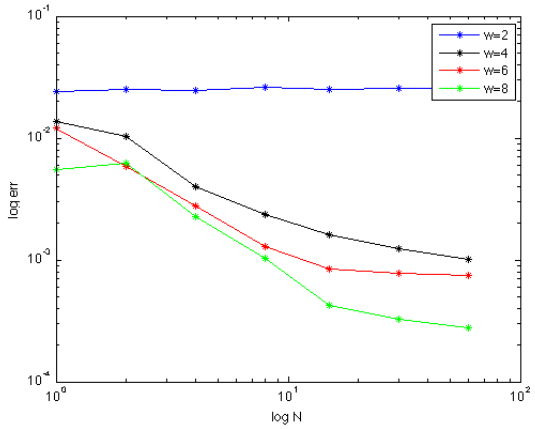


Figure 5.15: Heston, Strong probability error for Stochastic Collocation method coupled with Euler scheme and reflection fix: the error is given by $\left(E\left[(v_{\Delta t}(T) - v^w_{\Delta t}(T))^2\right]\right)^{\frac{1}{2}} \simeq \sqrt{\sum_{j=1}^{M} \frac{1}{M}(v_{\Delta t}(T;\omega_j) - v^w_{\Delta t}(T;\omega_j))^2}$.Positivity condition $2\chi\theta > \epsilon^2$ is satisfied

If we set $\chi = 0.3$, $\theta = 0.4$, $\epsilon = 0.5$ (the positivity condition is not satisfied) we get similar results (Fig. 5.16 and 5.17).

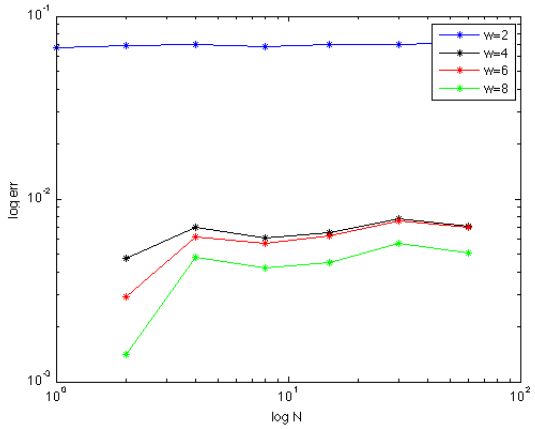


Figure 5.16: Heston, Strong probability error for Stochastic Collocation method coupled with Implicit Milstein scheme: the error is given by $\left(E\left[(v_{\Delta t}(T) - v^w_{\Delta t}(T))^2\right]\right)^{\frac{1}{2}} \simeq \sqrt{\sum_{j=1}^{M} \frac{1}{M}(v_{\Delta t}(T;\omega_j) - v^w_{\Delta t}(T;\omega_j))^2}$. Positivity condition $2\chi\theta > \epsilon^2$ is not satisfied

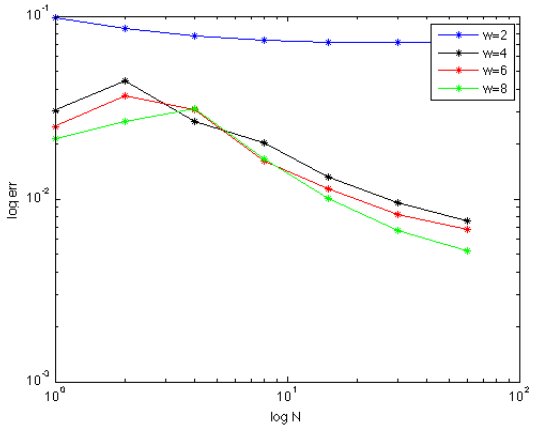


Figure 5.17: Heston, Strong probability error for Stochastic Collocation method coupled with Euler scheme and reflection fix: the error is given by $\left(E\left[(v_{\Delta t}(T) - v^w_{\Delta t}(T))^2\right]\right)^{\frac{1}{2}} \simeq \sqrt{\sum_{j=1}^{M} \frac{1}{M}(v_{\Delta t}(T;\omega_j) - v^w_{\Delta t}(T;\omega_j))^2}$. Positivity condition $2\chi\theta > \epsilon^2$ is not satisfied

## 5.6 A comparison with Monte Carlo method

In this section we want to compare the two possible probability approximation techniques: Monte Carlo and Stochastic Collocation method. From the well known Central Limit Theorem we know that the first method converges with order $1/\sqrt{M}$, where $M$ is the number of realizations of the random variables the approximating function depends on and this rate does not depend on the number of random variables $N$. The latter technique instead has a convergence rate which is still unknown: the convergence results provided in Chapter 2 are not related to Stochastic Differential Equations so we shall try to derivate a convergence result from our simulations.

For the Monte Carlo method we can define the total weak error ($\epsilon$) as the sum of the weak error (5.13) and the probability approximation error, which is thus given by $\epsilon \sim \frac{1}{N} + \frac{1}{\sqrt{M}}$. In particular for a given value of $N$ (number of variables/time discretization intervals) in order to have the same magnitude of the two contributions, we have to set $M = C_* N^2$, where $C_*$ is a constant, and obtain $\epsilon \sim \frac{1}{N}$. We now define the cost of one simulation as $C = MN$, where $C$ represents the number of function evaluations (and the total number of samples of normal random variables generated). From our choice of $M$ it follows $C \sim N^3$ and therefore $C \sim \frac{1}{\epsilon^3}$ (or equivalently $\epsilon \sim C^{-\frac{1}{3}}$). From this theoretical estimation we know how to choose optimally the value of $M$ given $N$ but we have also been able to define a relationship between the global error and the total cost: if we get an error $\epsilon_1$ with a certain cost $C_1$, in order to halve $\epsilon_1$ we expect the cost $C_2$ to be at least eight times $C_1$.

For the stochastic collocation technique these results are unknown and we will try to derive an estimation for the three volatility models we have dealt with so far.

Since the Monte Carlo method has been gradually improved we will compare our convergence

rates with the theoretical rates of more efficient techniques: the Multi-level Monte Carlo and the Quasi-Monte Carlo methods. The first one consists in using a geometric sequence of grids, each twice as fine as its predecessor; the relationship between error and computional cost is given by $C = (log(\epsilon))^2/\epsilon^2$ (see [29]). The second technique consists in generating quasi-random sequences of numbers instead of random ones (also known as low-discrepancy sequences). For Quasi-Monte Carlo methods we know that $\epsilon \sim (logM)^k/M$ for some $k$ related to $N$ [28]: in particular when integrating a function it has been proved that Quasi-Monte Carlo methods lose their effectiveness if the function has a non-smooth dependence from the random variables or the number of variables becomes too large.

### 5.6.1 Hull-White model

As a first equation we shall consider the Geometric Brownian motion which represents the Hull-White model of volatility. For different values of $N$ we will evaluate the weak error $|\mathbb{E}[\sigma(T)^2] - \mathbb{E}[(\sigma^w_{\Delta t}(T))^2]|$ (just like in Section 5.3) as a function of the above-mentioned computational cost $C$ (see the Fig. 5.18). We will use the usual Hyperbolic Cross rule and the Euler scheme with the same parameters of Section 5.3.

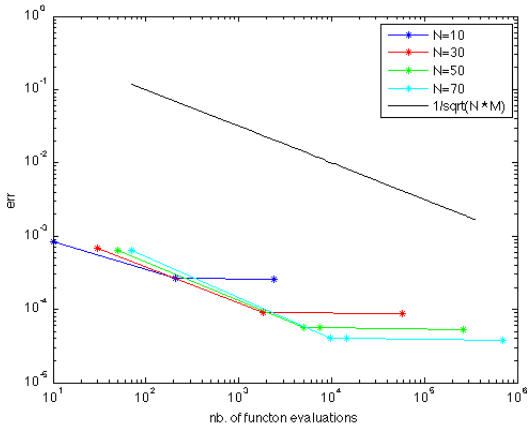


Figure 5.18: Hull-White, Weak error for Stochastic Collocation method coupled with Euler scheme: the error is given by $|\mathbb{E}[\sigma(T)^2] - \mathbb{E}[(\sigma^w_{\Delta t}(T))^2]|$

The probability approximation convergence is observable until the time discretization error related to the Euler scheme becomes dominant. We can fairly state that for every value of $N$ the cost corrisponding to the level $w = 2$ saturates the probability error since the difference between the level $w = 2$ and $w = 4$ has a magnitude of $10^{-6}$ (see Fig. 5.1).
For Monte Carlo simulations, given a value of $N$ and $M$ we repeat the same simulation 100 times and get thus a mean value of $|\mathbb{E}[\sigma(T)^2] - \mathbb{E}[(\sigma^w_{\Delta t}(T))^2]|$ in order to obtain a more accurate estimation and also a more meaningful graph. Fig. 5.19 shows the error versus cost curve obtained.

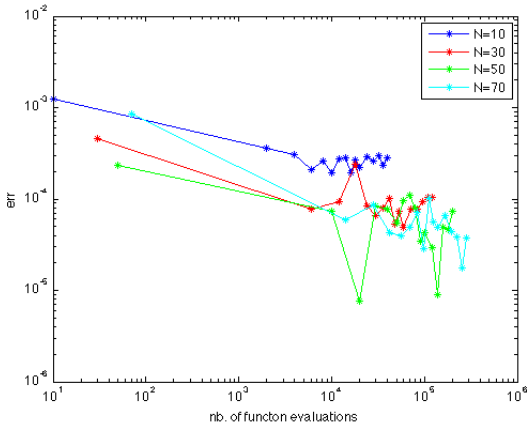


Figure 5.19: Hull-White, Weak error for Stochastic Collocation method coupled with Euler scheme: the error is given by $|\mathbb{E}[\sigma(T)^2] - \mathbb{E}[(\sigma^w_{\Delta t}(T))^2]|$

Fig. 5.20 shows a direct comparison between Monte Carlo and Stochastic Collocation for a fixed value $N = 32$. For a variable $N$, to compare the complexity of the two approaches (error versus total cost), one has to choose optimally the level $w$ (for Stochastic Collocation) and the number of samples $M$ in Monte Carlo. In Fig. 5.21 this is done in the following way: for Stochastic Collocation, we rely on plot 5.18 and take the minimum among all curves for $N = 10, 30, 50, 70$. This gives the blue line in Fig. 5.21. For Monte Carlo we have only plot the theoretical rate $\epsilon \sim C^{-\frac{1}{3}}$ since the curves in Fig. 5.19 are too noisy. Furthermore we also plot the function $C = (log(\epsilon))^2/(\epsilon)^2$ which represents the convergence rate of the Multi-level Monte Carlo method.

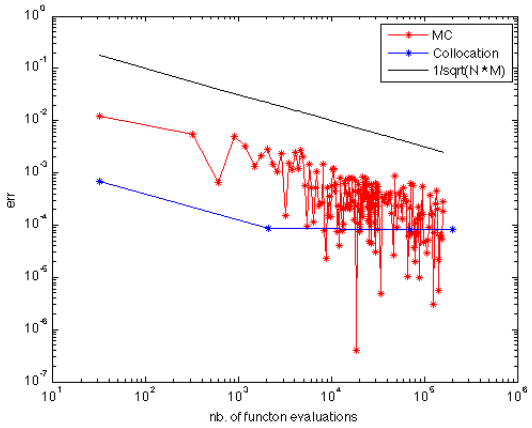


Figure 5.20: Hull-White, Weak error for Stochastic Collocation and Monte Carlo methods coupled with Euler scheme: the error is given by $|\mathbb{E}[\sigma(T)^2] - \mathbb{E}[(\sigma^w_{\Delta t}(T))^2]|$, N=32

The line corresponding to Stochastic Collocation is perfectly approximated by a straight line with angular coefficient $\sim -0.485$ and has therefore a convergence rate similar to the

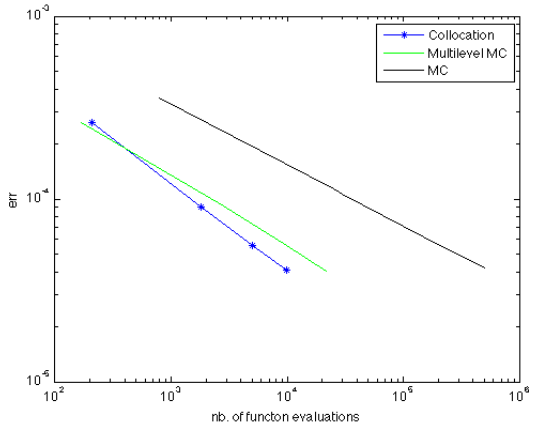


Figure 5.21: Hull-White, A comparison of the costs for different approximation techniques

Quasi-Monte Carlo and Multi-level Monte Carlo methods on this model.
With Milstein scheme we get analogous results.

### 5.6.2 Scott model

We now focus on the Ornstein-Uhlenbeck process which represents the Scott model of volatility. For different values of $N$ we will evaluate the weak error $|\mathbb{E}[\exp(y(T))] - \mathbb{E}[\exp(y^w_{\Delta t}(T))]|$ (just like in Section 5.4) as a function of the above-mentioned computational cost $C$ (see Fig. 5.22). We will use the same parameters of Section 5.4.

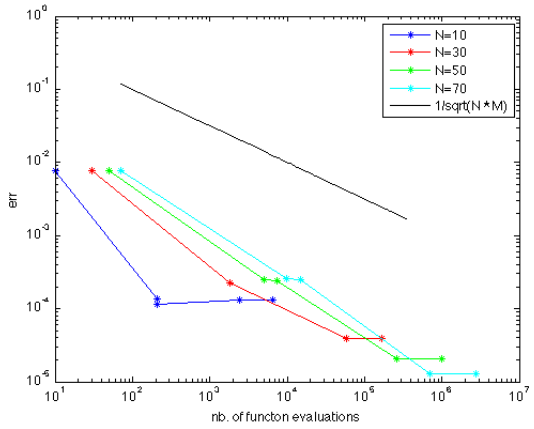


Figure 5.22: Scott, Weak error for Stochastic Collocation method: the error is given by $|\mathbb{E}[\exp(y(T))] - \mathbb{E}[\exp(y^w_{\Delta t}(T))]|$

The probability approximation convergence is observable until the time discretization error related to the time discrete scheme becomes dominant. We can state that for every value of

$N \geq 30$ the cost corrisponding to the level $w = 3$ saturates the probability error (see also Fig. 5.9). In this case we will not provide the graph analogous to Fig. 5.19 since the convergence curves are too noisy and no information can be extracted out of it. However in order to compare the two methods, next figure shows the results for $N = 32$ where Monte Carlo errors are computed on a single run just like in Fig. 5.20.

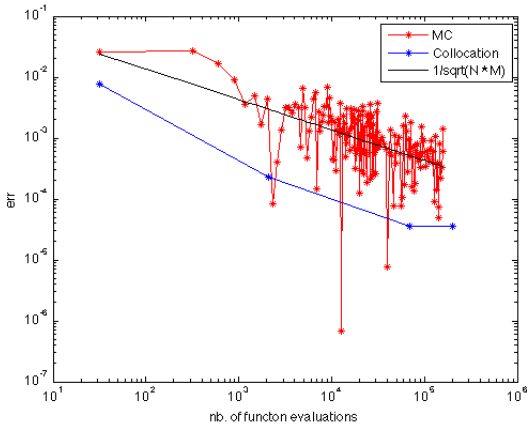


Figure 5.23: Scott, Weak error for Stochastic Collocation and Monte Carlo methods: the error is given by $|\mathbb{E}[\exp(y(T))] - \mathbb{E}[\exp(y^w_{\Delta t}(T))]|$, N=32

If we extract the minimum of all curves in Fig. 5.22 for $N = 10, 30, 50, 70$ this will be an estimate of the complexity of the algorithm. Such minimum is well interpolated by a quadratic curve given by $log(\epsilon) = -0.064(log(C))^2 + 0.237 * log(C) - 4.09$. In Fig. 5.24 the blue line represents the minimum among all curves for $N = 10, 30, 50, 70$, while for Monte Carlo we have only plot the theoretical rate $\epsilon \sim C^{-\frac{1}{3}}$. In this case the Collocation rate seems to be superior to the Monte Carlo one only asimptotically.

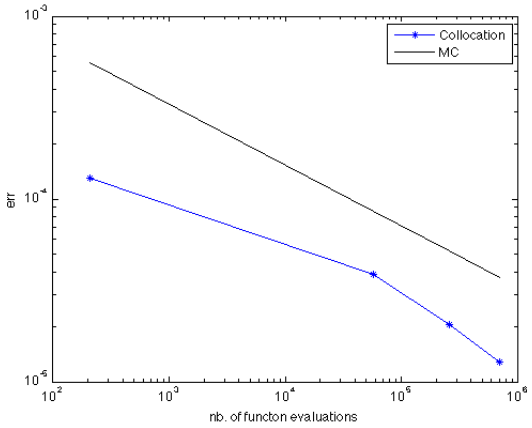


Figure 5.24: Scott, A comparison of the costs for different approximation techniques

### 5.6.3 Heston model

We now focus on the square root process which represents the Heston model of volatility. For different values of $N$ we will evaluate the weak error $|\mathbb{E}[v(T)] - \mathbb{E}[v^w_{\Delta t}(T)]|$ (just like in Section 5.5) as a function of the above-mentioned computational cost $C$ (see Fig. 5.25).
We will consider the Implicit Milstein scheme (5.30) and the parameters that satisfy the positivity condition of the continuous-time process ($\chi = 0.3$, $\theta = 0.4$, $\epsilon = 0.3$).

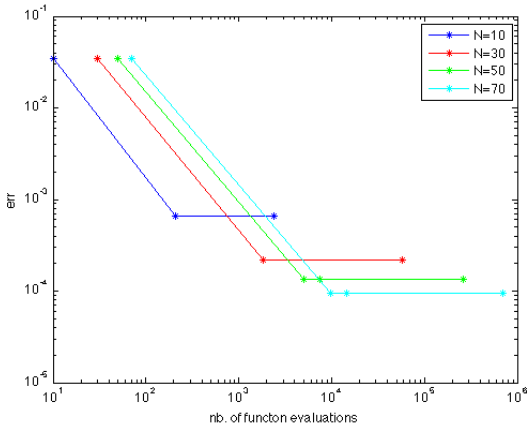


Figure 5.25: Heston, Weak error for Stochastic Collocation method coupled with Implicit Milstein scheme: the error is given by $|\mathbb{E}[v(T)] - \mathbb{E}[v^w_{\Delta t}(T)]|$

The probability approximation convergence is observable until the time discretization error related to the Milstein scheme becomes dominant. We can fairly state that for every value of $N$ the cost corrisponding to the level $w = 2$ saturates the probability error (see Fig. 5.12).
For Monte Carlo simulations, given a value of $N$ and $M$ we repeat the same simulation 20 times and get thus a mean value of $|\mathbb{E}[v(T)] - \mathbb{E}[v^w_{\Delta t}(T)]|$ (Fig. 5.26). In order to make a

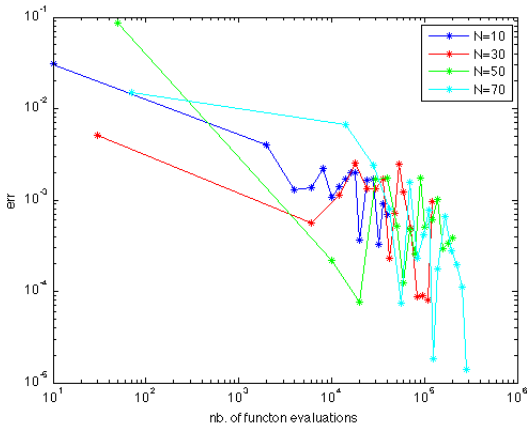


Figure 5.26: Heston, Weak error for Stochastic Collocation method coupled with Implicit Milstein scheme: the error is given by $|\mathbb{E}[v(T)] - \mathbb{E}[v^w_{\Delta t}(T)]|$

comparison between Stochastic Collocation and Monte Carlo method, we also provide the graph in Fig. 5.27 for the value $N = 32$ (Monte Carlo results are related to a single run). It is evident that the first method outperforms the latter for almost all the values of $M$.

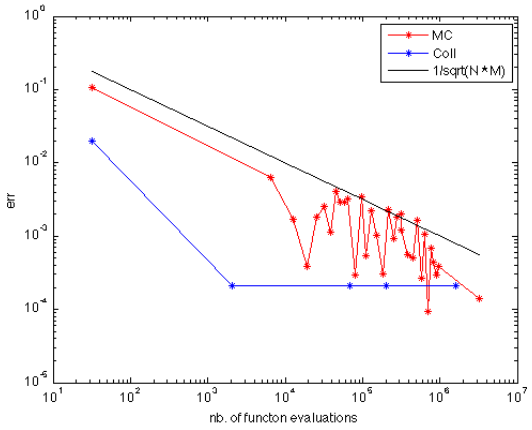


Figure 5.27: Heston, Weak error for Stochastic Collocation and Monte Carlo methods coupled with Implicit Milstein scheme: the error is given by $|\mathbb{E}[v(T)] - \mathbb{E}[v^w_{\Delta t}(T)]|$, N=32

Also the minimum of all curves in Fig. 5.25 is well interpolated by a straight line of slope $-0.5$ hence leading to an error versus cost relation $\epsilon \sim \sqrt{C}$ which is superior to Monte Carlo and comparable to Quasi-Monte Carlo and Multi-level Monte Carlo.

We have thus proved that the Stochastic Collocation method provides excellent results at least when the weak error is evaluated with simple functions ($g(x) = x^2$ for Hull-White model and $g(x) = x$ for Heston model). In fact with both Euler and Milstein schemes the relationship between error and computational costs seems to be very similar to the well-known Quasi-Monte Carlo and Multi-level Monte Carlo techniques. For Scott model where $g(x) = e^x$ further analysis could be necessary.

## 5.7 Brownian Bridge discretization

We have considered so far a random walk discretization for the Wiener paths:

$$W(t_{n+1}) = W(t_n) + \sqrt{\Delta t}\eta_{n+1} \tag{5.34}$$

where $W(0) = 0$ and $\eta_{n+1} \sim N(0,1), n = 0 \dots N-1$. With this scheme all the variables have the same variances $\Delta t$ and the natural choice in Stochastic Collocation is represented by an isotropic grid.
A different approach has been proposed by Caflish [28] and Griebel [31] and consists in genereting the path using a present and a past value, typically using a value of $N$ which is a

power of 2. Under this assumption we first generate $W(t_N) = \sqrt{T}\eta_N$ and define $m = \log_2(N)$. For every $i = 1 \dots m$, we also define $dt_i = \frac{N}{2^i}$ and then for every $k = 1 \dots 2^i$ such that $k$ is odd we generate:

$$W(kdt_i\Delta t) = \frac{W((k-1)dt_i\Delta t) + W((k+1)dt_i\Delta t)}{2} + \sqrt{\frac{dt_i\Delta t}{2}}\eta_{kdt_i}. \tag{5.35}$$

This is usually referred to as Brownian Bridge discretization and allows to use anisotropic grids since the total variance is lumped into the first random variables: in fact the variance of the random variables in (5.35) decreases with $i$.
This discretization is related to the Levy representation of the Brownian motion, based on certain polygonal functions (the Schauder functions, see [30]). We first define the Haar functions $H_n : [0,1] \to \mathbb{R}$:

$$H_0(t) = 1 \tag{5.36}$$

$$H_1(t) = \begin{cases} 1 & \text{if} \quad t \in [0, 1/2] \\ -1 & \text{if} \quad t \in (1/2, 1] \end{cases} \tag{5.37}$$

$$H_n(t) = \begin{cases} 2^{i/2} & \text{if} \quad t \in \left[\frac{n-2^i}{2^i}, \frac{n-2^i+1/2}{2^i}\right] \\ -2^{i/2} & \text{if} \quad t \in \left(\frac{n-2^i+1/2}{2^i}, \frac{n-2^i+1}{2^i}\right] \\ 0 & \text{elsewhere} \end{cases} \tag{5.38}$$

if $i \in \mathbb{N}$, $2^i \leq n < 2^{i+1}$.
The system of Schauder functions is obtained integrating the Haar functions on $[0,1]$

$$\tilde{H}_n(t) = \int_0^t H_n(s)ds \qquad n = 1, 2 \dots \tag{5.39}$$

A series representation for a Brownian sample path on $[0,1]$ is then given by

$$W_t = \sum_{n=1}^{\infty} \tilde{H}_n(t)\eta_n \tag{5.40}$$

where the convergence of the series is uniform for $t \in [0,t]$ and $\eta_n$ are standard gaussian variables.
Recent papers focus on dimension adaptive sparse grids integration (see [31],[33] and [34]) with Gauss-Hermite knots. These works are also based on the use of the effective dimension (see [32]), which allows to consider not the nominal dimension $N$ but a truncated dimension $N_t$ obtained through *ANOVA* decomposition.
We will use the Brownian Bridge construction more naively by anisotropic integration (see

Section 2.1) with weights chosen a priori: for example if $N = 64$ we will give a weight equal to 1 to the variable corresponding to $W(N\Delta t)$, weight 2 to the variable corresponding to $W(\frac{N}{2}\Delta t)$, weight 3 to the variables related to $W(\frac{N}{4}\Delta t)$ and $W(\frac{3N}{4}\Delta t)$ and so on until weight equal to 7.
In particular we will apply this construction to the Geometric Brownian Motion SDE with Euler and Milstein approximation:

$$\begin{aligned}\sigma_{n+1} &= \sigma_n[1 + \bar{\xi}\Delta t + \bar{\epsilon}\Delta W_n] \\ \sigma_{n+1} &= \sigma_n[1 + \bar{\xi}\Delta t + \bar{\epsilon}\Delta W_n + \frac{1}{2}\bar{\epsilon}^2(\Delta W_n^2 - \Delta t)]\end{aligned} \tag{5.41}$$

In Fig. 5.28 and 5.29 we compare the "usual" isotropic TD sparse grid approximation based on a random walk construction of the Brownian sample path with the new anisotropic sparse grid approximation based on Brownian Bridge discretization and evaluate the weak error on the second moment $|\mathbb{E}[\sigma(T)^2] - \mathbb{E}[(\sigma^w_{\Delta t}(T))^2]|$ with isotropic TD plus usual random walk construction and anisotropic TD with Brownian Bridge discretization (the parameters are $\sigma_0 = 0.2$, $\bar{\xi} = 0.2$ e $\bar{\epsilon} = 0.1$:

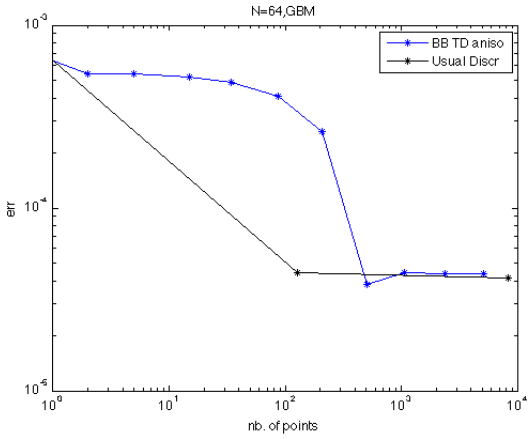


Figure 5.28: GBM, Weak error for TD plus usual random walk construction and anisotropic TD with Brownian Bridge discretization (Euler scheme): the error is given by $|\mathbb{E}[\sigma(T)^2] - \mathbb{E}[(\sigma^w_{\Delta t}(T))^2]|$, N=64

It seems that the Brownian Bridge construction loses his effectiveness when the Euler scheme is applied to the SDE whereas it is very efficient when combined to a Milstein method. Alternatively, (as in [33]), we could consider the Euler scheme applied to the process for $y = log(\sigma)$ and evaluate the weak error $|\mathbb{E}[\sigma(T)^2] - \mathbb{E}[\exp(2y^w_{\Delta t}(T))]|$. Fig 5.30 shows the results obtained.

It is evident that this approach is far superior than the usual scheme based on an anisotropic TD grid and a random walk discretization of the Brownian path. Similar results as those of Fig. 5.30 are obtained for the Ornstein-Uhlenbeck process (see Fig. 5.31).

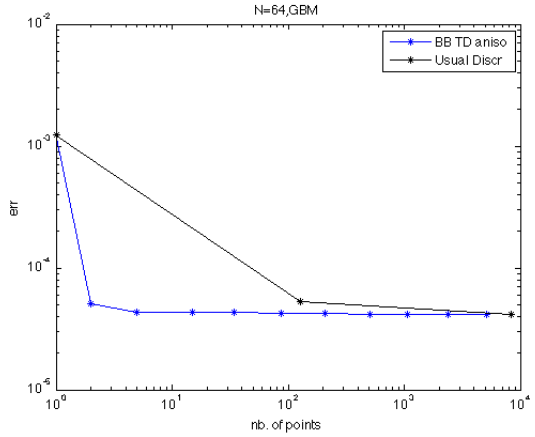


Figure 5.29: GBM, Weak error for TD plus usual random walk construction and anisotropic TD with Brownian Bridge discretization (Milstein scheme): the error is given by $|\mathbb{E}[\sigma(T)^2] - \mathbb{E}[(\sigma^w_{\Delta t}(T))^2]|$, N=64

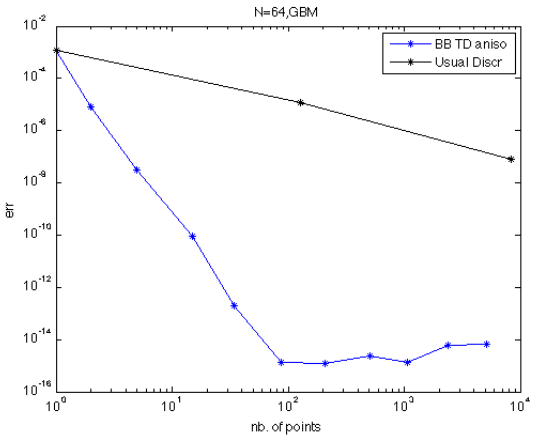


Figure 5.30: GBM, Weak error for TD plus usual random walk construction and anisotropic TD with Brownian Bridge discretization: the error is given by $|\mathbb{E}[\sigma(T)^2] - \mathbb{E}[\exp(2y^w_{\Delta t}(T))]|$, N=64

Considering the Heston model, the Euler scheme with reflection fix and the Implicit Milstein scheme with the Brownian Bridge discretization show the same trends of Fig. 5.28 and 5.29 respectively.

It is still not clear the reason why the Euler scheme on the GBM equation provides such bad results (Fig. 5.28): it seems to be due to the fact that the noise is multiplicative, although for the Milstein scheme the convergence is good. Certainly if the time discrete SDE shows only additive noise, just like in (5.23), then this new type of construction outperforms the usual random walk one (Fig. 5.30 and Fig. 5.31). Further and more sophisticated analysis are necessary since this new method seems to have great potentiality.

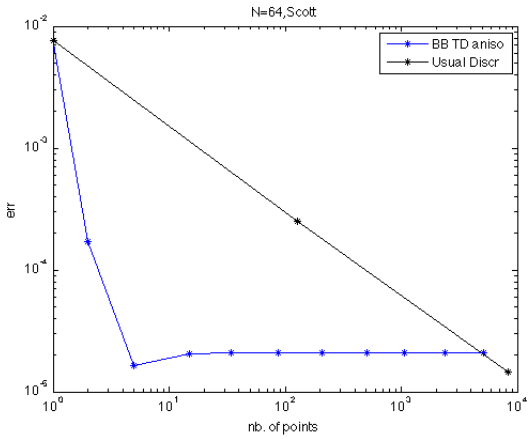


Figure 5.31: GBM, Weak error for TD plus usual random walk construction and anisotropic TD with Brownian Bridge discretization: the error is given by $|\mathbb{E}[\exp(y(T))] - \mathbb{E}[\exp(y^w_{\Delta t}(T)]|$, N=64

# Chapter 6

# Black and Scholes equation conditional to stochastic processes

One of the motivating ideas of this thesis is to solve the option pricing problem by a conditional Black and Scholes equation.
In the first part we will focus on Basket Options: the most known pricing strategies consist in simulating the processes of the underlyings with Monte Carlo method or solving a bidimensional PDE (presented in Section 6.1). With our new approach we introduce a new pricing method based on the approximation of the SDE related to one of the underlying and the solution of one monodimensional "conditional Black and Scholes" equation, which will depend on the simulated process (and the related Brownian motion). We will show that this method can be competitive with respect to the traditional ones, in particular we will make comparison with Monte Carlo methods.
In the second part we will focus instead on option pricing under Heston model of volatility: we will mainly deal with the pricing of European and Barrier vanilla options under the hypothesis of zero correlation between the underlying and the volatility processes using the Stochastic Collocation technique to approximate Heston's SDE. We will also provide a PDE for options pricing in both the uncorrelated and correlated cases. In the latter case the Stochastic Collocation technique might be unefficient compared to more traditional approaches and further analyses or method improvements should be necessary.
Our conditional PDE is based on [7] but there are still theoretical aspects to be investigated, although the numerical results are encouraging.

## 6.1 Basket Call

We will now focus on a European Basket Call written on two underlying whose dynamics are described by two Geometric Brownian motions (under the risk-neutral probability):

$$\begin{cases} dS_1(t) = rS_1(t)dt + \sigma_1 S_1(t)dW_1(t) \\ dS_2(t) = rS_2(t)dt + \sigma_2 S_2(t)dW_2(t) \end{cases} \tag{6.1}$$

with $dW_1(t)dW_2(t) = \rho dt$. The price of every contingent claim $F(t, S_1(t), S_2(t))$ with payoff $\phi(S_1(T), S_2(T))$ written on these underlyings can be obtained as the solution of the following PDE:

$$\begin{cases} \frac{\partial F}{\partial t} + rS_1\frac{\partial F}{\partial S_1} + rS_2\frac{\partial F}{\partial S_2} + \frac{1}{2}\left(S_1^2\sigma_1^2\frac{\partial^2 F}{\partial {S_1}^2} + 2\rho\sigma_1\sigma_2 S_1 S_2\frac{\partial^2 F}{\partial S_1 \partial S_2} + S_2^2\sigma_2^2\frac{\partial^2 F}{\partial {S_2}^2}\right) - rF = 0 \\ F(T, S_1(T), S_2(T)) = \phi(S_1(T), S_2(T)) \end{cases} \tag{6.2}$$

We will consider a Basket Call with payoff $\phi(S_1(T), S_2(T)) = max(S_1(T) + S_2(T) - K, 0)$.

### 6.1.1 Plain Monte Carlo method

In a classical Monte Carlo approach, one simulates the processes in (6.1) by generating M trajectories discretized e.g. by Euler or Milstein method. In particular, one can decouple the two correlated Wiener processes with $W_1(t) = \rho W_2(t) + \sqrt{1-\rho^2}\tilde{W}_1(t)$ and obtain the following discrete approximations related to Euler method

$$\begin{cases} S_{1,j}(t_{n+1}) = S_{1,j}(t_n) + rS_{1,j}(t_n)\Delta t + \sigma_1 S_{1,j}(t_n)(\rho\eta_{n+1,2,j} + \sqrt{1-\rho^2}\eta_{n+1,1,j}) \\ S_{2,j}(t_{n+1}) = S_{2,j}(t_{n+1}) + rS_{2,j}(t_n)\Delta t + \sigma_1 S_{2,j}(t_n)\eta_{n+1,2,j} \end{cases} \tag{6.3}$$

where $\eta_{n+1,k,\cdot}, n = 0 \dots N-1, k = 1, 2$ are uncorrelated standard gaussian variables. The price is given by the mean of the discounted payoff:

$$\mathbb{E}[TV] \simeq e^{-rT}\frac{1}{M}\sum_{j=1}^{M} max(S_{1,j}(t_N) + S_{2,j}(t_N) - K, 0) \tag{6.4}$$

In Table 6.1 we report the prices and the 95% confidence intervals (with the sample variance) obtained with a pure Monte Carlo simulation on (6.1) and the following parameters: $S_1(0) = 100, S_2(0) = 5, K = 105, T = 1, r = 0.05, \sigma_1 = 0.2, \sigma_2 = 0.15$ and $\rho = 0.8$.

| N \ M | $10^5$ | 2 $10^5$ | 3 $10^5$ | 5 $10^5$ |
|---|---|---|---|---|
| 100 | [10.84,10.93,11.02] | [10.71,10.78,10.85] | [10.74,10.79,10.85] | [10.72,10.77,10.81] |
| 200 | [10.80,10.89,10.99] | [10.73,10.80,10.86] | [10.76,10.81,10.87] | [10.75,10.79,10.83] |

Table 6.1: Prices and 95% confidence intervals solving two SDEs with Euler method

It is evident that in order to obtain a good result, a value of $M$ at least equal to 200000 should be chosen and $N = 100$ is also a good choice which is not computationally too expensive.
Instead of solving (6.3), since we are approximating the Geometric Brownian Motion we can even speed up the simulations used in (6.4) by simulating the exact dynamics just like in Chapter 4:

$$\begin{cases} S_{1,j}(t_N) = S_1(0)\exp((r - \frac{1}{2}\sigma_1^2)T + \sigma_1 T(\rho_1\eta_{2,j} + \sqrt{1-\rho^2}\eta_{1,j})) \\ S_{2,j}(t_N) = S_2(0)\exp((r - \frac{1}{2}\sigma_2^2)T + \sigma_2 T\eta_{2,j}) \end{cases} \tag{6.5}$$

With this scheme we get the prices in Table 6.4. In this way we have a faster pricing method,

| M | Prices and confidence intervals |
|---|---|
| 30000 | [10.55,10.72,10.89] |
| 50000 | [10.76,10.89,11.03] |
| 100000 | [10.73,10.82,10.92] |
| 150000 | [10.76,10.84,10.92] |
| 200000 | [10.76,10.83,10.89] |
| 300000 | [10.77,10.83,10.88] |
| 500000 | [10.80,10.84,10.88] |

Table 6.2: Mean Prices and 95% confidence approximating the exact dynamics

while the variances in Table 6.2 vary with $M$ just like in Table 6.1. Notice also that these results indicate that the exact price is probably slightly higher than 10.80.

### 6.1.2 Conditional Black and Scholes with Monte Carlo

The alternative approach that we explore here consists in approximating $M$ times only the second equation of (6.1) with Euler or Milstein method by random samples (Monte Carlo) or sparse grid samples (Stochastic Collocation). The dynamics of the first underlying is kept continuous.

$$dS_1(t) = rS_1(t)dt + \rho\sigma_1 S_1(t)dW_2(t) + \sqrt{1-\rho^2}\sigma_1 S_1(t)d\tilde{W}_1(t). \tag{6.6}$$

We can now recast equation (6.6) on a logarithmic variable $x = log(S_1)$ and obtain

$$dx(t) = rdt - \frac{1}{2}\sigma_1^2 dt + \rho\sigma_1 dW_2(t) + \sigma_1\sqrt{1-\rho^2}d\tilde{W}_1(t). \tag{6.7}$$

Considering intervals $[t_i, t_{i+1}]$ of length $\Delta t$ and having simulated a trajectory of the process $S_2$, the trajectory corresponding to $W_2(t)$ is known as well so we get:

$$dx(t) = (r - \frac{1}{2}\sigma_1^2)dt + \rho\sigma_1\mu_t dt + \sigma_1\sqrt{1-\rho^2}d\tilde{W}_1(t) \tag{6.8}$$

$$\mu_t = \frac{W_2(t_{i+1}) - W_2(t_i)}{\Delta t} \quad for\ t\ in\ [t_i, t_{i+1}) \tag{6.9}$$

where we are considering $dW_2(t) = \dot{W}_2(t)dt$ and approximating the Gaussian white noise $\dot{W}_2(t)$ with (6.9). The Feynman-Kac equation for $u(t,x) = F(t,x,S_2)|_{S_2} = \mathbb{E}[\phi(S_1(T), S_2(T))|S_1(t) = x, S_2(\tau), 0 \leq \tau \leq T]$ (the expected value of the payoff conditional to the knowledge of the second underlying) related to (6.8) is given by the following PDE:

$$\begin{cases} \frac{\partial u}{\partial t} + (r - \frac{1}{2}\sigma_1^2)\frac{\partial u}{\partial x} + \rho\sigma_1\mu_t\frac{\partial u}{\partial x} + \frac{1}{2}(1-\rho^2)\sigma_1^2\frac{\partial^2 u}{\partial x^2} = ru \\ F(x,T) = \phi(x, S_2(T)) \end{cases} \tag{6.10}$$

with payoff varying with each of the M realizations $S_{2,j}(t_N), j = 1 \dots M$ so equation (6.15) can be thought of as a SPDE with random final time condition and random drift.
In (6.10) we have assumed that the brownian motion $W_2(t)$ or more precisely its discrete version $\mu_t$ represents just a drift term[1] as in [7]. We finally remark that we are not sure about the well posedness of (6.10) since a white noise process appears as a drift term.
In our simulations we solve with Euler scheme and Monte Carlo sampling the SDE corresponding to $S_2$ and then solve with the Rannacher method PDE (6.10) on a physical space grid with step $\sim 10^{-3}$. With this method we get the mean prices and 95% confidence intervals (obtained with the sample variance) varying the values of $N$ and $M$ shown in Table 6.3.
Using Milstein scheme to solve the SDE we get analogous results (Table 6.4).
If we assume from Table 6.1 that the exact price is $\sim 10.80$ all these simulations show a relative error close to 1%.

[1]We are not able to provide a rigorous proof of the fact that the process $\mu_t$ does not contribute to the diffusion but only to the drift. Therefore the derivation of PDE (6.15) has to be considered heuristic: it is based mostly on numerical results rather than on a rigorous theoretical proof

| $N, M$ | Prices and confidence intervals |
|---|---|
| $N = 20, M = 20000$ | [10.74,10.90,11.06] |
| $N = 30, M = 20000$ | [10.77,10.93,11.10] |
| $N = 30, M = 25000$ | [10.79,10.94,11.08] |
| $N = 35, M = 15000$ | [10.62,10.81,11] |
| $N = 35, M = 20000$ | [10.71,10.87,11.03] |
| $N = 35, M = 25000$ | [10.74,10.88,11.02] |

Table 6.3: Mean Prices and 95% confidence intervals solving one SDE with Euler scheme and one PDE

| $N, M$ | Prices and confidence intervals |
|---|---|
| $N = 20, M = 20000$ | [10.78,10.95,11.11] |
| $N = 30, M = 20000$ | [10.74,10.90,11.06] |
| $N = 35, M = 20000$ | [10.77,10.93,11.09] |
| $N = 35, M = 25000$ | [10.72,10.87,11.01] |

Table 6.4: Mean Prices and 95% confidence intervals solving one SDE with Milstein scheme and one PDE

From Table 6.1 we can also derive the sampling variances that vary between 195 and 255. The sampling variances from Table 6.3 vary instead between 131 and 146, while the variances from Table 6.4 are included in $[133, 150]$. Therefore we can certainly state that by solving one SDE and one PDE instead of two SDEs we get a mild variance reduction effect.
This reduction however does not allow to enstablish that the PDE strategy is more efficient than plain MC, since the computational time involved in a PDE solution is much higher than a SDE solve. The choice of a spacial step $h$ close to $10^{-3}$ might be too conservative so we have also tried the choices $h = 1/300$, $N = 35$ and $M = 25000$ and got the following price and confidence intervals $[10.74, 10.88, 11.03]$ comparable to those in Tables 6.1 and 6.2. This method in conclusion is still not competitive with the plain MC (Section 6.1.1) since, conversely to these simulations, it takes just a few seconds to get the values in Table 6.2.

### 6.1.3 Conditional Black and Scholes with Stochastic Collocation

We now want to solve (6.10) relying on schemes and results presented in Chapter 5. In particular we have observed a better convergence rate with respect to the weak error using the Brownian Bridge discretization rather than the random walk one. We solve the SDE for $y = log(S_2)$:

$$y_{n+1} = y_n + (r - \frac{1}{2}\sigma_2^2)\Delta t + \sigma_2 \Delta W_{2,n} \tag{6.11}$$

then, on an anisotropic grid with the weights considered in Section 5.7 we plug the value $\exp(y(t_N))$ into the payoff $max(S_1(T) + S_2(T) - K, 0)$. The increments of the Wiener process in (6.11) will also appear in (6.10). With this method we get the results of Table 6.5 computed

with $h = 1/300$:

| w \ N | 16 | 32 | 64 |
|---|---|---|---|
| 0 | 6.77 | 6.77 | 6.77 |
| 1 | 11.36 | 11.36 | 11.36 |
| 2 | 10.68 | 10.66 | 10.66 |
| 3 | 10.92 | 10.89 | 10.88 |
| 4 | 10.89 | 10.83 | 10.81 |
| 5 | 11.01 | 10.88 | 10.85 |
| 6 | 10.94 | 10.93 | 10.87 |
| 7 | 10.98 | 10.90 | 10.93 |
| 8 | 10.96 | 10.92 | 10.90 |
| 9 | 10.97 | 10.91 | N/A |

Table 6.5: Mean Prices solving one SDE and one PDE with Stochastic Collocation (anisotropic TD rule) and Brownian Bridge discretization

To give an idea of the computational costs involved in the simulations which is mainly represented by the number of PDE to be solved, we provide the number of $N$ dimensional points in the sparse grids corresponding to Table 6.5.

| w \ N | 16 | 32 | 64 |
|---|---|---|---|
| 0 | 1 | 1 | 1 |
| 1 | 2 | 2 | 2 |
| 2 | 5 | 5 | 5 |
| 3 | 15 | 15 | 15 |
| 4 | 33 | 33 | 34 |
| 5 | 85 | 85 | 87 |
| 6 | 171 | 203 | 208 |
| 7 | 363 | 427 | 506 |
| 8 | 749 | 909 | 1076 |
| 9 | 1471 | 1951 | N/A |

Table 6.6: Number of Collocation Points (anisotropic TD rule) related to Table 6.5

Notice that considering $N = 64$ and $w = 8$ (1076 Collocation Points) we get an approximation comparable to the one obtained with $M = 150000$ in Table 6.2.

## 6.2 Heston model of volatility: non correlated case

In this section we present the results obtained using the Stochastic Collocation method under the assumption of independent Brownian motions driving the dynamics of the underlying and the volatility for European Vanilla and Barrier Options. The dynamics are given by:

$$\begin{cases} dS(t) = rS(t)dt + \sqrt{v(t)}S(t)dW_1(t) \\ dv(t) = \chi(\theta - v(t))dt + \epsilon\sqrt{v(t)}dW_2(t) \end{cases} \tag{6.12}$$

with $\mathbb{E}[dW_1(t)dW_2(t)] = 0$. A comparison between Euler and Milstein scheme is also presented.

In order to apply the Stochastic Collocation method we consider a discretization of the second SDE with Implicit Milstein scheme (5.30): the maturity of the option T is divided into N time steps of length $\Delta t$ and on each of them we consider this approximation

$$v_{n+1} = \frac{v_n + \chi\theta\Delta t + \epsilon\sqrt{v_n}\eta_{n+1}\sqrt{\Delta t} + 0.25\epsilon^2\Delta t(\eta_{n+1}^2 - 1)}{1 + \chi\Delta t} \tag{6.13}$$

where $\eta_{n+1} \sim N(0,1),\ \forall n = 0 \dots N-1$ is a standard normal variable. We can finally use a PDE to evaluate the price of the option for each realization $\sigma_t^m = \{\sqrt{v_i^m}\}_{i=1}^N$: for each $m$ we solve numerically the BS PDE with respect to the variables S and t conditionally to the volatility realization $\sigma_t^m$

$$\frac{\partial u}{\partial t} + rS\frac{\partial u}{\partial S} + \frac{1}{2}(\sigma_t^m)^2\frac{\partial^2 u}{\partial S^2} = ru \tag{6.14}$$

where $u(x,t) = F(x,t)|_{\sigma_\tau^m} = \mathbb{E}[\phi(S(T))|\sigma_\tau^m, 0 \le \tau \le T]$ is the expected value of the payoff conditional to the knowledge of the volatility realization. In this case we find exactly the usual Black and Scholes equation and the Wiener paths related to $W_2$ do not appear in (6.14).

### 6.2.1 Numerical results

We will now show the efficacy of the Stochastic Collocation under the assumption of zero correlation considering a European and a Barrier Call.

### 6.2.2 European Vanilla

As a first example we consider the price of a European Vanilla call with the following parameters: $S_0 = 100$, $K = 100$, $r = 0.03$, $T = 1$, while Heston parameters are: $\chi = 0.3$, $\epsilon = 0.2$, $\theta = 0.4$, $v_0 = 0.2$. The analytical TV for this option is 19.968719. After dividing the maturity in N intervals we get the prices in Table 6.7 with respect to different values of the level (w) and different numbers of variables

| w \ N | 5 | 10 | 15 | 20 | 25 |
|---|---|---|---|---|---|
| 1 | 20.8948 | 19.5272 | 20.0067 | 19.8370 | 19.8903 |
| 2 | 21.0319 | 19.5784 | 20.0972 | 19.9049 | 19.9696 |
| 3 | 21.0325 | 19.5786 | 20.0973 | 19.9050 | 19.9697 |
| 4 | 21.0325 | 19.5798 | 20.0983 | 19.9063 | 19.9709 |
| 5 | 21.0325 | 19.5798 | 20.0983 | N/A | N/A |
| 6 | 21.0325 | 19.5798 | 20.0983 | N/A | N/A |
| 7 | 21.0325 | 19.5798 | N/A | N/A | N/A |
| 8 | 21.0325 | 19.5798 | N/A | N/A | N/A |

Table 6.7: European Vanilla prices with Stochastic Collocation and Implicit Milstein scheme, $\rho = 0$

Fig. 6.1 and 6.2 show the error as a function of the number of variables and levels (6.1) and the number of N dimensional probability points involved in the approximations (6.2).

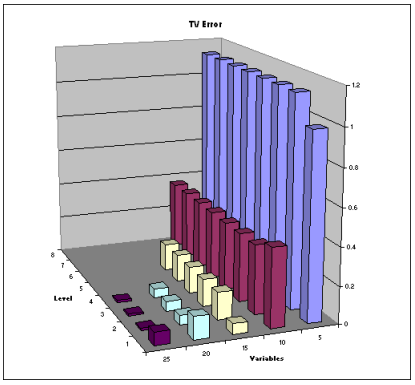


Figure 6.1: TV error vs Number of variables and Level

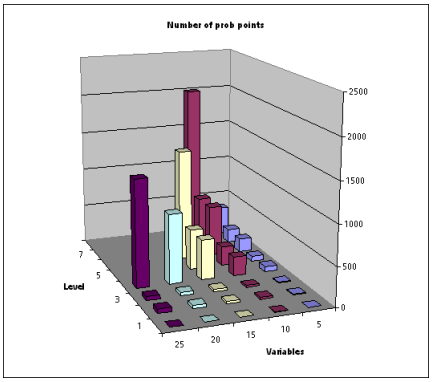


Figure 6.2: Number of probability points vs Number of variables and Level

The charts in Fig. 6.1 and 6.2 show the convergence of Stochastic Collocation method with respect to the number of variables used for the computation (N) and the level, which is related to the number of knots corresponding to each variable. It also follows that one of best choices is represented by $N = 25$ and $w = 2, 3$: for $w = 3$ in fact we have to solve just 51 BS equation with a time step equal to $1/25$ in order to get a relative error close to 0.005%. This conclusion is consistent with the results presented in Chapter 5 and Fig. 5.14, 5.15: in particular from Fig. 5.14 we have been able to conclude that the level $w = 4$ was the right choice, but we had not taken into account the level $w = 3$, which seems to be the best one. To prove that the results do not depend on the particular value of $S_0$ chosen, we evaluate the price of four out-of-the money and four in-the-money options with level and number of

intervals equal to 3 and 25 respectively (Fig. 6.3)

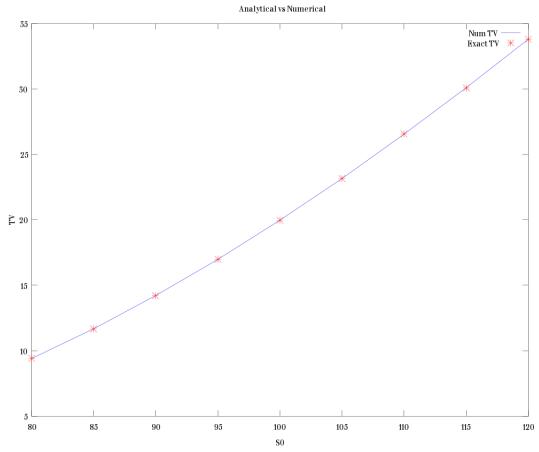


Figure 6.3: TV with Stochastic Collocation method and Implicit Milstein scheme vs spot, $N = 25, w = 3$

The graph shows that the TV computed with these parameters for every spot value are very close to the analytical solutions.

### 6.2.3 European Barrier

As a second example we consider the price of a European Barrier call with the same parameters of the first case and a lower and upper barrier equal to 80 and 160 respectively. In this case a closed-form solution does not exist, so we assume as the exact value the one computed with 200 space points, time step equal to 1/40 and level 4. For this choice of parameters the TV is 1.64963. Using the Stochastic Collocation method, we get the prices in Table 6.8 with respect to different values of the level (w) and different numbers of variables

| w \ N | 5 | 10 | 15 | 20 | 25 |
|---|---|---|---|---|---|
| 1 | 1.5604 | 1.5409 | 1.53719 | 1.53681 | 1.53641 |
| 2 | 1.6876 | 1.6637 | 1.65873 | 1.65787 | 1.6572 |
| 3 | 1.6855 | 1.6627 | 1.65809 | 1.65739 | 1.65682 |
| 4 | 1.6799 | 1.6560 | 1.65108 | 1.65022 | 1.64956 |
| 5 | 1.6799 | 1.6560 | 1.65108 | N/A | N/A |
| 6 | 1.6799 | 1.6560 | 1.65105 | N/A | N/A |
| 7 | 1.6799 | 1.6560 | N/A | N/A | N/A |
| 8 | 1.6798 | 1.6559 | N/A | N/A | N/A |

Table 6.8: European Barrier prices with Stochastic Collocation and Implicit Milstein scheme

The previous values have been obtained not with the standard Crank-Nicholson method but with the Rannacher one [8] since the initial condition is discontinuous (see Appendix). The following graphs show the precentage error as a function of the number of variables and levels (Fig. 6.4) and the number of N dimensional probability points involved in the approximations (Fig. 6.5).

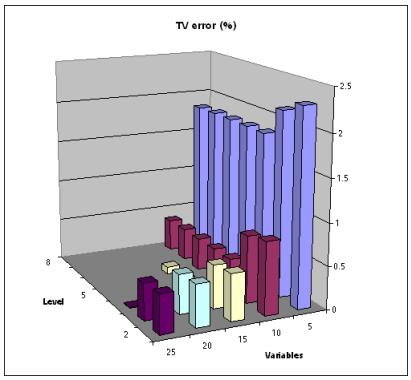


Figure 6.4: TV error vs Number of variables and Level

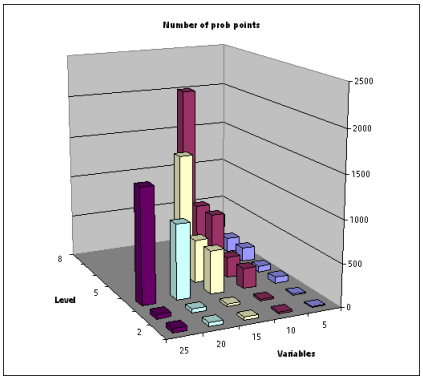


Figure 6.5: Number of probability points vs Number of variables and Level

Also in this case the use of 25 variables and a level equal to 3 provides the right trade off between number of problem to solve and accurancy of the solution (percentage error close to 0.5%).
The Stochastic Collocation method represents therefore an important tool when considering options without an analytical pricing formula.

### 6.2.4 Euler scheme vs Milstein scheme

In Section 6.2, in order to approximate the second equation of (6.12) we have considered the Milstein scheme and thus equation (5.30). In this subsection we make a comparison with Euler scheme and reflection fix presented in the previous chapter. Using this scheme and focusing on the case of the European Vanilla option, we would get the following prices:

| w \ N | 5 | 10 | 15 | 20 | 25 |
|---|---|---|---|---|---|
| 1 | 21.193563 | 19.71509 | 20.216852 | 20.02936 | 20.084921 |
| 2 | 21.097584 | 19.600228 | 20.115247 | 19.917502 | 19.980398 |
| 3 | 21.096663 | 19.599696 | 20.114948 | 19.917252 | 19.980204 |
| 4 | 21.096665 | 19.600041 | 20.115176 | 19.917644 | 19.98044 |
| 5 | 21.096663 | 19.600041 | 20.115176 | N/A | N/A |
| 6 | 21.096671 | 19.600054 | 20.115182 | N/A | N/A |
| 7 | 21.096671 | 19.600054 | N/A | N/A | N/A |
| 8 | 21.096681 | 19.600078 | N/A | N/A | N/A |

Table 6.9: European Vanilla prices with Stochastic Collocation and Euler scheme plus reflection fix, $\rho = 0$

It is evident that the two schemes provide the same TVs and at least for the purpose of our analysis they are equivalent.

## 6.3 Heston model of volatility: correlated case

In this section we aim at deriving a PDE for the correlated case, which has already been presented in [7] and is very similar to (6.10).
Thus we obtain an equation for $u(x,t) = F(x,t)|_{\sigma_\tau^m} = \mathbb{E}[\phi(S(T))|\sigma_\tau^m, 0 \leq \tau \leq T]$, which is given by

$$\frac{\partial u}{\partial t} + (r - \frac{1}{2}\sigma_t^2)\frac{\partial u}{\partial x} + \rho\sigma_t\mu_t\frac{\partial u}{\partial x} + \frac{1}{2}(1-\rho^2)\sigma_t^2\frac{\partial^2 u}{\partial x^2} = ru \tag{6.15}$$

where we are assuming a piecewise constant volatility over each time interval and $\mu$ defined as in (6.9).

### 6.3.1 Numerical results

In this framework we will not use the random walk discretization on an isotropic grid that would give erratic results. Most likely this kind of construction requires a huge value of $w$ to approximate properly the term $\rho\sigma_t\mu_t$ in (6.15), since both the volatility process and the options prices seem to be well approximated with a moderate value of $w$ (see Section 5.5 and 6.2.1). We will therefore focus on the Brownian Bridge construction and the Implicit Milstein scheme that fits to this kind of discretization more than the Euler schemes as stated in Section 5.7.
With the same parameters of Section 6.2.2 and $\rho = 0.8$ we want to obtain the price of a European Vanilla Call, whose exact price evaluated with Heston formula is 20.21. Solving

(6.15) with a value of $h = 1/300$ we get the following prices

| w \ N | 32 | 64 | 128 |
|---|---|---|---|
| 0 | 8.85 | 8.85 | 8.85 |
| 1 | 22.26 | 22.27 | 22.27 |
| 2 | 22.21 | 22.22 | 22.23 |
| 3 | 22.49 | 22.55 | 22.59 |
| 4 | 22.19 | 22.39 | 22.46 |
| 5 | 21.75 | 22.16 | 22.39 |
| 6 | 20.77 | 21.65 | 22.13 |
| 7 | 20.39 | 20.62 | 21.62 |
| 8 | 20.46 | 20.24 | 20.57 |
| 9 | 20.43 | 20.30 | 20.19 |
| 10 | 20.46 | 20.27 | 20.24 |

Table 6.10: Mean Prices with Stochastic Collocation (anisotropic TD rule) and Brownian Bridge discretization, Implicit Milstein, $\rho = 0.8$

In this case a value of $N = 64$ seems to be necessary and also a level $w \geq 8$: for example, for $\rho = -0.8$ the exact price is given by 19.75 and with $N = 64$ and $w = 7, 8, 9, 10$ we get the following mean prices respectively 19.60, 19.85, 19.79 and 19.82.
In this case for $N = 64$ and $w = 10$, for example, we have to solve 5124 PDEs and obtain a mean price with a relative error $< 0.5\%$. Conversely to the Basket Call where the drift term $\rho\sigma_1\mu_t$ is given by a discretized white noise (multiplied by a constant), in this framework the term $\rho\sigma_t\mu_t$ consists in a white noise multiplied by a process and this term may require a higher value of evaluation points. For the Heston models maybe solving the bidimensional PDE (5.3) represents the most efficient strategy.

# Chapter 7

# Conclusions and possible developments

In this work we have presented many successful applications of the Stochastic Collocation technique to financial problems. In the case of Black and Scholes PDE with random coefficients the results are positive as it was expected, since parabolic SPDEs have already been studied and some convergence results are available.
As far as Stochastic Differential Equations are concerned the usual random walk discretization coupled with sparse grids seems quite effective and the Stochastic Collocation is competitive with the most recent probability approximation methods (e.g. Quasi-Monte Carlo and Multilevel Monte Carlo). The Brownian Bridge discretization appears more trickier and some questions have arised during the writing of this thesis, for example it is still not clear why this new scheme works properly with the Milstein scheme but sometimes not with the Euler one (see Section 5.7).
Furthermore, given $N$, we have been able to define $log_2(N)+1$ orders of approximation (each of them related to a different magnitude of the variance in (5.35)) and set the corresponding weights as $[1,\dots,N+1]$. Letting the level of the sparse grid vary $w=1,2\dots$ we are progressively activating these orders. In recent papers as [32] and [33] through an *ANOVA* approach the authors have been able to define an effective dimension of problem (which is related to the value of $N$ considered). With our choice of the weights, given $N$ we might be able to determine a sort of effective dimension as well. For example, for the GBM equation, simulating the exact dynamics and evaluating $\mathbb{E}[\exp(2y^w_{\Delta t}(T))]$ (see Section 5.7) we notice that the effective dimension is actually 1 in this case (see Fig. 7.1): this is not surprising since the exact solution for the GBM $\mathbb{E}[\sigma(T)] = \sigma_0 e^{(\bar{\xi}-\frac{1}{2}\bar{\epsilon}^2)T+\bar{\epsilon}W(T)}$ can be actually parametrized and simulated with just one variable. The interesting thing is that the error corresponding to different values of $N$ and same $w$ are superimposed and this shows the progressive activation process explained before (unless the machine error order is reached). Notice in fact that for example the lines corresponding to $N=4$ and $N=8$ coincide up to $w=3$, if $w=4$ for

$N = 4$ we are adding points in the same dimensions while for $N = 8$ we are taking into account the variables related to a smaller approximation scale and the two lines diverge as the line corresponding to $N = 8$ still coincides with those corresponding to $N \geq 8$.

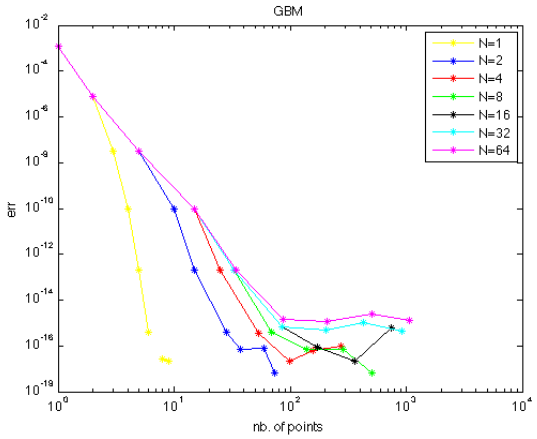


Figure 7.1: GBM, Weak error for anisotropic TD with Brownian Bridge discretization: the error is given by $|\mathbb{E}[\sigma(T)^2] - \mathbb{E}[\exp(2y^w_{\Delta t}(T))]|$, $w = 0 \dots 8$

Conversely evaluating the weak error on the Ornstein-Uhlenbeck SDE this effect is less evident since the time discretization error plays an important role in this case (Fig. 7.2):

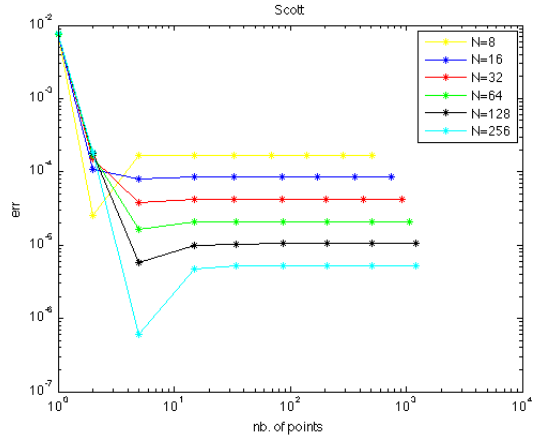


Figure 7.2: Scott, Weak error for anisotropic TD with Brownian Bridge discretization: the error is given by $|\mathbb{E}[\exp(y(T))] - \mathbb{E}[\exp(y^w_{\Delta t}(T)]|$,$w = 0 \dots 8$

The results obtained with the conditional PDE has to be considered also positive, in particular for Basket Options but we want to underline once again that further efforts have to be done in the direction of well posedness results. It is not even clear why the discretized Gaussian White Noise does not contribute to the diffusion coefficient in (6.10) and (6.15).
As a further development we remark that the choice of $N$ and $w$ can be related when the Brownian Bridge construction is used. As shown in Fig. 7.1, given $N$, increasing $w$ there is

a sort of progressive activation process of the smaller approximation scales; for this reason a possible rule of thumb might consist in choosing $w = log_2(N) + 1$, unless above this value the time discretization error becomes dominant. Actually in Table 6.10 the right choice seems to be $w = log_2(N) + 2$, but obviously a more accurate analysis is necessary.

# Appendix A

# Rannacher method

We consider a convection-diffusion equation:

$$\begin{aligned} &\frac{\partial u}{\partial t} + Au = f \qquad \text{in} \qquad \Omega \times [0, \text{inf}] \\ &A = -div(\alpha \nabla) + \beta \cdot \nabla + \gamma \end{aligned} \tag{A.1}$$

with initial condition

$$u|_{t=0} = u_0 \tag{A.2}$$

and boundary condition

$$u|_\Gamma = d, \qquad \alpha \partial_\nu + \delta u = g \tag{A.3}$$

If we assume that $\alpha > 0$ on $\Omega$, all coefficients are regular and the initial value is "rough" the initial-boundary value problem (A.1)-(A.3), has unique solution $u$ which is regular for all $t > 0$, while at $t = 0$ it exhibits a certain singular behavior. Bacause of this local loss of regularity, a certain decrease of accurancy should be expected using the standard discretization methods. We will focus on $A$-stable single step Padé schemes, combined with a Galerkin scheme for the spatial variable. Using a finite elements discretization we get the following finite dimensional problem

$$\frac{du_h}{dt} + A_h u_h = F_h \qquad t \in (0, \infty), \qquad u_h(0) = u_h^0 \tag{A.4}$$

If we consider the homogeneous problem, as in a classical Black and Scholes equation, we can recast our problem into an initial value problem for a linear system of ODE with solution given by:

$$u_h(t) = e^{-tA_h} u_h(0) \tag{A.5}$$

which satisfies a two terms recurrence formula between the time $t_n$, $t_n + k$

$$u_h(t_n + k) = e^{-kA_h} u_h(t_n) \tag{A.6}$$

Tha main idea consists in approximating the exponential function between two time steps with a rational function:

$$u_h(t_{n+1}) = R(kA_h)u_h(t_n) \tag{A.7}$$

where the functions $R(kA_h)$ are rational functions bounded on the spectrum of $kA_h$, uniformly in $k$.

Many technicals details and possible choices of $R$ are given in ([**?**]). We will focus on (m,n)-Padé approximations of $e^{-z}$. We define $R_{m,n}(z) := P_{m,n}(z)/Q_{m,n}(z)$ where

$$\begin{aligned} P_{m,n}(z) &= \sum_{j=0}^{m} \frac{(m+n-j)!m!}{(m+n)!j!(m-j)!}(-z)^j \\ Q_{m,n}(z) &= \sum_{j=0}^{n} \frac{(m+n-j)!n!}{(m+n)!j!(n-j)!}(-z)^j \end{aligned} \tag{A.8}$$

These schemes have the property that $R_{m,n}(z) = e^z + O(|z|^{m+n+1})$ as $z \to 0$, $z \in \mathbb{C}$ which means that the global discretization error has the order $m+n$. A classical choice is represented by $m = n$ but these schemes are just $A$-stable, while those with $n < m$ are strongly $A$-stable. Since $A$-stable schemes don't always damp the error in the computation conversely to those strongly $A$-stable, we expect that Padé schemes with $n = m$ propagate the high frequency errors generated by local irregularities of the data. The following Padé approximations are the most used, though we will focus on the first two:

$$\begin{aligned} R_{01}(z) &= (1-z)^{-1} \qquad \text{Backward Euler Scheme} \\ R_{11}(z) &= (1-\frac{1}{2}z)^{-1}(1+\frac{1}{2}z) \qquad \text{Crank Nicolson scheme} \\ R_{12}(z) &= (1-\frac{2}{3}z+\frac{1}{6}z^2)^{-1}(1+\frac{1}{3}z) \\ R_{22}(z) &= (1-\frac{1}{2}z+\frac{1}{12}z^2)^{-1}(1+\frac{1}{2}z+\frac{1}{12}z^2) \end{aligned} \tag{A.9}$$

In the general case ($f \neq 0$), let $U^n_{h,k}$ be the fully discretized solution, then the Padé schemes with $n = m = 1$ or $n = m = 2$ can be written as follows:

$$(I + \frac{k}{2}A_h)U^n_{h,k} = (I - \frac{k}{2}A_h)U^{n-1}_{h,k} + \frac{k}{2}(F^n_h + F^{n-1}_h) \tag{A.10}$$

$$(I+\frac{k}{2}A_h+\frac{k^2}{12}A_h^2)U^n_{h,k} = (I-\frac{k}{2}A_h+\frac{k^2}{12}A_h^2)U^{n-1}_{h,k}+\frac{k}{2}(F^n_h+F^{n-1}_h)+\frac{k^2}{12}(A_hF^n_h-F^n_{ht}-A_hF^{n-1}_h+F^{n-1}_h) \tag{A.11}$$

where $F^n_h = F_h(t_n)$, $F_h$ is the $L^2$ projection of $F$ on the Galerkin space and $F^n_{ht} = \partial_t F_h(t_n)$. If the data are all regular we expect the usual convergence rate with respect to the time step:

a quadratic convergence for Crank-Nicolson method and a fourth order convergence for the last presented scheme (plus the usual convergence rate with respect to the physical space approximation). For rough data instead the convergence may dramatically decrease in the extreme cases. It's possible to recover the usual convergence rate thanks to the following result:

*Let $U^0_{h,k}$ be the $L^2$-projection of the initial datum $u^0$ onto the finite dimensional space and suppose that, at $t = 0$, $2m$ of the diagonal $(m, m)$-Padé steps are replaced by subdiagonal $(m-1, m)$-Padé steps. Then the approximation $U^n_{h,k}$ converges to the exact solution $u(t_n)$ with the optimal order, for all times $t_n$ sufficiently far from $t_0$.*

The damping effect of the strongly $A$-stable $(m-1, m)$-Padé schemes is able to counter the high frequancy errors without diminishing the global order of the apporximation, since the subdiagonal scheme is used just in a finite number of steps. Often the time advancement is performed using an LU-factorization of the matrix related to the operator $Q_{mn}$ and the dumping procedure would increase the computational cost. However applying a Crank-Nicolson scheme we may use the subdiagonal scheme (backward Euler) for the first two steps with half the time step $k$, thanks to the identity:

$$Q_{01}(-\frac{k}{2}A_h) = I + \frac{k}{2}A_h = Q_{11}(-kA_h) \tag{A.12}$$

Therefore the above-mentioned increase can be avoided and the dumping procedure is equivalent to computing the average:

$$\bar{U}^1_{h,k} = \frac{1}{4}(U^0_{h,k} + 2U^1_{h,k} + U^2_{h,k}) \tag{A.13}$$

for the Crank-Nicolson solution.

We can also consider non costant data in (A.1)-(A.3). In this case we assume that some of the coefficients/boundary values are constant functions up to a finite number of jump times at which they are discontinuous in time. The same optimality convergence result previously stated holds, provided that the $2m$ dumping steps are applied after every jump time.
In the non autonomous case, A=A(t), the convergence result holds has been proven for at least the Crank-Nicolson scheme. There are examples of non autonomous cases with $\alpha = \alpha(t)$ that violate the usual convergence result once a high order Padé scheme is applied (order 3 or higher). If just the convection or the reaction terms are time-dependent, higher order schemes are still optimals. For further details see ([8]).

# Bibliography


[1] I. Babuska, F. Nobile, R. Tempone *A Stochastic Collocation Method for Elliptic Partial Differential Equations with Random Input Data*, SIAM Review, Volume 52, Issue 2, pp. 317-355, 2010.

[2] F.Nobile, R.Tempone *Analysis and implementation issues for the numerical approximation of parabolic equations with random coefficients*, International Journal for Numerical Methods in Engineering, 2009, vol. 80/6-7, pp. 979-1006.

[3] I. Babuska, F. Nobile, R. Tempone *A Stochastic Collocation Method for Elliptic Partial Differential Equations with Random Input Data*, SIAM J. Num. Anal., 2007, vol. 45/3, pp. 1005–1034.

[4] J. Back, F. Nobile, L. Tamellini, R. Tempone, *Stochastic spectral Galerkin and collocation methods for PDEs with random coefficients: a numerical comparison, in Spectral and High Order Methods for Partial Differential Equations*, Lecture Notes in Computational Science and Engineering, Volume 76, pp. 43-62, 2011.

[5] F. Nobile, R. Tempone, C. G. Webster, *A sparse grid stochastic collocation method for partial differential equations with random input data* SIAM Journal, Vol. 46, No.5, pp. 2309-2345

[6] S. Heston *A closed form solution for options with stochastic volatility with application to bond and currency options*, Review with Financial studies, pp.327-343, 1993.

[7] G. Loeper, O. Pironneau *A mixed PDE/Monte-Carlo Method for Stochastic Volatility Models*, Comptes Rendus de l'Academie des Sciences C.R.A.S. Vol. 347 - N 9-10 - p. 559-563, 2009.

[8] R. Rannacher *Finite element solution of diffusion problems with irregular data*, Numerische Mathematik, 43:309-327, 1984.

[9] R. Lord, R. Koekkoek, D. van Dijk *A comparison of biased simulation schemes for stochastic volatility models* working paper, Erasmus University Rotterdam, Rabobank International and Robeco Alternative Investments.

[10] P.E. Kloeden , E. Platen *Numerical solution of stochastic differential equations*, $1^{st}$ edition, Springer Verlag, New York.

[11] D.J. Higham, X. Mao *Convergence of the Monte Carlo simulations involving the mean reverting square root process*, Journal of Computational Finance, vol.8, no.3, pp.35-62, 2005.

[12] A. Berkaoui, M. Bossy, A. Diop *Euler schemes for SDEs with non-Lipschitz diffusion coefficient: strong convergence*, INRIA working paper no. 5637, 2006.

[13] M. Bossy, A. Diop *An efficient discretization scheme for one dimensional SDEs with a diffusion coefficient of the form* $|x|^{\alpha}, \alpha \in [1/2, 1)$, INRIA working paper no. 5396, 2004.

[14] G. Deelstra, F. Delbaen *Convergence of discretized stochastic (interest rate) processes with stochastic drift term*, Applied stochastic models and Data analysis, 1998, Volume: 14, no. 1, pp. 77-84.

[15] L. Andersen, P. Jackel, C. Kahl *Simulation of square root processes*, Enciclopedia of Quantitative Finance John Wiley and Sons, 2010.

[16] C. Kahl, H. Schurz *Balanced Milstein methods for Ordinary SDEs* Monte Carlo Methods and Applications, Volume: 12, Issue: 2, Pages: 143-170, 2006.

[17] D. Higham, X. Mao, A. Stuart *Strong convergence of Euler-type methods for nonlinear stochastic differential equations* SIAM J. Num. Anal., Volume 40, Issue 3, pp. 1041-1063, 2002.

[18] O. P. Le Maitre, O. M. Knio *Spectral Methods for Uncertainty Quantification*, Springer, 2010.

[19] A. Quarteroni *Modellistica numerica per problemi differenziali*, $3^{rd}$ edition, Springer, 2006.

[20] J. Hull, A. White *The pricing of Options on Assets with Stochastic volatility*, The Journal of Finance, Volume: 42, Issue: 2. pp. 281-300, 1987.

[21] L. Scott. *Option Pricing When the Variance Changes Randomly: Theory, Estimation and An Application* Journal of Financial and Quantitative Analysis, 22:419 438, 1987.

[22] M. Chesney, L. Scott. *Pricing European Currency Options: A comparison of the modified Black-Scholes model and a random variance model* Journal of Financial and Quantitative Analysis, 24:267 284, 1989.

[23] S. A. Smolyak *Quadrature and interpolation formulas for tensor products of certain classes of functions*, Dokl. Akad. Nauk SSSR, 4 (1963), pp. 240 243.

[24] P. Wilmott, *Derivatives. The Theory and Practice of Financial Engineering* chap. 23, 1998, pp. 299 304 (John Wiley Sons: Chichester).

[25] A. Alfonsi *On the discretization schemes for the CIR (and Bessel squared) processes*, CERMICS report 2005-279, 2005.

[26] R. Pulch, C. van Emmerich *Polynomial chaos for simulating random volatilities*, Mathematics and Computers in Simulation, 2009, Volume: 80, pp. 245 - 255.

[27] L. Andersen *Efficient Simulation of the Heston Stochastic Volatility Model*, Banc of America Securities, Research Paper, 2007.

[28] R. Caflisch *Monte Carlo and Quasi-Monte Carlo methods*, Acta Numerica vol.7, pp.1-49, Cambridge University Press, 1998.

[29] M.B. Giles *Multilevel Monte Carlo path simulation*, Operations Research 56: 981 986, 2008.

[30] T. Mikosch *Elementary Stochastic Calculus with Finance in view*, World Scientific Publishing Company, 1999.

[31] H.-J. Bungartz, M. Griebel *Sparse grids*, Acta Numerica vol.13 ,pp.1-121, Cambridge University Press, 2004.

[32] X. Wang, K.-T. Fang. *The effective dimension and quasi-Monte Carlo integration*, J. Complexity, 19(2):101 124, 2003.

[33] T. Gerstner, M. Griebel, M. Holtz *Efficient deterministic numerical simulation of stochastic asset-liability management models in life insurance*, Insurance: Math. Economics, 44:434-446, 2009

[34] T. Grestner, M. Griebel *Dimension adaptive tensor product quadrature*, Computing, 71: 6587, 2003.

[35] Feller, W. *Two singular diffusion problems*, Annals of Mathematics 54, 173182, 1951.