UNIVERSITÀ DEGLI STUDI DI TRIESTE

XXVIII CICLO DEL DOTTORATO DI RICERCA IN

ASSICURAZIONE E FINANZA: MATEMATICA E GESTIONE

PRICING AND HEDGING GLWB AND GMWB IN THE HESTON AND IN THE BLACK-SCHOLES WITH STOCHASTIC INTEREST RATE MODELS

Settore scientifico-disciplinare: SECS-S/06

DOTTORANDO
DR. ANDREA MOLENT

COORDINATORE
PROF. GIANNI BOSI

SUPERVISORE DI TESI
PROF. ANTONINO ZANETTE

ANNO ACCADEMICO 2014 / 2015

Contents

Introduction

In 2008 following the subprime crisis, financial markets have suffered the upheavals that have affected the entire world economy. Since then, these markets were extremely volatile: this situation could last a while and perhaps become the new standard. After many failures, the gap between the different interest rates applied to different transmitters has become larger and larger and a discussion on the identification of the risk-free rate is open. The ECB and the Fed's rates gradually declined, while the rate on sovereign debt increased gradually, and then dropped.

For customers, it is difficult to balance risk and return. In this context, clients seek protection for their savings, and the ability to take advantage of the positive changes in the market. With regard to social problematic, following the increase in life expectancy, annuities for retirement dropped.

The mission of insurance companies is to answer the request for protection and compensation of their customers. The solution is to provide the customer an investment account and cover its value with guarantees. These products are called Variable Annuities. In the words of François Robinet, CEO of AXA Life Invest, "*These products, unit of account guaranteed will become a solution to solve the long-term investment problems with security, and prepare for retirement*". Variable Annuities are insurance life contracts in account units with guaranteed revenues or capital. They were launched by AXA in the United States in 1995, and appeared in Italy in 2008. Variable Annuities are mainly diffused in USA, Japan, and North Europe; they are attractive products especially with the retirement reforms and the new sales may reach $22 billion by 2018, with a 57% increase from 2012 (*Think Advisor*, 2014). Variable Annuities are nevertheless exotic products with hidden options (the main one is the lapse one) and two kinds of risk: the market risk and the actuarial risk. Because of these characteristics, it's difficult to price these new kinds of optional products, and manage the ingrained risks.

Among all the Variable Annuities types, there are two ones that are particularly relevant because they are both the most required by the customers and, at the same time, the hardest to be priced: GLWB (*Guaranteed Lifelong Withdrawal Benefit*) and GMWB (*Guaranteed Minimum Withdrawal Benefit*).

In this PhD. thesis I present my research results about Variable Annuities pricing and hedging. I personally did a six months internship at *AXA Life Invest* (see [48]), and during that time I worked in the Risk Management Team. In such a work place I could appreciate how pricing and hedging problems are real problems daily solved by the employees. According to this, research targets of my PhD. have been chosen thinking to practical relevance.

The main contribution of my research has been the development of efficient numerical methods to extends GLWB and GMWB pricing and Greeks calculation to stochastic models including stochastic volatility (the Heston model) or stochastic interest rate (the Black-Scholes Hull-White model).

The thesis is based on two research papers (see [23] and [24]), available at arXiv website. These papers are joint works with Professor Antonino Zanette (my thesis supervisor), and PhD. Ludovic Goudenège (Féderation de Mathématiques de l'École Centrale Paris).

In these papers, we priced GLWB and GMWB guarantees, and we found the no-arbitrage fee, in the Heston model and the Black-Scholes with stochastic interest rate model. First, we treated a static withdrawal strategy: the policy holder (hereafter *PH*) withdraws at the contract rate. Then, taking the point of view of the worst case for the hedger, we priced the guarantees assuming that the PH follows an optimal withdrawal strategy. We also used these methods to calculate the Greeks for hedging and Risk Management. Moreover, in the GLWB case, we performed a mortality shock useful in risk management framework.

To achieve these targets, we developed four numerical methods: a hybrid tree-finite difference method and a Hybrid Monte Carlo method (both introduced by Briani et al. [10]), an ADI finite difference scheme (Haentjens and Hout [25]), and a Standard Monte Carlo method with Longstaff-Schwartz least squares regression (Longstaff and Schwartz [33]).

The main results of our research papers are:

- We formulated the determination of the no-arbitrage fee (i.e. the cost of maintaining a replicating hedging portfolio) in the Heston model and in the Black-Scholes Hull-White model using different pricing methods;
- We presented the effects of stochastic volatility and stochastic interest rate on pricing and Greeks calculation, and the sensitivity of the GLWB and GMWB fee to various modeling parameters;
- We used different numerical methods to price the GLWB and GMWB contract;
- We presented numerical examples which show the convergence of these methods.

My personal contributions to these papers are:

- Concept and development of quadrinomial trees (important on long maturities products to preserve convergence of hybrid methods);
- Improvement of MC Hybrid method (improvement of convergence through the introduction of a splitting method);
- Improvement of the PDE Hybrid method (improvement of convergence through the use of quadrinomial trees and reduction of the computation time by cutting methods and general remarks on the method);
- Use of splines for GMWB (useful to reduce the number of points used in the grid);
- Concept and development of the methods "*Regression by Lines*" and "*Full Regression*" to price GMWB with Monte Carlo methods;

- Use of "*Similarity Reduction*" property in GMWB pricing;
- C++ coding of the numerical methods (excepts ADI method).

I will now present a more detailed summary of the models and numerical methods introduced in the thesis.

The GLWB and GMWB Variable Annuities

The Variable Annuity on which I focused my research work are of type GLWB and GMWB. About this kind of products, a reference article is that of Bacinello and al. [5]: in this paper they classified main GMxBs Variable Annuities, and they computed and compared contract values and fair fee rates under "static" and "mixed" valuation approaches, via ordinary and least squares Monte Carlo methods, respectively. We briefly see how these products work.

GLWB

These insurance products have been analyzed by several authors. among them, we cite the work of Forsyth and Vetzal [13], that performed respectively pricing of GLWB and GMWB Variable Annuities in Black and Scholes framework via PDE methods. They considered both "static" and "dynamic withdrawals", pricing different versions of the products.

The PH who buys a GLWB policy pays at time $t = 0$ an initial gross premium GP from which can be deducted some entry fees, resulting the net premium P. That amount is fully invested in a Investment Fund, and its value is indicated by S_t. The state parameters of the policy are essentially two: the *account value* A_t and the *base benefit* B_t. These parameters are initialized in the following way: $A_0 = P$, $B_0 = GP$. The account value evolves over time according to the Investment Fund

$$dA_t = \frac{A_t}{S_t} dS_t - \alpha_{tot} A_t dt. \tag{0.0.1}$$

The parameter α_{tot} defines the amount (*fees*) that is subtracted from the account value and continuously used by the insurance company to cover the guarantees of the product. Our goal was to calculate the fair value of this parameter, that expresses the cost to guarantee coverage. The contract defines a set of dates (usually at each anniversary of the start of the contract) called *event times*, when the following procedure activates:

1. if the fees have not been taken continuously (variant of the standard contract) they are picked;
2. if the insured died in the previous year, the heirs of the insured receives a guaranteed capital (death benefits) usually defined as the residual value of the account value;
3. If the insured is alive, he (she) is entitled to withdraw a minimum guaranteed sum, defined as a fixed percentage of the base benefits B_t;
4. The contract may provide for mechanisms (roll-up or ratchets) which, under certain circumstances, increase the base benefits.

The activation or not of the previous mechanisms is subject to the fact that the insured is or is not alive; the mortality of insured persons has been modeled according to a function of intensity of mortality $\mathcal{M} : [0, T] \to \mathbb{R}$ such that the probability for a insured person to die in the time lag $[t, t + dt]$ is equal to $\mathcal{M}(t)\, dt$.

In the static version of the product, the still alive PH withdraws at each event time t_i a fixed amount $W_{t_i} = G\Delta t \cdot B^{2+}_{t_i}$, where G is a constant specified by the contract. In the dynamic case, however, the withdrawn amount is characterized by a parameter $\gamma_i \in [0, 2]$ chosen by the PH and it is worth $W_{t_i} = \gamma_i G\Delta t \cdot B^{2+}_{t_i}$. In correspondence with a withdrawal of the PH, the state parameters A_t and B_t change according to γ_i. We denote by the superscript (2+) the state parameters before the withdrawal, and by the superscript (3+) after the withdrawal. Then:

- If $\gamma_i = 0$ the withdrawal in fact does not take place and the contract may provide a bonus:
$$\left(A^{3+}_{t_i}, B^{3+}_{t_i}, t_i\right) = \left(A^{2+}_{t_i}, B^{2+}_{t_i}\left(1 + b_{t_i}\right), t_i\right).$$
- If $0 < \gamma_i \le 1$ the PH withdraws at a lower rate than the standard rate, and the new state variables are
$$\left(A^{3+}_{t_i}, B^{3+}_{t_i}, t_i\right) = \left(\max\left(0, A^{2+}_{t_i} - W_{t_i}\right), B^{2+}_{t_i}, t_i\right).$$
- If $\gamma_i \in \,]1, 2]$ the PH withdraws more than the maximum admitted and some charges may be applied. Moreover, the case $\gamma_i = 2$ corresponds to a total surrender. We define
$$A' = \max\left(0, A^{2+}_{t_i} - G\Delta t \cdot B^{2+}_{t_i}\right).$$
The withdrawn amount is
$$W_{t_i} = G\Delta t \cdot B^{2+}_{t_i} + (\gamma_i - 1)\, A' \left(1 - \kappa_{t_i}\right).$$
where $\kappa_{t_i} \in [0, 1]$ is a penalty for withdrawal above the contract amount. The new state variables are
$$\begin{aligned}\left(A^{3+}_{t_i}, B^{3+}_{t_i}, t_i\right) &= \left(\max\left(0, A^{2+}_{t_i} - G\Delta t \cdot B^{2+}_{t_i} - (\gamma_i - 1)\, A'\right), (2 - \gamma_i)\, B^{2+}_{t_i}, t_i\right) \\ &= \left((2 - \gamma_i)\, A', (2 - \gamma_i)\, B^{2+}_{t_i}, t_i\right).\end{aligned}$$

GMWB

These insurance products have been analyzed by several authors; among them, we cite the work of Chen et Forsyth [13] that performed pricing of GMWB Variable Annuities in Black and Scholes framework via PDE methods. They considered both “static” and “dynamic withdrawals”, pricing different versions of the products.

Another research article is that of Yang and Dai [46] that used a tree based model to price GMWB in Black and Scholes model. They considered both “static withdrawal” and “optimal surrender”.

The products of Chen-Forsyth (GMWB-CF) and Yang-Dai (GMWB-YD) exhibit differences and in my research I have analyzed both versions of these two products. Now let’s see the main features of these contracts.

At the beginning $t = 0$ the customer pays an initial premium P to the insurance company. The premium P is invested in a Investment Fund whose value is S_t. For both the contract types, we suppose that there exists a set of discrete times $\{t_i, i = 1, \dots N\}$, called as in the GLWB case *event times*; at these moments, the customer can make withdrawals. Let's suppose $\Delta t_i = t_{i+1} - t_i$ constant e let's denote it by Δt.

Both contracts could include a death benefit, as for GLWB, but for simplicity we have neglected this aspect as they did in [21]. Both the two contract define an account value A_t that evolves as in equation (0.0.1). We denote by the exponent minus $(-)$ the value of the parameters of state before the withdrawal, and the exponent plus $(+)$ the value of these parameters after that.

Now let's see in detail the two contracts.

GMWB-CF The state parameters of the policy are:

- Account value: $A_t, A_0 = P$.
- Base benefit: $B_t, B_0 = P$.

Both of these variables are initialized by taking the value of the initial premium. We define $T_1 = 0$ the time of the contract beginning, and $T_2 = t_N$ the time of the last possible withdrawal. Usually the first withdrawal takes place in $t_1 = 1\ y$ or $t_1 = 0.5\ y$.

At each event time t_i, the PH withdraws a sum $W_i \in \left[0, B_{t_i}^-\right]$. In reality, the PH doesn't always receives the whole amount withdrawn: if W_i exceeds the guaranteed amount a penalty is applied. The sum actually received by the customer is:

$$f(W_i) = \begin{cases} W_i & \text{if } W_i \leq G \\ W_i - \kappa (W_i - G) & \text{if } W_i > G, \end{cases}$$

while the new state variables are

$$\left(A_{t_i}^+, B_{t_i}^+, t_i\right) = \left(\max\left(A_{t_i}^- - W_i\right), B_{t_i}^- - W_i, t_i\right).$$

At time $T = t_2$ the last withdrawals takes place and the PH gains a final payoff equal to

$$FP = \max\left(A_{T_2}, (1 - \kappa) B_{T_2}\right).$$

GMWB-YD The policy state parameters are:

- Account value: A_t, $A_0 = P$.
- Guaranteed minimum withdrawal: G.

The state variable A_t is initially equal to the premium, while G isn't defined up to time T_1. The payments take place at time $T_1 + \Delta t$ up to T_2. To treat this type of products is not necessary to define the Base Benefit, because its value is constant equal to the premium P until the contract ends. Usually, the contract sets a minimum rate of return i for the deferred period $[0, T_1]$. At time T_1 (if $T_1 > 0$) the account value is reset and the value of G is set:

$$A^+_{T_1} = \max\left[P\,(1+i)^{T_1}, A^-_{T_1}\right], \quad G = \frac{A^+_{T_1}}{m\,(T_2 - T_1)}.$$

For this type of product, the PH withdraws at the contract dates the amount guaranteed G for the duration of the contract. The evolution of state parameters is the following:

$$\left(A^+_{t_i}, G^+, t_i\right) = \left(\max\left(0, A^+_{t_i} - G^-\right), G^-, t_i\right).$$

In the case of the dynamic approach, the PH may terminate the contract early, let's say in $\bar{t}$, collecting a final payoff equal to

$$FP = G + (1-\kappa)\max\left(0, A^-_{\bar{t}} - G\right).$$

The stochastic models for the fund S

We recap the two models for the investment fund S that we treated.

The Heston model

The Heston model [26] is one of the most important models for stochastic volatility. Its dynamics is

$$\begin{cases} dS_t = rS_t dt + \sqrt{v_t} S_t dZ^S_t & S_0 = \bar{S}_0, \\ dv_t = k\,(\theta - v_t)\,dt + \omega\sqrt{v_t} dZ^v_t & v_0 = \bar{v}_0, \end{cases}$$

where Z^S and Z^v are Brownian motions and $d\left\langle Z^S_t, Z^v_t\right\rangle = \rho dt$. We also define $\bar{\rho} = \sqrt{1-\rho^2}$.

The Black-Scholes Hull White model

The Hull and White model [28] is one of the most important models used to describe the stochastic interest rate process. The dynamics of the couple underlying-interest rate is

$$\begin{cases} dS_t = r_t S_t dt + \sigma S_t dZ^S_t & S_0 = \bar{S}_0, \\ dX_t = -kX_t dt + dZ^r_t & X_0 = 0, \\ r_t = \omega X_t + \beta\,(t), \end{cases}$$

where Z^S and Z^r are Brownian motions and $d\left\langle Z^S_t, Z^r_t\right\rangle = \rho dt$. We also define $\bar{\rho} = \sqrt{1-\rho^2}$.

Pricing methods

The pricing methods that we used are four: two Monte Carlo methods and two PDE methods.

Two of these methods are "Hybrid" as they combine trees and MC simulations or trees and PDE. Before presenting a summary of these four methods it is worth spending a few words on the trees used.

Quadrinomial trees

Given the long maturity of the products studied, the classic trees presented in [35] or [4] are not suitable to discretize the stochastic volatility and the stochastic interest rate. We have therefore introduced new quadrinomial trees with the aim to combine exactly the first three moments of the associated processes. In the thesis it is exposed in detail the procedure for the construction of the trees.

The Hybrid Monte Carlo method

This method has been introduced by Briani and al, [11].

The method involves the use of a tree to define a Markov chain for the volatility (respectively the interest rate): using a discrete variable distributed according to the transition probabilities of the Markov chain, a discrete process $\bar{v}$ (resp. $\bar{r}$ and $\bar{X}$) is defined in order to approximate the volatility (resp. the interest rate).

For the Heston model, let $N \sim \mathcal{N}(0,1)$ and $B \sim \mathcal{B}(0.5)$. The value of the process $\bar{S}_{t+\Delta t}$ that simulates the underlying is given by

$$\bar{S}_{t+\Delta t} = \begin{cases} \bar{S}_t \exp\Big[\big(r - \frac{\rho}{\sigma} k\theta\big)\Delta t + \big(\frac{\rho}{\sigma}k - \frac{1}{2}\big)\Big(\frac{\bar{v}_{t+\Delta t}+\bar{v}_t}{2}\Big)\Delta t + \frac{\rho}{\sigma}(\bar{v}_{t+\Delta t} - \bar{v}_t) + \sqrt{(1-\rho^2)\,\Delta t \bar{v}_t}N\Big] & \text{if } B = 0, \\ \bar{S}_t \exp\Big[\big(r - \frac{\rho}{\sigma} k\theta\big)\Delta t + \big(\frac{\rho}{\sigma}k - \frac{1}{2}\big)\Big(\frac{\bar{v}_{t+\Delta t}+\bar{v}_t}{2}\Big)\Delta t + \frac{\rho}{\sigma}(\bar{v}_{t+\Delta t} - \bar{v}_t) + \sqrt{(1-\rho^2)\,\Delta t \bar{v}_{t+\Delta t}}N\Big] & \text{if } B = 1. \end{cases}$$

For the Black-Scholes Hull-White model, we have

$$\bar{S}_{t+\Delta t} = \bar{S}_t \exp\left[\left(\frac{\bar{r}_{t+\Delta t} + \bar{r}_t}{2} - \frac{\sigma^2}{2}\right)\Delta t + \sigma\left(\left(\bar{X}_{t+\Delta t} + \bar{X}_t\,(k\Delta t - 1)\right)\rho + \sqrt{\Delta t}\bar{\rho}N\right)\right].$$

The Standard Monte Carlo method

The Standard Monte Carlo method is defined using the best methods to generate scenarios in the two models that we considered. In the case of the Heston model, the reference method is a third order scheme by Alfonsi [2]. For the Black-Scholes Hull-White model, we have made reference to an exact discretization scheme presented in Ostrovski [39].

The Hybrid PDE method

This method has been presented in Briani [10] and [11]. The method uses a tree to define a Markov chain for the volatility (the interest rate). Then, a one variable PDE (in addition to the time variable) is solved at the tree nodes.

To solve the problem of the correlation between the volatility (or the interest rate) and the underlying, an initial transformation is required. For the Heston model we set

$$Y_t^E = \ln(S_t) - \frac{\rho}{\omega}v_t,\ Y_0^E = \ln(S_0) - \frac{\rho}{\omega}v_0.$$

Then

$$dY_t^E = \left(r - \frac{v_t}{2} - \frac{\rho}{\omega}k\,(\theta - v_t)\right)dt + \bar{\rho}\sqrt{v_t}d\bar{Z}_t^S.$$

The Y^E process can be used easily because it is not correlated with the process v. Freezing the value of v at each node, it can be used to define a PDE that is going to be solved by the tree.

Similarly, the Black-Scholes Hull-White we defined

$$Y_t^U = \ln(S_t) - \rho\sigma X_t,\; Y_0^U = \ln(S_0)\,,$$

that verifies

$$dY_t^U = \left(r_t - \frac{\sigma^2}{2} + \sigma\rho k X_t\right)dt + \sigma\bar{\rho}d\bar{Z}_t^S.$$

The Y^U process can be used easily because it is not correlated with the process X. Freezing the value of X and r at each node, it can be used to define a PDE that is going to be solved by the tree.

Within the GLWB Variable Annuities, a couple of the PDEs for the two models is given by

$$\mathcal{V}_t^{He} + \frac{\bar{\rho}^2\bar{v}}{2}\mathcal{V}_{EE}^{He} + \left(r - \frac{\bar{v}}{2} - \frac{\rho}{\omega}k(\theta - \bar{v}) - \alpha_{tot}\right)\mathcal{V}_E^{He} - r\mathcal{V}^{He} + \alpha_m\mathcal{R}(t)\exp\left(E_t + \frac{\rho}{\omega}\bar{v}\right) = 0\,, \quad \text{(He 2)}$$

$$\mathcal{V}_t^{HW} + \frac{\bar{\rho}^2\sigma^2}{2}\mathcal{V}_{UU}^{HW} + \left(\bar{r} - \frac{\sigma^2}{2} + \sigma\rho k\bar{X} - \alpha_{tot}\right)\mathcal{V}_U^{HW} - \bar{r}\mathcal{V}^{HW} + \alpha_m\mathcal{R}(t)\exp\left(U_t + \rho\sigma\bar{X}\right) = 0\,. \quad \text{(HW 2)}$$

Hybrid PDE method has proved to be very powerful. It features greater simplicity of coding, stability of results and good speed of convergence.

The PDE ADI method

The ADI method is a method introduced in the 1950s by Peaceman and Rachford (see [40]) to solve parabolic PDEs. Since then, the method has been used in many sectors.

It can be proved that, in the Heston model, the value of an option $\mathcal{V}$ verifies the PDE

$$\frac{\partial\mathcal{V}}{\partial t} + \frac{vS^2}{2}\frac{\partial^2\mathcal{V}}{\partial S^2} + \frac{\omega^2 v}{2}\frac{\partial^2\mathcal{V}}{\partial v^2} + rS\frac{\partial\mathcal{V}}{\partial S} + \rho\omega Sv\frac{\partial^2\mathcal{V}}{\partial S\partial v} + k(\theta - v)\frac{\partial\mathcal{V}}{\partial v} - r\mathcal{V} = 0,$$

while in the Black-Scholes Hull-White model

$$\frac{\partial\mathcal{V}}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2\mathcal{V}}{\partial S^2} + \frac{\omega^2}{2}\frac{\partial^2\mathcal{V}}{\partial r^2} + rS\frac{\partial\mathcal{V}}{\partial S} + \rho\omega S\sigma\frac{\partial^2\mathcal{V}}{\partial S\partial r} + k(\theta - r)\frac{\partial\mathcal{V}}{\partial r} - r\mathcal{V} = 0.$$

Such multidimensional equations can be solved using the composition of two schemes that discretize the pair (S, v) (respectively (S, r)) implicitly w.r.t a variable and explicitly to the other.

In particular, for our applications, we used the Douglas scheme with $\theta = 1/2$.

Results

All the four numerical methods have been implemented and used to calculate the fair value of the parameter α_{tot}. The various numerical tests focused on different variations of both products and they have been done fixing the computational time for all methods: working with a fixed time, we were able to compare the quality of the results obtained by the different methods.

For both policies GLWB and GMWB, in the static case, the results were very good with all methods: the values obtained are consistent and they differ from each other in small relative differences. Things have been rather different in the dynamic case. Monte Carlo methods have suffered from the problem of least squares regression: the difficulty of approximating accurately the value function $\mathcal{V}$ in multiple dimensions has meant that the values obtained by Monte Carlo methods were lower than those obtained with the PDE methods (the withdrawal strategy deducted by the MC methods is not fully optimal). These latter, however, proved to be more stable and efficient to address problems of optimal withdrawal. In particular, the Hybrid PDE method provided, in general, the best results.

All results of the several numerical tests are presented in the thesis, and they are often accompanied by graphics that facilitate their reading.

The thesis is organized as follows.

Chapter 1 presents the main numerical methods: Monte Carlo, Trees, PDE, and finally Hybrid methods. I widely used these methods to achieve Variable Annuities pricing and Greeks calculation. Then, Chapter 2 presents the Variable Annuities products: the different types, the products structure and mechanism, the guarantees and the fees. Chapter 3 and 4 are the most relevant since they are devoted to the main research topic: pricing and hedging of GLWB and GMWB in the Heston model and in the Black-Scholes Hull-White model. These two last chapters contain conclusions where the main research targets achieved are pointed out.

To conclude I would thank my supervisor, Prof. Antonino Zanette, who encouraged and guided me in research work. His suggestions and hints were very valuable to me. I would also thank PhD. Ludovic Goudenège for his remarkable tips about LS MC methods and ADI development.

Chapter 1

Numerical Methods in Finance and Insurance

In this Chapter, we describe the main numerical methods that we are going to use to perform Variable Annuities pricing. These methods belong to different type: Monte Carlo, Trees, PDE and finally Hybrid methods.

The purpose of this Chapter is not giving a detailed and rigorous description of the methods, but briefly introduce the concepts that will be used to develop the research matter. The reader interested in more detail will find in the bibliographic references more information and the proofs of the propositions here stated.

1.1 Monte Carlo methods

This part is inspired by [15], [29] and [31].

Monte Carlo methods are extensively used in financial institutions to compute European options prices, to evaluate sensitivities of portfolios to various parameters and to compute risk measurements. Let us describe the principle of the Monte Carlo methods on an elementary example. Let

$$I = \int_{[0,1]^d} f(x)\, dx$$

where $f(\cdot)$ is a bounded real valued function.

We can represent I as $E[f(U)]$, where U is an uniformly distributed random variable on $[0,1]^d$. By the Strong Law of Large numbers, if $(U_i, i \geq 1)$ is a family of uniformly distributes independent random variables on $[0,1]^d$, then the average

$$S_N = \frac{1}{N}\sum_{i=1}^{N} f(U_i) \tag{1.1.1}$$

converges to $E[f(U)]$ almost surely when N tends to infinity. This suggests a very simple algorithm to approximate I: call a random number generator N times and compute the average

(1.1.1). Observe that the method converges for any integrable function on $[0,1]^d$: f is not necessarily a smooth function.

In order to efficiently use the above mentioned Monte Carlo method, we need to know its rate of convergence and to determine when it is more efficient than deterministic algorithms. The Central Limit Theorem provides the asymptotic distribution of $\sqrt{N}\,(S_N - I)$ where N tends to $+\infty$. Various refinements of the Central Limit Theorem, such as Berry-Essen and Bikelis Theorems, provide non asymptotic estimates. The preceding consideration shows that the convergence rate of a Monte Carlo method is rather slow $\left(1/\sqrt{N}\right)$. Moreover, the approximation error is random and may take large values even if N is large (however, the probability of such an event tends to 0 when N tends to infinity). Nevertheless, the Monte Carlo methods are useful in practice, especially when the dimension of the integration domain is high. For instance, consider an integral in a hypercube $[0,1]^d$, with d large ($d = 40$, e.g.). It is clear that the quadrature methods require too many points. A Monte Carlo method does not have such disadvantages: it requires the simulation of independent random vectors $(X_1, \ldots, X_d)$, whose coordinates are independent.

In addition, another advantage of the Monte Carlo methods is their parallel nature: each processor of a parallel computer can be assigned the task of making a random trial.

To summarize the preceding discussion: probabilistic algorithms are used in situations where the deterministic methods are inefficient, especially when the dimension of the state space is very large. Obviously, the approximation error is random and the rate of convergence is slow, but in these cases it is still the best method known.

1.1.1 On the convergence rate of Monte Carlo methods

In this Section we present results which justify the use of Monte Carlo methods and help to choose the appropriate number of simulations N of a Monte Carlo method in terms of the desired accuracy and the confidence interval on the accuracy.

Theorem 1.1 (Strong Law of Large Numbers). *Let $(X_i, i \geq 1)$ be a sequence of independent identically distributed random variables such that $E\left[|X_1|\right] < +\infty$. Then,*

$$\lim_{n \to +\infty} \frac{1}{n}\left(X_1 + \cdots + X_n\right) = E\left[X_1\right] \ a.s.$$

Remark 1.2. The random variables X_1 needs to be integrable. Therefore the Strong Law of Large numbers does not apply when X_1 is Cauchy distributed, that is when its density is $\frac{1}{\pi(1+x^2)}$.

1.1.1.1 Convergence rate

We now seek estimates on the error

$$\epsilon_n = E\left[X\right] - \frac{1}{n}\left(X_1 + \cdots + X_n\right).$$

The Central Limit Theorem precises the asymptotic distribution of $\sqrt{N}\epsilon_N$.

Theorem 1.3 (Central Limit Theorem)**.** *Let $(X_i, i \geq 1)$ be a sequence of independent identically distributed random variables such that $E\left[X_1^2\right] < +\infty$. Let σ^2 denote the variance of X_1, that is*

$$\sigma^2 = E\left[\left(X_1 - E\left[X_1\right]\right)^2\right] = E\left[X_1^2\right] - E\left[X_1\right]^2.$$

Then:

$$\left(\frac{\sqrt{n}}{\sigma}\epsilon_n\right)$$

converges in distribution to G, where G is a Gaussian random variable with mean 0 *and variance* 1.

Remark 1.4. From this theorem it follows that for all $c_1 < c_2$

$$\lim_{n \to +\infty} P\left(\frac{\sigma}{\sqrt{n}}c_1 \leq \epsilon_n \leq \frac{\sigma}{\sqrt{n}}c_2\right) = \int_{c_1}^{c_2} e^{-\frac{x^2}{2}} \frac{dx}{\sqrt{2\pi}}.$$

In practice, we can apply the following approximate rule: for n large enough, the law of ϵ_n is a Gaussian random variable with mean 0 and variance σ^2/n.

Note that it is impossible to bound the error, since the support of any (non degenerate) Gaussian random variable is $\mathbb{R}$. Nevertheless the preceding rule allows us to define a confidence interval: for instance, observe that

$$P\left(|G| \leq 1.96\right) \approx 0.95.$$

Therefore, with a probability closed to 0.95, for n is large enough, we get:

$$|\epsilon_n| \leq 1.96\frac{\sigma}{\sqrt{n}}$$

1.1.1.2 How to estimate the variance

The previous result shows that it is crucial to estimate the standard deviation σ of the random variable. Its easy to do this by using the same samples as for the expectation. Let X be a random variable, and $(X_1, \ldots, X_n)$ a sample drawn along the law of X. We will denote by $\bar{X}_N$ the Monte Carlo estimator of $E\left[X\right]$ given by

$$\bar{X}_N = \frac{1}{N}\sum_{i=1}^{N} X_i.$$

A standard estimator for the variance is given by

$$\bar{\sigma}_N^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(X_i - \bar{X_N}\right)^2$$

and this expression is often called the *empirical variance* of the sample. Note that $\bar{\sigma}_N^2$ can be rewritten as

$$\bar{\sigma}_N^2 = \frac{N}{N-1}\left(\frac{1}{N}\sum_{i=1}^{N}\left(X_i^2\right) - \bar{X_N}^2\right).$$

On this last formula, it is obvious that $\bar{X}_N$ and $\bar{\sigma}_N$ can be computed using only $\sum_{i=1}^N X_i$ and $\sum_{i=1}^N X_i^2$.

Moreover, one can prove that if $E\left[X^2\right] < +\infty$ then $\lim_{N\to+\infty} \bar{\sigma}_N^2 = \sigma^2$, almost surely, and that $E\left[\bar{\sigma}_N^2\right] = \sigma^2$ (the estimator is unbiased). This leads to an (approximate) confidence interval by replacing σ by $\bar{\sigma}_N$ in the standard confidence interval. With a probability near to 0.95, $E\left[X\right]$ belongs to the (random) interval given by

$$\left[\bar{X}_N - \frac{1.96\bar{\sigma}_N}{\sqrt{N}}, \bar{X}_N + \frac{1.96\bar{\sigma}_N}{\sqrt{N}}\right].$$

So, with very little additional computations, (we only have to compute $\bar{\sigma}_N$ on a sample already drawn) we can give an reasonable estimate of the error done by approximating $E\left[X\right]$ with $\bar{X}_N$. The possibility to give an error estimate with a small numerical cost, is a very useful feature of Monte Carlo methods.

1.1.2 Simulation methods of classical laws

The aim of this Section is to give a short introduction to sampling methods used in Finance. Our aim is not to be exhaustive on this broad subject (for this we refer to, e.g., [41] or [42]) but to describe methods needed for the simulation of random variables widely used in Finance. Thus we concentrate on Uniform and Gaussian random variables and Gaussian vectors.

1.1.2.1 Simulation of the Uniform law

In this Section we present basic algorithms producing sequences of "pseudo random numbers", whose statistical properties mimic those of sequences of independent and identically uniformly distributed random variables. For a recent survey on random generators see, for instance, [20] and for mathematical treatment of these problems, see [37] and the references therein.

To generate a deterministic sequence which "looks like" independent random variables uniformly distributed on $[0, 1]$, the simplest (and the most widely used) methods are congruential methods. They are defined through four integers a, b, m and U_0. The integer U_0 is the seed of the generator, m is the order of the congruence, a is the multiplicative term. A pseudo random sequence is obtained from the following inductive formula:

$$U_n = (aU_{n-1} + b) \,(\mathrm{mod}\ m).$$

In practice, the seed is set to U_0 at the beginning of a program and must never be changed inside the program.

Observe that a pseudo random number generator consists of a completely deterministic algorithm. Such an algorithm produces sequences which statistically behaves (almost) like sequences of independent and identically uniformly distributed random variables. There is no theoretical criterion which ensures that a pseudo random number generator is statistically acceptable. Such a property is established on the basis of empirical tests. For example, one builds a sample from successive calls to the generator, and one then applies the Chi–square test or the Kolmogorov–Smirnov test in order to test whether one can reasonably accept the hypothesis that the sample results from independent and uniformly distributed random variables. A generator is good when no severe test has rejected that hypothesis. Good choice for a, b, m are given in [32].

1.1.2.2 Simulation of the Gaussian law

The standard Gaussian law (mean 0 and variance 1) on $\mathbb{R}$ is the law with the density given by

$$\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right).$$

The most widely used simulation method of a Gaussian law is the Box-Muller method. This method is based upon the following result:

Proposition 1.5. *Let U and V be two independent random variables which are uniformly distributed on $[0,1]$. Let X and Y be defined by*

$$X = \sqrt{-2\ln(U)}\sin(2\pi V),$$
$$Y = \sqrt{-2\ln(U)}\cos(2\pi V).$$

Then X and Y are two independent Gaussian random variables with mean 0 and variance 1.

Of course, the method can be used to simulate N independent realizations of the same real Gaussian law. The simulation of the two first realizations is performed by calling a random number generator twice and by computing X and Y as above. Then the generator is called two other times to compute the corresponding two new values of X and Y, which provides two new realizations which are independent and mutually independent of the two first realizations, and so on.

1.1.2.3 Simulation of a Gaussian vector

To simulate a Gaussian vector

$$X = \left(X^1, \ldots, X^d\right)$$

with mean zero and with a $d \times d$ covariance matrix $C = (c_{i,j})_{1 \le i,j \le n}$ with $c_{i,j} = E\left[X^i X^j\right]$ it's possible to proceed as follows.

The matrix C is a covariance matrix, so it is positive (since, for each $v \in \mathbb{R}^d$, $v.Cv = E\left[(v.X)^2\right] \ge 0$). Standard results of linear algebra prove that there exists a $d \times d$ matrix A, called a square root of C such that

$$AA^* = C,$$

where A^* is the transposed matrix of $A = (a_{i,j}, 1 \le i,j \le n)$. Moreover one can compute a square root of a given positive symmetric matrix by specifying that $a_{ij} = 0$ for $i < j$ (i.e. A is a lower triangular matrix). Under this hypothesis, its easy to see that A is uniquely determined by the following algorithm.

Algorithm 1.6 (Cholevsky).

$$a_{11} = \sqrt{c_{11}}$$

For $2 \leq i \leq d$

$$a_{i1} = \frac{c_{i1}}{a_{11}}.$$

Then increasing i *from* 2 *to* d,

$$a_{ii} = \sqrt{c_{ii} - \sum_{j=1}^{i-1} |a_{ij}|^2},$$

For $j < i \leq d$

$$a_{ij} = \frac{c_{ij} - \sum_{k=1}^{j-1} a_{ik} a_{jk}}{a_{jj}},$$

For $1 < i < j$

$$a_{ij} = 0.$$

This way of computing a square root of a positive symmetric matrix is known as the Cholesky algorithm.

Now, if we assume that $G = \left(G^1, \ldots, G^d\right)$ is a vector of independent Gaussian random variables with mean 0 and variance 1 (which are easy to sample as we have already seen), one can check that $Y = AG$ is a Gaussian vector with mean 0 and with covariance matrix given by $AA^* = C$. As X and Y are two Gaussian vectors with the same mean and covariance matrix, the law of X and Y are the same. This leads to the following simulation algorithm.

Algorithm 1.7 (Gaussian vector generation). *Simulate the vector* $\left(G^1, \ldots, G^d\right)$ *of independent Gaussian variables as explained above. Then return the vector* $X = AG$.

1.1.3 A variance reduction method: antithetic variates.

All the results of the preceding lecture show that the ratio $\sigma/\sqrt{N}$ governs the accuracy of a Monte Carlo method with N simulations. An obvious consequence of this fact is that one always has interest to rewrite the quantity to compute as the expectation of a random variable which has a smaller variance: this is the basic idea of variance reduction techniques. The use of antithetic variates is widespread in Monte Carlo simulation. This technique is often efficient but its gains are less dramatic than others.

We begin by considering a simple and instructive example. Let

$$I = \int_0^1 g(x)\, dx.$$

If U follows a uniform law on the interval $[0, 1]$, then $1 - U$ has the same law as U, and thus

$$I = \frac{1}{2} \int_0^1 \left(g(x) + g(1-x)\right) dx = E\left[\frac{1}{2}\left(g(U) + g(1-U)\right)\right].$$

Therefore one can draw n independent random variables $U_1, \ldots, U_n$ following a uniform law on $[0,1]$, and approximate I by

$$I_{2n} = \frac{1}{2n}\left(g\left(U_1\right) + g\left(1-U_1\right) + \cdots + g\left(U_n\right) + g\left(1-U_n\right)\right).$$

We need to compare the efficiency of this Monte Carlo method with the standard one with $2n$ drawings

$$I_{2n}^0 = \frac{1}{2n}\left(g\left(U_1\right) + g\left(U_2\right) + \cdots + g\left(U_{2n-1}\right) + g\left(U_{2n}\right)\right).$$

We will now compare the variances of I_{2n} and I_{2n}^0. Observe that in doing this we assume that most of numerical work relies in the evaluation of f and the time devoted to the simulation of the random variables is negligible. This is often a realistic assumption. An easy computation shows that the variance of the standard estimator is

$$\mathrm{Var}\left(I_{2n}^0\right) = \frac{1}{2n}\mathrm{Var}\left(g\left(U_1\right)\right),$$

whereas

$$\begin{aligned}\mathrm{Var}\left(I_{2n}\right) &= \frac{1}{n}\mathrm{Var}\left(\frac{1}{2}\left(g\left(U_1\right) - g\left(1-U_1\right)\right)\right) \\ &= \frac{1}{4n}\left(\mathrm{Var}\left(g\left(U_1\right) + \mathrm{Var}\left(g\left(1-U_1\right)\right) + 2\mathrm{Cov}\left(g\left(U_1\right), g\left(1-U_1\right)\right)\right)\right) \\ &= \frac{1}{2n}\left(\mathrm{Var}\left(g\left(U_1\right)\right) + \mathrm{Cov}\left(g\left(U_1\right), g\left(1-U_1\right)\right)\right).\end{aligned}$$

Then, $\mathrm{Var}\left(I_{2n}\right) \leq \mathrm{Var}\left(I_{2n}^0\right)$ if and only if $\mathrm{Cov}\left(g\left(U_1\right), g\left(1-U_1\right)\right) \leq 0$. One can prove that if f is a monotone function this is always true and thus the Monte Carlo method using antithetic variates is better than the standard one. This ideas can be generalized in dimension greater than 1, in which case we use the transformation

$$\left(U_1, \ldots, U_d\right) \to \left(1-U_1, \ldots, 1-U_d\right).$$

More generally, if X is a random variable taking its values in $\mathbb{R}^d$ and T is a transformation of $\mathbb{R}^d$ such that the law of $T(X)$ is the same as the law of X, we can construct an antithetic method using the equality

$$E\left[g\left(X\right)\right] = \frac{1}{2}E\left[g\left(X\right) + g\left(T\left(X\right)\right)\right].$$

Namely, if $\left(X_1, \ldots, X_n\right)$ are independent and sampled along the law of X, we can consider the estimator

$$I_{2n} = \frac{1}{2n}\left(g\left(X_1\right) + g\left(T(X_1\right) + \cdots + g\left(X_n\right) + g\left(T\left(X_n\right)\right)\right)$$

and compare it to

$$I_{2n}^0 = \frac{1}{2n}\left(g\left(X_1\right) + g\left(X_2\right) + \cdots + g\left(X_{2n-1}\right) + g\left(X_{2n}\right)\right).$$

The same computations as before prove that the estimator I_{2n} is better than the crude one if and only if $\mathrm{Cov}\left(g\left(X\right), g\left(T\left(X\right)\right)\right) \leq 0$.

1.1.4 Stochastic differential equations

1.1.4.1 Existence and uniqueness, applications in finance

Let $T > 0$ be a fixed time frame (an option maturity for example). Let's consider a probability space $(\Omega, \mathcal{A}, P)$, provided with a filtration $(\mathcal{F}_t)_{t\geq 0}$, and a n-dimensional $\mathcal{F}_t$-brownian motion defined in that space: $W_t = \left(W_t^1, \ldots, W_t^d\right)$. We also consider $Y : \Omega \to \mathbb{R}^n$ a real-valued random variable $\mathcal{F}_0$-measurable, and two functions $\sigma : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n\times d}$, $b : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n$. Our attention focuses on the following SDE (stochastic differential equation)

$$\begin{cases} dX_t &= \sigma(t, X_t)\, dW_t + b(t, X_t)\, dt \\ X_0 &= Y \end{cases} \tag{1.1.2}$$

Definition 1.8 (SDE solution)**.** We call solution of the SDE (1.1.2) a $\mathcal{F}_t$-adapted continuous $\mathbb{R}^n$-valued process $(X_t)_{t\in[0,T]}$ such that

- $\int_0^T |b(s, X_s)| + |\sigma(s, X_s)|^2\, ds < +\infty$ a.s.,
- $\forall t \in [0, T], X_t = Y + \int_0^t \sigma(s, X_s)\, dW_s + \int_0^t b(s, X_s)\, ds.$

The main result of existence and uniqueness for the SDEs is the stochastic version of the Cauchy-Lipschitz theorem for the ODEs. They have similar hypothesis.

Theorem 1.9 (Itô)**.** *We suppose that*

$$(Lip)\quad \exists K > 0, \forall t \in [0, T], \begin{cases} \forall x \in \mathbb{R}^n, & |\sigma(t, x)| + |b(t, x)| \leq K(1 + |x|) \\ \forall x \in \mathbb{R}^n, & |\sigma(t, x) - \sigma(t, y)| + |b(t, x) - b(t, y)| \leq K|x - y|. \end{cases}$$

Then the SDE in (1.1.2) has one and only one solution $(X_t)_{t\in[0,T]}$ (if X'_t is another solution, then $\forall t \in [0, T], X_t = X'_t$ a.s.). Moreover, if $E\left[|Y|^2\right] < +\infty$, then

$$E\left[\sup_{t\leq T} |X_t|^2\right] \leq C\left(1 + E\left[|Y|^2\right]\right)$$

where C is a constant that doesn't depend from Y.

The price of a Call option, having maturity T and strike K, defined over the underlying X, is given by $C = E\left[e^{-rT}(X_T - K)^+\right]$. In the Black-Scholes model, we know a closed formula for such a type of options and we also know a way to perform exact simulations the underlying $X_t = X_0 e^{\sigma W_T + \left(r - \frac{\sigma^2}{2}\right)T}$ to estimate C by the Monte Carlo method. But if the model complexity increases a little bit, adding for example stochastic volatility, these pleasant properties get lost. To calculate the price, then, we can discretize the SDE: this gives a way to simulate a random variable $\bar{X}_T$ approximating X_T. The Monte Carlo method consists then in approximating C by

$$\frac{1}{M}\sum_{i=1}^{M} e^{-rT}\left(\bar{X}_T^i - K\right)^+$$

where the $\bar{X}_T^i$ variables are i.i.d.. The error over the result can be written as a sum of two terms: a discretization error and a statistic error

$$C - \frac{1}{M}\sum_{i=1}^{M} e^{-rT}\left(\bar{X}_T^i - K\right)^+ = E\left[e^{-rT}\left(X_T - K\right)^+\right] - E\left[e^{-rT}\left(\bar{X}_T - K\right)^+\right] + \\ + E\left[e^{-rT}\left(\bar{X}_T - K\right)^+\right] - \frac{1}{M}\sum_{i=1}^{M} e^{-rT}\left(\bar{X}_T^i - K\right)^+ .$$

The behavior of statistical error term is well known and stems from the central limit theorem: because of $E\left[\left(\left(X_T - K\right)^+\right)^2\right] < +\infty$, as M becomes bigger an bigger, this term behaves as

$$e^{-rT}\sqrt{\frac{\text{Var}\,\left(\left(\bar{X}_T - K\right)^+\right)}{M}}G$$

where $G \sim \mathcal{N}_1(0,1)$.

1.1.4.2 The Euler method

Let's consider a partition of the time lag $[0,T]$ in N sub lags, using $N-1$ equispaced points in $(0,T)$: for all $k \in \{0,\dots,N\}$, we define $t_k = \frac{kT}{N}$. Then, all the subsets have the same length equal to T/N. Consider the SDE

$$dX_t = \sigma\left(t, X_t\right)dW_t + b\left(t, X_t\right)dt,\ X_0 = y \in \mathbb{R}^n.$$

The Euler method is based on the following discretization:

$$\begin{cases} \bar{X}_0 & = y \\ \bar{X}_{t_{k+1}} & = \bar{X}_{t_k} + \sigma\left(t_k, \bar{X}_{t_k}\right)\left(W_{t_{k+1}} - W_{t_k}\right) + b\left(t_k, \bar{X}_{t_k}\right)\left(t_{k+1} - t_k\right). \end{cases}$$

We fix the coefficient values at their value at the beginning of the time lapse to move to the next time instant. To implement this scheme, it is enough to know the increments $\left(W_{t_{k+1}} - W_{t_k}\right)_{0\le k\le N-1}$ that are i.i.d. following a Gaussian law $\mathcal{N}_d\left(0, \frac{T}{N}I_d\right)$ where I_d is the d-sized identity matrix. To present the main results about convergence, it is useful to introduce the continue version of the Euler method, defined by

$$\bar{X}_t = \bar{X}_{t_k} + \sigma\left(t_k, \bar{X}_{t_k}\right)\left(W_t - W_{t_k}\right) + b\left(t_k, \bar{X}_{t_k}\right)\left(t - t_k\right),\ \forall 0 \le k \le N-1, \forall t \in \left[t_k, t_{k+1}\right].$$

We present now two theorems showing some interesting results about the convergence of the Euler method.

1.1.4.3 Strong rate

In the scheme, the coefficients b and σ are fixed at their value at the beginning of each time step: we need to introduce some regularity hypothesis for these functions.

Theorem 1.10 (Strong rate). *Let's suppose that the coefficient σ, b verify the hypothesis (Lip) (see theorem 1.9, and let's also suppose that*

$$\exists \alpha, K > 0, \forall x \in \mathbb{R}^n, \forall (s,t) \in [0,T], |\sigma(t,s) - \sigma(s,x)| + |b(t,x) - b(s,x)| \leq K(1+|x|)(t-s)^\alpha .$$

Then for $\beta = \min(\alpha, 1/2)$,

$$\forall p \geq 1, \exists C_p > 0, \forall y \in \mathbb{R}^n, \forall N \in \mathbb{N}, E\left[\sup_{t \leq T} \left|X_t - \bar{X}_t^N\right|^{2p}\right] \leq \frac{C_p\left(1+|y|^{2p}\right)}{N^{2\beta p}}.$$

Moreover if $\gamma < \beta$, $N^\gamma \sup_{t \leq T} \left|X_t - \bar{X}_t^N\right|$ converges almost surely to 0 as N moves to infinity.

As we have

$$\left\|\sup\left|X_t - \bar{X}_t^N\right|\right\|_{2p} = \left(E\left[\sup_{t \leq T}\left|X_t - \bar{X}_t^N\right|^{2p}\right]\right)^{\frac{1}{2p}} \leq \frac{C}{N^\beta}$$

we say that the strong rate of the Euler method is in $\frac{1}{N^\beta}$. Particularly when the coefficients of the SDE don't depend on time or if $\alpha \geq \frac{1}{2}$, the strong rate is in $\frac{1}{2}$.

1.1.4.4 Weak rate

Let f be a Lipschitz function, with Lipschitz constant K, and let's suppose we want to calculate $E[f(X_T)]$. The theorem 1.10 ensures that

$$\left|E\left[f(X_T) - E\left(f\left(\bar{X}_T^N\right)\right)\right]\right| \leq \frac{C}{N^\beta}.$$

The first inequality consists in bounding from above the absolute value of the difference of the expected values by the expected value of the absolute values of the difference: this is very rough. Actually, our aim is to know if $E\left[f\left(\bar{X}_T^N\right)\right]$ is close to $E[f(X_T)]$: this is equivalent to clarify if the law of $\bar{X}_T^N$ is close to the law of X_T. So we focus on the problem of the convergence of the law of the Euler scheme. The convergence in distribution is based on the analysis over some tests functions as the function f introduced before. That's why we speak about weak rate. The following result was first proved by Talay and Tubaro [45].

Theorem 1.11 (Weak rate). *Let b and σ be two functions in C^∞ over $[0,T] \times \mathbb{R}^n$ with bounded derivatives et let $f : \mathbb{R}^n \to \mathbb{R}$ be C^∞ with derivatives having polynomial growth*

$$\forall a = (a_1, \ldots, a_n) \in \mathbb{N}^n, \exists p, C > 0, \forall x \in \mathbb{R}^n, \left|\frac{\partial^{a_1 + \cdots + a_n} f}{\partial x_1^{a_1} \ldots \partial x_n^{a_n}}(x)\right| \leq C(1+|x|^p).$$

Then, there exists a sequence $(C_l)_{l \geq 1}$ of real numbers such that for all $L \in \mathbb{N}^*$, the error can be developed as

$$E\left[f\left(\bar{X}_T^N\right)\right] - E[f(X_T)] = \frac{C_1}{N} + \frac{C_2}{N^2} + \cdots + \frac{C_L}{N^L} + \mathcal{O}\left(\frac{1}{N^{L+1}}\right)$$

In particular this implies that $\left|E\left[f\left(\bar{X}_T^N\right)\right]-E\left[f\left(X_T\right)\right]\right| \leq \frac{C}{N}$ whereas using the strong rate results, we would bound from above

$$\left|E\left[f\left(\bar{X}_T^N\right)\right]-E\left[f\left(X_T\right)\right]\right| \leq \frac{C}{\sqrt{N}}.$$

In fact the hypothesis of Theorem 1.11 leads to have $\alpha = 1$.

1.1.5 Longstaff-Schwartz algorithm for American option pricing

The computation of American option prices is a challenging problem, especially when several underlying assets are involved. The mathematical problem to solve is an optimal stopping problem. In classical diffusion models, this problem is associated with a variational inequality, for which, in higher dimensions, classical PDE methods are ineffective.

Various authors introduced numerical methods based on Monte Carlo techniques. The starting point is to replace the time interval of exercise dates by finite a subsets. This amounts to approximate the American option by a so called Bermuda option. The solution of the corresponding discrete optimal stopping problem reduces to an effective implementation of the dynamic programming principle. The conditional expectations involved in the iterations of dynamic programming cause the main difficulty for the development of Monte Carlo techniques. One way of treating this problem is to use least squares regression on a finite set of functions as a proxy for conditional expectation. This is the method used by Longstaff and Schwartz [33].

1.1.5.1 Description of the algorithm

As mentioned in the introduction, the first step in all probabilistic approximation methods is to replace the original optimal stopping problem in continuous time by an optimal stopping problem in discrete time. Therefore we will present the Longstaff-Schwartz algorithm in the context of discrete optimal stopping.

We will consider a probabilistic space $(\Omega, \mathcal{A}, \mathbb{P})$, equipped with a discrete filtration $\left(\mathcal{F}_j\right)_{j=0,\ldots,L}$. Here the positive integer L denotes the (discrete) time horizon. Given an adapted payoff $\left(Z_j\right)_{j=0,\ldots,L}$ process, where $Z_0, Z_1, \ldots, Z_L$ are square integrable random variables, we are interested in computing

$$\sup_{\tau \in \mathcal{T}_{0,L}} E\left[Z_\tau\right],$$

where $\mathcal{T}_{j,L}$ denotes the set of all stopping times with values in $\{j, \ldots, L\}$.

Following classical optimal stopping theory (for which we refer to [36], chapter 6), we introduced the Snell envelope $\left(U_j\right)_{j=0,\ldots,L}$ of the payoff process $\left(Z_j\right)_{j=0,\ldots,L}$, defined by

$$U_j = \text{ess } \sup_{\tau \in \mathcal{T}_{j,L}} E\left[Z_\tau | \mathcal{F}_j\right], \ j = 0, \ldots, L.$$

The dynamic programming principle can be written as follows:

$$\begin{cases} U_L & = Z_L \\ U_j & = \max\left(Z_j, E\left[U_{j+1} | \mathcal{F}_j\right]\right), \ 0 \leq j \leq L-1. \end{cases}$$

We also have $U_j = E\left[Z_{\tau_j}|\mathcal{F}_j\right]$, with

$$\tau_j = \min\left\{k \geq j | U_k = Z_k\right\}.$$

In particular $E\left[U_0\right] = \sup_{\tau\in\mathcal{T}_{0,L}} E\left[Z_\tau\right] = E\left[Z_{\tau_0}\right]$.

The dynamic programming principle can be written in terms of optimal stopping times τ_j, as follows:

$$\begin{cases} \tau_L &= L \\ \tau_j &= j 1_{\left\{Z_j \geq E\left[Z_{\tau_{j+1}}|\mathcal{F}_j\right]\right\}} + \tau_{j+1} 1_{\left\{Z_j < E\left[Z_{\tau_{j+1}}|\mathcal{F}_j\right]\right\}},\ j \leq L-1. \end{cases}$$

This formulation in terms of stopping rules (rather than in terms of value functions) plays an essential role in the Longstaff-Schwartz method.

The method also requires that the underlying model be a Markov chain. Therefore, we will assume that there is an $(\mathcal{F}_j)$-Markov chain $(X_j)_{j=0,\ldots,L}$ with state space $(E, \mathcal{E})$ such that, for $j = 0, \ldots, L$

$$Z_j = f\left(j, X_j\right),$$

for some Borel function $f\left(j,\cdot\right)$. We then have $U_j = V\left(j, X_j\right)$ for some function $V\left(j,\cdot\right)$ and $E\left[Z_{\tau_{j+1}}|\mathcal{F}_j\right] = E\left[Z_{\tau_{j+1}}|X_j\right]$. We will also assume that the initial state $X_0 = x$ is deterministic, so that U_0 is also deterministic.

The first step of the Longstaff-Schwartz algorithm is to approximate the conditional expectation with respect to X_j by the orthogonal projection on the space generated by a finite number of functions of X_j. Let us consider a sequence $\left(e_k\left(x\right)\right)_{k\geq1}$ of measurable real valued functions defined on E and satisfying the following conditions:

Assumption 1. *(A1)*
For $j = 0$ to $j = L-1$, the sequence $\left(e_k\left(X_j\right)\right)_{k\geq1}$ is total in $L^2\left(\sigma\left(X_j\right)\right)$.

Assumption 2. *(A2)*
For $j = 0$ to $j = L-1$ and $m \geq 1$, if $\sum_{k=1}^m \lambda_k e_k\left(X_j\right) = 0$ a.s. then $\lambda_k = 0$ for $k = 1$ to m.

Then, for $j = 1$ to $L-1$, we denote by P_j^m the orthogonal projection from $L^2\left(\Omega\right)$ onto the vector space generated by $\left\{e_1\left(X_j\right), \ldots, e_m\left(X_j\right)\right\}$ and we introduce the stopping times $\tau_j^{[m]}$:

$$\begin{cases} \tau_j^{[m]} &= L \\ \tau_j^{[m]} &= j 1_{\left\{Z_j \geq P_j^m\left(Z_{\tau_{j+1}^{[m]}}\right)\right\}} + \tau_{j+1}^{[m]} 1_{\left\{Z_j < P_j^m\left(Z_{\tau_{j+1}^{[m]}}\right)\right\}},\ j \leq L-1. \end{cases}$$

From these stopping times, we obtain an approximation of the value function:

$$U_0^m = \max\left(Z_0, E\left[Z_{\tau_1^{[m]}}\right]\right).$$

Recall that $Z_0 = f\left(0, x\right)$ is deterministic. The second step of the algorithm is then to evaluate numerically $E\left[Z_{\tau_1^{[m]}}\right]$ by a Monte Carlo procedure. We assume that we can simulate N independent paths $\left(X_j^{(1)}\right), \left(X_j^{(2)}\right), \ldots, \left(X_j^{(N)}\right)$ of the Markov chain $\left(X_j\right)$ and we denote by $Z_j^{(n)}$

the associated payoff for $j = 0$ to L and $n = 1$ to N $\left(Z_j^{(n)} = f\left(j, X_j^{(n)}\right)\right)$. For each path n, we estimate recursively the stopping times $\left(\tau_j^{[m]}\right)$ by:

$$\begin{cases} \tau_L^{n,m,N} & = L \\ \tau_L^{n,m,N} & = j1_{\left\{Z_j^{(n)} \geq \alpha_j^{(m,N)} \cdot e^m\left(X_j^{(n)}\right)\right\}} + \tau_{j+1}^{n,m,N} 1_{\left\{Z_j^{(n)} < \alpha_j^{(m,N)} \cdot e^m\left(X_j^{(n)}\right)\right\}}, j \leq L-1. \end{cases}$$

Here, $x \cdot y$ denotes the usual inner product in $\mathbb{R}^m$, e^m is the vector valued function $(e_1, \ldots, e_m)$ and $\alpha_j^{(m,N)}$ is the least square estimator:

$$\alpha_j^{(m,N)} = \arg \min_{a \in \mathbb{R}^m} \sum_{n=1}^{N} \left(Z^{(n)}_{\tau_{j+1}^{n,m,N}} - a \cdot e^m\left(X_j^{(n)}\right)\right)^2 .$$

Remark that for $j = 1$ to $L-1$, $\alpha_j^{(m,N)} \in \mathbb{R}^m$. Finally, from the variables $\tau_j^{n,m,N}$, we derive the following approximation for U_0^m:

$$U_0^{m,N} = \max\left(Z_0, \frac{1}{N} \sum_{n=1}^{N} Z^{(n)}_{\tau_1^{n,m,N}}\right).$$

In [15], the authors proved that for any fixed m, $U_0^{m,N}$ converges almost surely to U_0^m as N goes to infinity, and that U_0^m converges to U_0 as m goes to infinity.

1.1.6 Simulation of the Heston Model

In this Section we describe the most relevant method actually used to perform Monte Carlo simulations for the Heston model: the Alfonsi's third order scheme (see [2]).

1.1.6.1 The Model

The Heston model [26] is one of the most known and used models in finance to describe the evolution of the volatility of an underlying asset and the underlying asset itself. In order to fix the notation, we report its dynamics:

$$\begin{cases} dS_t = rS_t dt + \sqrt{v_t} S_t dZ_t^S & S_0 = \bar{S}_0, \\ dv_t = k(\theta - v_t) dt + \omega\sqrt{v_t} dZ_t^v & v_0 = \bar{v}_0, \end{cases} \tag{1.1.3}$$

where Z^S and Z^v are Brownian motions, and $d\left\langle Z_t^S, Z_t^v\right\rangle = \rho dt$.

1.1.6.2 The Alfonsi's third order scheme

Alfonsi's paper ([2]) presents weak second and third order schemes for the Cox-Ingersoll-Ross (CIR) process, without any restriction on its parameters. At the same time, it gives a general recursive construction method to get weak second-order schemes that extends the one introduced by Ninomiya and Victoir [38]. Combining these both results, this allows to propose a second-order scheme for more general affine diffusions. Simulation examples are given to illustrate the convergence of these schemes on CIR and Heston models. Algorithms are stated in a pseudo-code language.

The main difficulty when discretizing the CIR process is located in 0, where the square root is not Lipschitzian. Usual schemes such as the Euler scheme or the Milshtein scheme are in general not well defined. They can indeed lead to negative values for which the square root is not defined. One has therefore to modify them or to create ad-hoc schemes. A possible criteria to chose the scheme may be its capacity to support large values of σ (we mean here $\omega^2 \gg 4k$). In finance, such large values do not occur when the CIR diffusion is used to represent the short interest rate. They are instead often observed when the CIR stands for the default intensity in credit risk or the stock volatility like in the Heston model. Heuristically, the larger is ω, the more the CIR process spends time in the neighborhood of 0 where the square-root is very sensitive. This is intuitively why most of the schemes fail to be accurate for large ω.

Now, we point out the main features of the paper in order to present the scheme.

Assumptions on the SDE and notations We consider a d_W-dimensional standard Brownian motion $(W_t, t \geq 0)$ and we will denote in the sequel $(\mathcal{F}_t)_{t\geq 0}$ its augmented associated filtration that satisfies the usual conditions. Let $d \in \mathbb{N}^*$, and $\mathbb{D} \subset \mathbb{R}^d$ a domain that we assume for sake of simplicity to be a product of d intervals. Typically, we will consider $\mathbb{D} = \mathbb{R}_+^{d_1} \times \mathbb{R}^{d_2}$ with $d_1 + d_2 = d$. For any multi-index $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}^d$, we define $\partial_\alpha = \partial_1^{\alpha_1} \ldots \partial_d^{\alpha_d}$ and $|\alpha| = \sum_{l=1}^d \alpha_l$. We introduce the following functional space:

$$\mathcal{C}^\infty_{\text{pol}}(\mathbb{D}) = \left\{ f \in \mathcal{C}^\infty(\mathbb{D}, \mathbb{R}), \forall \alpha \in \mathbb{N}^d, \exists C_\alpha > 0, e_\alpha \in \mathbb{N}^*, \forall x \in \mathbb{D}, |\partial_\alpha f(x)| \leq C_\alpha \left(1 + \|x\|^{e_\alpha}\right) \right\}$$

where $\|\cdot\|$ is a norm on $\mathbb{R}^d$. We will say that $(C_\alpha, e_\alpha)_{\alpha\in\mathbb{N}^d}$ is a good sequence for $f \in \mathcal{C}^\infty_{\text{pol}}(\mathbb{D})$ if one has $\forall x \in \mathbb{D}, |\partial_\alpha f(x)| \leq C_\alpha (1 + \|x\|^{e_\alpha})$.

We do the following assumptions.

Assumptions 1.12. *We assume that $b : \mathbb{D} \to \mathbb{R}^d$ and $\sigma : \mathbb{D} \to \mathcal{M}_{d\times d_W}(\mathbb{R})$ are such that for $1 \leq i, j \leq d$, the functions $x \in \mathbb{D} \mapsto b_i(x)$ and $x \in \mathbb{D} \mapsto (\sigma\sigma^*)_{i,j}(x)$ are in $\mathcal{C}^\infty_{pol}(\mathbb{D})$. For $x \in \mathbb{D}$, we introduce the general $\mathbb{R}^d$-valued SDE:*

$$t \geq 0, X_t^x = x + \int_0^t b(X_s^x)\, ds + \int_0^t \sigma(X_s^x)\, dW_s. \tag{1.1.4}$$

We assume that for any $x \in \mathbb{D}$, there is a unique weak solution defined for $t \geq 0$, and therefore

$$P(\forall t \geq 0, X_t^x \in \mathbb{D}) = 1.$$

The differential operator associated to the SDE is given by

$$f \in \mathcal{C}^2(\mathbb{D}, \mathbb{R}), Lf(x) = \sum_{i=1}^d b_i(x)\, \partial_i f(x) + \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \sum_{k=1}^{d_W} \sigma_{i,k}(x)\, \sigma_{j,k}(x)\, \partial_i \partial_j f(x). \tag{1.1.5}$$

If $f \in \mathcal{C}^{\infty}_{pol}(\mathbb{D})$, thanks to the regularity assumptions made on b and σ, all the iterated functions $L^k f(x)$ are well defined on $\mathbb{D}$ and belong to $\mathcal{C}^{\infty}_{pol}(\mathbb{D})$ for any $k \in \mathbb{N}$.

Definition 1.13. We will say (for short) that the operator L satisfies the required assumptions on $\mathbb{D}$ if it is defined by (1.1.5) for some functions $b(x)$ and $\sigma(x)$ and satisfies all assumptions above.

Now, let us turn to discretization schemes for the SDE (1.1.4). Let us fix a time horizon $T > 0$. We will consider in the whole Section the time interval $[0, T]$ and the regular time discretization $t^n_i = iT/n$ for $i = 0, \dots n$.

Definition 1.14. A family of transition probabilities $(\hat{p}_x(t)(dz), t > 0, x \in \mathbb{D})$ on $\mathbb{D}$ is such that $\hat{p}_x(t)$ is a probability law on $\mathbb{D}$ for $t > 0$ and $x \in \mathbb{D}$.

A discretizations scheme with transition probabilities $(\hat{p}_x(t)(dz), t > 0, x \in \mathbb{D})$ is a sequence $\left(\hat{X}^n_{t^n_i}, 0 \le i \le n\right)$ of $\mathbb{D}$-valued random variables such that:

- for $0 \le i \le n$, $\hat{X}^n_{t^n_i}$ is a $\mathcal{F}_{t^n_i}$-measurable random variable on $\mathbb{D}$.
- the law of $\hat{X}^n_{t^n_{i+1}}$ is given by $\mathbb{E}\left[f\left(\hat{X}^n_{t^n_{i+1}}\right)|\mathcal{F}_{t^n_i}\right] = \int_{\mathbb{D}} f(z)\,\hat{p}_{\hat{X}^n_{t^n_i}}(T/n)(dz)$ and thus only depends on $\hat{X}^n_{t^n_i}$ and T/n .

For convenience, we will denote, for $t > 0$ and $x \in \mathbb{D}$, $\hat{X}^x_t$ a random variable distributed according to the probability law $\hat{p}_x(t)(dz)$. The law of discretization scheme $(\hat{X}^n_{t^n_i},\ 0 \le i \le n)$ is thus entirely determined by its initial value and its transition probabilities.

Definition 1.15 (Weak ν th-order scheme)**.** Let us denote $\mathcal{C}^{\infty}_K(\mathbb{D}, \mathbb{R})$ the set of $\mathcal{C}^{\infty}$ real valued functions with a compact support in $\mathbb{D}$. Let $x \in \mathbb{D}$. A discretization scheme $\left(\hat{X}^n_{t^n_i}, 0 \le i \le n\right)$ is a weak ν th-order scheme for the SDE $\left(X^X_t, t \in [0, T]\right)$ if :

$$\forall f \in \mathcal{C}^{\infty}_K(\mathbb{D}, \mathbb{R}), \exists K > 0, \left|E\left[f\left(X^x_T\right)\right] - E\left[f\left(\hat{X}^n_T\right)\right]\right| \le K/n^{\nu}.$$

The quantity $E\left[f\left(X^x_T\right)\right] - E\left[f\left(\hat{X}^n_T\right)\right]$ is called the weak error associated to f.

The third order scheme We present now the Alfonsi's scheme. A proof of its properties is available in [2].

First we write the CIR process as

$$\begin{cases} dX^x_t &= (a - kX^x_t)\,dt + \sigma\sqrt{X^x_t}\,dW_t \\ X^x_0 &= x \end{cases}$$

with parameters $(a, k, \sigma) \in \mathbb{R}^*_+ \times \mathbb{R} \times \mathbb{R}_+$. We remember that it is a non-negative process. Moreover, if $x > 0$ and $2a \ge \sigma^2$ the process $(X_t, t \ge 0)$ is always positive. We will exclude the trivial case $\sigma = 0$ and assume $\sigma > 0$. This process has $d_W = 1$ and $\mathbb{D} = \mathbb{R}_+$. We introduce its operator

$$f \in \mathcal{C}^2(\mathbb{R}_+, \mathbb{R}), \; L^{CIR} f(x) = (a - kx)\,\partial_x f(x) + \frac{1}{2}\sigma^2 x \partial_x^2 f(x)$$

that satisfies the required assumptions on $\mathbb{D}$.

We define the following quantities:

$$\psi_k(t) = \frac{1 - e^{-kt}}{k}, \; \psi_0(t) = t$$

$$\begin{aligned} K_3(t) = \psi_{-k}(t)\, 1_{\{4a/3<\sigma^2<4a\}} \left(\sqrt{\frac{\sigma^2}{4} - a + \frac{\sigma}{\sqrt{2}}\sqrt{a - \frac{\sigma^2}{4}}} + \frac{\sigma}{2}\sqrt{3+\sqrt{6}} \right)^2 + \\ + \psi_{-k}(t)\, 1_{\{4a<\sigma^2\}} \left[\frac{\sigma^2}{4} - a \left(\sqrt{\frac{\sigma}{\sqrt{2}}\sqrt{\frac{\sigma^2}{4} - a}} + \frac{\sigma}{2}\sqrt{3+\sqrt{6}} \right) \right] + \\ + \psi_{-k}(t)\, 1_{\{\sigma^2 \le 4a/3\}} \frac{\sigma}{\sqrt{2}} \sqrt{a - \sigma^2/4} \end{aligned} \quad (1.1.6)$$

Let Y be a random discrete variable such that $P\left[Y = \sqrt{3+\sqrt{6}}\right] = P\left[Y = -\sqrt{3+\sqrt{6}}\right] = \frac{\sqrt{6}-2}{4\sqrt{6}}$, and $P\left[Y = \sqrt{3-\sqrt{6}}\right] = P\left[Y = -\sqrt{3-\sqrt{6}}\right] = \frac{1}{2} - \frac{\sqrt{6}-2}{4\sqrt{6}}$. This variable is useful because it fits the first seven first moments of a standard Gaussian variable.

We consider the three following discretization schemes

$$\begin{aligned} X_0^{CIR}(t,x) &= xe^{-kt} + \left(a - \sigma^2/4\right)\psi_k(t), \\ X_1^{CIR}(t,x) &= \left(\left(\sqrt{x} + \frac{\sigma}{2} t \right)^+ \right)^2 \\ \tilde{X}(t,x) &= x + t\frac{\sigma}{\sqrt{2}}\sqrt{\left| a - \frac{\sigma^2}{4} \right|} \end{aligned}$$

The first two operator are obtained applying Ninomiya-Victoir's theorem (see [38]) and their composition can define a second order scheme:

$$\begin{aligned} &X_0^{CIR}\left(t/2, X_1^{CIR}\left(N, X_0^{CIR}(t/2, x)\right)\right) = \\ &= e^{-\frac{kt}{2}} \left(\sqrt{\left(a - \frac{\sigma^2}{4}\right)\psi\left(\frac{t}{2}\right) + xe^{-\frac{kt}{2}}} + \frac{w\sigma}{2} \right)^2 + \left(a - \frac{\sigma^2}{4}\right)\psi\left(\frac{t}{2}\right), \; N \sim \mathcal{N}(0,1) \end{aligned}$$

We can now state the scheme for those starting point x that are far enough from zero.

Proposition 1.16. *Let ε and ζ be respectively uniform r.v. on $\{-1,1\}$ and $\{1,2,3\}$, and Y be sampled independently according to the previous definition. Then, for $\sigma^2 \le 4a$ (resp. $\sigma^2 > 4a$), the following scheme*

$$\hat{X}_t^{x,k=0} = \begin{cases} \tilde{X}\left(\varepsilon t, X_0^{CIR}\left(t, X_1^{CIR}\left(\sqrt{t}Y, x\right)\right)\right) & \left(resp.\ \tilde{X}\left(\varepsilon t, X_1^{CIR}\left(\sqrt{t}Y, X_0^{CIR}\left(t,x\right)\right)\right)\right) \text{ if } \zeta = 1, \\ X_0^{CIR}\left(t, \tilde{X}\left(\varepsilon t, X_1^{CIR}\left(\sqrt{t}Y, x\right)\right)\right) & \left(resp.\ X_1^{CIR}\left(\sqrt{t}Y, \tilde{X}\left(\varepsilon t, X_0^{CIR}\left(t,x\right)\right)\right)\right) \text{ if } \zeta = 2, \\ X_0^{CIR}\left(t, X_1^{CIR}\left(\sqrt{t}Y, \tilde{X}\left(\varepsilon t, x\right)\right)\right) & \left(resp.\ X_1^{CIR}\left(\sqrt{t}Y, X_0^{CIR}\left(t, \tilde{X}\left(\varepsilon t, x\right)\right)\right)\right) \text{ if } \zeta = 3, \end{cases}$$

is well defined and non-negative for $t \ge 0$ and $x \ge K_3(t)\, t/\psi_{-k}(t)$. Then, for $x \ge K_3(t)$, the scheme

$$\hat{X}_t^x = e^{-kt} \hat{X}_{\psi_{-k}(t)}^{x,k=0}$$

is a potential third-order scheme.

On $x \in [0, K_3(t)]$ we will approximate the CIR with a discrete random variable that matches the three first moments of the CIR. we approximate the CIR near 0 and keep non-negativity. We remember the moments of the CIR

$$\begin{aligned} \mu_{1,t}^x &= E\left[(X_t^x)^1\right] = xe^{-kt} + a\psi_k(t) \\ \mu_{2,t}^x &= E\left[(X_t^x)^2\right] = \left(\mu_{1,t}^x\right)^2 + \sigma^2\psi_k(t)\left[a\psi_k(t)/2 + xe^{-kt}\right] \\ \mu_{3,t}^x &= E\left[(X_t^x)^3\right] = \mu_{1,t}^x\mu_{2,t}^x + \sigma^2\psi_k(t)\left[2x^2e^{-2kt} + \psi_k(t)\left(a + \frac{\sigma^2}{2}\right)\left(3xe^{-kt} + a\psi_k(t)\right)\right] \end{aligned}$$

We define

$$s = \frac{\mu_{3,t}^x - \mu_{1,t}^x\mu_{2,t}^x}{\mu_{2,t}^x - \left(\mu_{1,t}^x\right)^2},\ p = \frac{\mu_{1,t}^x\mu_{3,t}^x - \left(\mu_{2,t}^x\right)^2}{\mu_{2,t}^x - \left(\mu_{1,t}^x\right)^2}$$

$$x_{\pm}(t,x) = \frac{s \pm \sqrt{s^2 - 4p}}{2},\ \pi(t,x) = \frac{\mu_{1,t}^x - x_-(t,x)}{x_+(t,x) - x_-(t,x)}$$

Proposition 1.17. *Let $U \sim \mathcal{U}([0,1])$ and consider the scheme*

$$\hat{X}_t^x = x_+(t,x)\, 1_{\{U \le \pi(t,x)\}} + x_-(t,x)\, 1_{\{U > \pi(t,x)\}}.$$

This scheme is a positive potential third order scheme on $x \in [0, K_3(t)]$.

We can then conclude with the following

Theorem 1.18 (Alfonsi's third order scheme)**.** *Let $K_3(t)$ be defined as in (1.1.6), $\hat{X}_t^x$ the scheme defined in Proposition 1.16 (resp. Proposition 1.17) for $x \ge K_3(t)$ (resp. $x < K_3(t)$) and $\hat{p}_x(t)(dz)$ the law of $\hat{X}_t^x$. Then, $\hat{p}_x(t)(dz)$ is a potential third order scheme for L^{CIR} on $\mathbb{R}_+$. Moreover, the scheme $\left(\hat{X}_{t_0^n}^n, 0 \le i \le n\right)$ associated to the transition probabilities $(\hat{p}_x(t)(dz), t > 0)$ and starting from $\hat{X}_{t_0^n}^n = x \in \mathbb{R}_+$ is a third order scheme:*

$$\forall f \in \mathcal{C}_{pol}^{\infty}(\mathbb{R}_+), \exists K > 0, \forall n \in \mathbb{N}^*, \left|E\left[f\left(X_T^x\right)\right] - E\left[f\left(\hat{X}_T^n\right)\right]\right| \le K/n^3$$

1.1.6.3 An efficient scheme for the Heston model

In this part, we are going to apply the ideas developed about CIR process to the Heston model [26]. This approach has already been used by Ninomiya and Victoir [38], but the difference here is that we have at our disposal a third-order scheme for the CIR, without restriction on its parameters. Thus, we will use a different splitting of the Heston SDE that allows to use directly our CIR discretization.

First, we rewrite the Heston model using Alfonsi's notation. Let W and Z be two independent Brownian motions. We would like to discretize the following SDE:

$$\begin{cases} X_t^1 &= X_0^1 + \int_0^t \left(a - kX_s^1\right) ds + \sigma \int_0^t \sqrt{X_s^1} dW_s \\ X_t^2 &= \int_0^t X_s^1 ds \\ X_t^3 &= X_0^3 + \int_0^t rX_s^3 ds + \int_0^t \sqrt{X_s^1} X_s^3 \left(\rho dW_s + \sqrt{1-\rho^2} dZ_s\right) \\ X_t^4 &= \int_0^t X_s^3 ds \end{cases}$$

with $X_0^1 \geq 0$, $X_0^3 > 0$, $r \in \mathbb{R}$, $\rho \in [-1,1]$ and $(a,k,\sigma) \in \mathbb{R}_+^* \times \mathbb{R} \times \mathbb{R}_+^*$. The processes X^1 and X^3 are respectively the volatility process and the stock process, and X^2 and X^4 their respective integrals. From a financial point of view, it is common to assume moreover $r > 0$, $k > 0$, and $\rho \leq 0$, but these assumptions are not required for what follows.

First, we have to say that there is no hope that the theory developed in the previous Section works for the Heston model. Indeed, all that theory is thought to work when the discretization scheme has uniformly bounded moments. Since the discretization scheme is supposed to stick rather closely to the SDE, this roughly amounts to assume that the SDE has uniformly bounded moments, which holds when the drift $b(x)$ and the volatility function $\sigma(x)$ have a sub-linear growth. In the Heston model the diffusion coefficient $\sigma(x)$ has not a sub-linear growth, and it is proved indeed that the moments explode in a finite time. Therefore, the framework developed in Alfonsi's paper is not well suited to get a rigorous estimate of the weak error within the Heston model. However, it is not meaningless to apply the results stated in the previous Section to the Heston model. The recursive construction of third-order scheme is a way to cancel many biased terms of order 1, and improve really the convergence.

We will then apply the results of the previous Section in a non rigorous manner. To do so, we split the operator of the SDE $L = L^W + L^Z$, where the two operators are associated to the following respective SDEs:

$$\begin{cases} dX_t^1 &= \left(a - kX_t^1\right) dst + \sigma \sqrt{X_t^1} dW_t \\ dX_t^2 &= X_t^1 dt \\ dX_t^3 &= rX_t^3 dt + \sqrt{X_t^1} X_t^3 \rho dW_t \\ dX_t^4 &= X_t^3 ds \end{cases} \quad \text{and} \quad \begin{cases} dX_t^1 &= 0 \\ dX_t^2 &= 0 \\ dX_t^3 &= \sqrt{\left(1-\rho^2\right) X_t^1} X_t^3 \star dZ_{st} \\ dX_t^4 &= 0. \end{cases}$$

Here $\star$ denotes the Stratonovich integral. The second SDE is easy to integrate exactly. Concerning the first SDE, we use the third order scheme described in this paper for CIR. To discretize X_t^2, we use the trapezoidal rule. Then, we observe that X^3 can be integrated exactly in function of the increments of X^1 and X^2:

$$X_t^3 = X_0^3 \exp\left[\left(r - \frac{\rho}{\sigma} a\right) t + \left[\frac{\rho}{\sigma} k - \frac{1}{2}\right] \left(X_t^2 - X_0^2\right) + \frac{\rho}{\sigma} \left(X_t^1 - X_0^1\right)\right],$$

```
function X0(x):
```
$x \leftarrow x + (a - \sigma^2/4)\,\psi_{-k}(t)$
```
function X1(x):
```
$x \leftarrow \left(\left(\sqrt{x} + \sigma\sqrt{\psi_{-k}(t)}Y/2\right)^+\right)^2$
```
function Xt(x):
```
$x \leftarrow x + \frac{\sigma}{\sqrt{2}}\sqrt{|a - \sigma^2/4|}\varepsilon\psi_{-k}(t)$
```
function CIR3(x):
```
$\mathtt{if}(x \geq K_3(t))\{$

$if\,(\zeta = 1)\left\{if\left(\sigma^2 \leq 4a\right)\{X1(x)\,X0(x)\,Xt(x)\}\,else\,\{X0(x)\,X1(x)\,Xt(x)\}\right\}$

$if\,(\zeta = 2)\left\{if\left(\sigma^2 \leq 4a\right)\{X1(x)\,Xt(x)\,X0(x)\}\,else\,\{X0(x)\,Xt(x)\,X1(x)\}\right\}$

$if\,(\zeta = 3)\left\{if\left(\sigma^2 \leq 4a\right)\{Xt(x)\,X1(x)\,X0(x)\}\,else\,\{Xt(x)\,X0(x)\,X1(x)\}\right\}$

$x \leftarrow xe^{-kt}\}$ `else` $\{$

$s \leftarrow \frac{\mu_3 - \mu_1\mu_2}{\mu_2 - \mu_1^2},\ p \leftarrow \frac{\mu_1\mu_3 - \mu_2^2}{\mu_2 - \mu_1^2},\ \delta = \sqrt{s^2 - 4p},\ \pi \leftarrow \frac{\mu_1 - (s-\delta)/2}{\delta}$

$\mathtt{if}(U < \pi)\{x \leftarrow (s + \delta)/2\}$

$\mathtt{else}\ \{x \leftarrow (s - \delta)/2\}$

$\}$

Table 1.1: Algorithm computing the 3^{rd} order scheme value at the next time-step, starting from x with a time-step. Here, U is sampled uniformly on $[0, 1]$ and ε, ζ and Y as stated in Section 1.1.6.2.

and we use this formula with the increments of the discretization. Last, we discretize X^4 like X^2 using the trapezoidal scheme. Instead of writing the cumbersome formula or Alfonsi's scheme, we prefer to write here directly the algorithm that computes the discretization at the next time-step. The function `HW` (resp. `HZ`) calculates the discretization of the SDE associated to L^W (resp. L^Z). See Table 1.1 and 1.2.

```
function HW(x_1, x_2, x_3, x_4):
Δx_1 ← −x_1, CIR3(x_1), Δx_1 ← Δx_1 + x_1
x_2 ← x_2 + (x_1 + 0.5Δx_1) t
x_4 ← x_4 + 0.5 x_3 t
x_3 ← x_3 exp[(r − ρa/σ) t + ρΔx_1/σ + (ρk/σ − 0.5)(x_1 + 0.5Δx_1) t]
x_4 ← x_4 + 0.5 x_3 t
x_1 ← x_1 + Δx_1
function HZ(x_1, x_2, x_3, x_4):
x_3 ← x_3 exp( sqrt((1 − ρ^2) x_1 t) N )
function Heston(x_1, x_2, x_3, x_4):
if (B = 1) {HZ(x_1, x_2, x_3, x_4), HW(x_1, x_2, x_3, x_4)}
else {HW(x_1, x_2, x_3, x_4), HZ(x_1, x_2, x_3, x_4)}
```

Table 1.2: Algorithm for the Heston model, B being a Bernoulli sample or parameter $1/2$ and N an independent standard Gaussian variable.

1.1.7 Simulation of the Black-Scholes Hull-White Model

The Hull-White model [28] is one of historically most important interest rate models, which is nowadays often used for risk-management purposes. The important advantage of the HW model is the existence of closed formulas to calculate the prices of bonds, caplets and swaptions.

1.1.7.1 The model

In order to fix the notation, we report the dynamics of the BS HW model:

$$\begin{cases} dS_t = r_t S_t dt + \sigma S_t dZ_t^S & S_0 = \bar{S}_0, \\ dr_t = k\left(\theta_t - r_t\right) dt + \omega dZ_t^r & r_0 = \bar{r}_0, \end{cases}$$

where Z^S and Z^r are Brownian motions, and $d\left\langle Z_t^S, Z_t^r \right\rangle = \rho dt$.

The process r is a generalized Ornstein-Uhlenbeck (hereafter OU) process: here θ_t is not constant but it is a deterministic function which is completely determined by the market values of the zero-coupon bonds (*ZCBs*) by calibration (see Brigo and Mercurio [12]): in this case the theoretical prices of the ZCBs match exactly the market prices.

Let $P^M(0,T)$ denote the market price of the ZCB at time 0 for the maturity T. The market instantaneous forward interest rate is then defined by

$$f^M(0,T) = -\frac{\partial \ln P^M(0,T)}{\partial T}.$$

It is well known that the short rate process r can be written as

$$r_t = \omega X_t + \beta(t),$$

where X is a stochastic process given by

$$dX_t = -k X_t dt + dZ_t^r, \ X_0 = 0,$$

and $\beta(t)$ is a function

$$\beta(t) = f^M(0,t) + \frac{\omega^2}{2k^2}(1-\exp(-kt))^2.$$

Then, the BS HW model is described by

$$\begin{cases} dS_t = r_t S_t dt + \sigma S_t dZ_t^S & S_0 = \bar{S}_0, \\ dX_t = -kX_t dt + dZ_t^r & X_0 = 0, \\ r_t = \omega X_t + \beta(t). \end{cases} \tag{1.1.7}$$

A particular case is called *flat curve.* In this case, we assume $P^M(t,T) = e^{-\bar{r}_0(T-t)}$ and $f^M(0,T) = \bar{r}_0$. Then

$$\beta(t) = \bar{r}_0 + \frac{\omega^2}{2k^2}(1-\exp(-kt))^2,$$

and

$$\theta_t = \bar{r}_0 + \frac{\omega^2}{2k^2}(1-\exp(-2kt)).$$

1.1.7.2 The exact scheme

For this model it is easy to define an exact scheme. We refer to [39], where an algorithm for the rates generation is explained. First we recall the following theorem:

Theorem 1.19. *Let $0 \le s < t$. The variables X_t and $\int_s^t X_u du$ are bivariate normal distributed conditionally on $\mathcal{F}_s$ with*

$$\begin{aligned} E[X_t|\mathcal{F}_s] &= X_s e^{-k(t-s)} \\ E\left[\int_s^t X_u du|\mathcal{F}_s\right] &= \frac{1}{k}X_s\left(1-e^{-k(t-s)}\right) \\ Var[X_t|\mathcal{F}_s] &= \frac{1}{2k}\left(1-e^{-2k(t-s)}\right) \\ Var\left[\int_s^t X_u du|\mathcal{F}_s\right] &= \frac{1}{k^2}\left(t-s+\frac{2}{k}e^{-k(t-s)}-\frac{1}{2k}e^{-2k(t-s)}-\frac{3}{2k}\right) \\ Cov\left[X_t, \int_s^t X_u du|\mathcal{F}_s\right] &= \frac{1}{2k^2}\left(1-e^{-k(t-s)}\right)^2. \end{aligned}$$

Then, we observe that

$$\begin{aligned} \mathrm{Var}[X_t|\mathcal{F}_s] &= \mathrm{Var}\left[\int_s^t -kX_u du + \int_s^t dZ_u^r|\mathcal{F}_s\right] \\ &= \mathrm{Var}\left[\int_s^t -kX_u du|\mathcal{F}_s\right] + \mathrm{Var}\left[\int_s^t dZ_u^r|\mathcal{F}_s\right] + 2\mathrm{Cov}\left[\int_s^t -kX_u du; \int_s^t dZ_u^r|\mathcal{F}_s\right] \\ &= k^2\mathrm{Var}\left[\int_s^t X_u du|\mathcal{F}_s\right] + (t-s) - 2k\mathrm{Cov}\left[\int_s^t X_u du; \int_s^t dZ_u^r|\mathcal{F}_s\right], \end{aligned}$$

and then

$$\begin{aligned}\mathrm{Cov}\left[\int_s^t X_u du, Z_t^r - Z_t^r|\mathcal{F}_s\right] &= \frac{k^2\mathrm{Var}\left[\int_s^t X_u du|\mathcal{F}_s\right] + (t-s) - \mathrm{Var}\left[X_t|\mathcal{F}_s\right]}{2k} \\ &= \frac{e^{-k(t-s)} + k(t-s) - 1}{k^2}.\end{aligned}$$

Moreover

$$\begin{aligned}\mathrm{Cov}\left[X_t - X_s, Z_t^r - Z_s^r|\mathcal{F}_s\right] &= \mathrm{Cov}\left[\int_s^t dX_u, Z_t^r - Z_t^r\right] \\ &= \mathrm{Cov}\left[\int_s^t -kX_u du, Z_t^r - Z_t^r\right] + \mathrm{Cov}\left[\int_s^t dZ_u^r, Z_t^r - Z_t^r\right] \\ &= -k\mathrm{Cov}\left[\int_s^t X_u du, Z_t^r - Z_t^r\right] + \mathrm{Var}\left[Z_t^r - Z_t^r\right] \\ &= -\frac{e^{-k(t-s)} + k(t-s) - 1}{k} + (t-s) \\ &= \frac{1 - e^{-k(t-s)}}{k}.\end{aligned}$$

Then, the variance-covariance matrix associated to the Gaussian vector

$$\left(Z_t^r - Z_t^r, X_t - X_s, \int_s^t X_u du\right)$$

conditionally on $\mathcal{F}_s$ is

$$\begin{pmatrix} t-s & \frac{1}{k}\left(1-e^{-k(t-s)}\right) & \frac{1}{k^2}\left(e^{-k(t-s)} + k(t-s) - 1\right) \\ \frac{1}{k}\left(1-e^{-k(t-s)}\right) & \frac{1}{2k}\left(1-e^{-2k(t-s)}\right) & \frac{1}{2k^2}\left(1-e^{-k(t-s)}\right)^2 \\ \frac{1}{k^2}\left(e^{-k(t-s)} + k(t-s) - 1\right) & \frac{1}{2k^2}\left(1-e^{-k(t-s)}\right)^2 & \frac{1}{k^2}\left(t-s+\frac{2}{k}e^{-k(t-s)} - \frac{1}{2k}e^{-2k(t-s)} - \frac{3}{2k}\right) \end{pmatrix}$$

and it is possible to prove that its Cholesky decomposition is:

$$C = \begin{pmatrix} \sqrt{t-s} & -\frac{-1+e^{k(s-t)}}{k\sqrt{t-s}} & \frac{k(t-s)+e^{k(s-t)}-1}{k^2\sqrt{t-s}} \\ 0 & \frac{e^{k(s-t)}\omega_k(t,s)}{\sqrt{2k}\sqrt{t-s}} & -\frac{e^{k(s-t)}\omega_k(t,s)}{\sqrt{2k^2}\sqrt{t-s}} \\ 0 & 0 & 0 \end{pmatrix}$$

where $\omega_k(t,s) = \sqrt{k(s-t) + 4e^{k(t-s)} + e^{2k(t-s)}(k(t-s)-2) - 2}$.

In Table 1.3 we report a pseudo-code for the generation of the Black-Scholes Hull-White process. Similarly to the Heston case, we use the following variables: X_t^1 represents S_t, X_t^2 represents X_t and X_t^3 represents $\int_s^t X_u du$.

`function` $BSHW\,(x_1, x_2, x_3)$:
$x_3 \leftarrow x_2 \cdot \left(1 - e^{-k(t-s)}\right)/k + C_{1,3}G_0 + C_{2,3}G_1$
$x_2 \leftarrow x_2 \cdot e^{-k(t-s)} + C_{1,2}G_0 + C_{2,2}G_1$,
$x_1 \leftarrow x_1 \exp\left[x_3 + \int_s^t \beta\,(u)\,du - \sigma^2(t-s)/2 + \sigma\sqrt{t-s}\left(\rho G_0 + \sqrt{1-\rho^2}G_3\right)\right]$

Table 1.3: Algorithm for the Black-Scholes Hull-White model, G_1, G_2, G_3 being 3 independent standard Gaussian variables.

1.2 Lattice methods

This part is inspired by [9].

Lattice methods are built to implement discrete models such as the Cox-Ross-Rubinstein market model (CRR model, see [17]). The CRR model is an example of a multi-period market model of the stock price. This model known as the binomial model which has as a limiting case the Black-Scholes formula. The binomial model assumes that the stock price at each time moment can go either up or down by the multiplication of two factors called u and d.

1.2.1 The uniperiodal model

Consider a single time step of length Δt. We know the asset price S_0 at the beginning of the time step; the price S_1 at the end of the periods is a random variable. The simplest model we may think of specifies only two possible values, accounting, for example for the possibility of an increase and a decrease in the stock price. To be specific, let us consider Figure 1.2.1. We start with a price S_0; at next time instant we assume that the price may take either value $S_0 u$ or $S_0 d$, where $d < u$, with probabilities p_u and $p_d = 1 - p_u$ respectively. This is a discrete-time model as well, but it is also discrete-state. Now, imagine an option whose unknown value now is denoted by V_0. If the option can only be exercised after Δt, it is easy to find its values f_u and f_u corresponding to the outcomes. They are simply the option payoffs, which are determined by the type of the contract. Exploiting the no-arbitrage principle, we can calculate the price of the option. Let us set up a portfolio consisting of two assets: a riskless bond with initial price $B_0 = 1$ and future price $B_1 = e^{r\Delta t}$, and the underlying asset with initial value S_0. We denote the number of stock shares in the portfolio by δ and the number of bonds by Ψ. The initial value of this portfolio is

$$\Pi_0 = \delta S_0 + \Psi,$$

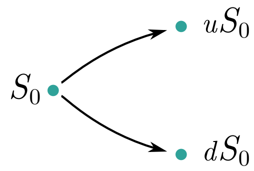

Figure 1.2.1: Simple single-period binomial lattice.

and its future value, depending on the realized state, will be either

$$\Pi_u = \delta S_0 u + \Psi e^{r\Delta t}, \text{ or } \Pi_d = \delta S_0 d + \Psi e^{r\Delta t}.$$

Now let us try to find a portfolio which will exactly replicate the *option* payoff,

$$\begin{cases} \Pi_u & = V_u \\ \Pi_d & = V_d. \end{cases}$$

Solving this system of two linear equations in two unknown variables, we get

$$\delta = \frac{V_u - V_d}{S_0 (u - d)}$$
$$\Psi = e^{-r\Delta t} \frac{uV_u - dV_d}{u - d}.$$

But in order to avoid arbitrage, the initial value of this portfolio must be exactly V_0:

$$\begin{aligned} V_0 &= \delta S_0 + \Psi \\ &= \frac{V_u - V_d}{u - d} + e^{-r\Delta t} \frac{uV_u - dV_d}{u - d} \\ &= e^{-r\Delta t} \left\{ \frac{e^{r\Delta t} - d}{u - d} V_u + \frac{u - e^{r\Delta t}}{u - d} V_d \right\}. \end{aligned} \tag{1.2.1}$$

It is important to note that this relationship does not depend on the objective probabilities p_u and p_d. In particular, the option price is not the discounted expected value of the payoff, which could have been a seemingly reasonable guess; nevertheless, we can interpret equation (1.2.1) as an expected value. Indeed, if we set

$$\pi_u = \frac{e^{r\Delta t} - d}{u - d}, \ \pi_d = \frac{u - e^{r - \Delta t}}{u - d}$$

we may notice that these probabilities define a risk neutral probability $\mathbb{Q}$. The option price can be interpreted as the discount expected value of the payoff under those probabilities:

$$V_0 = e^{-r\Delta t} E_{\mathbb{Q}} [V_1] = e^{-r\Delta t} (\pi_u V_u + \pi_d V_d), \tag{1.2.2}$$

where the notation $E_{\mathbb{Q}}$ is used to point out that expectation is taken with respect to the probability measure $\mathbb{Q}$.

The previous model is also known as uniperiodal model because we consider only one time lag, neglecting the middle values.

To allow for a better model of uncertainty, we should increase the number of states; to replicate the option payoff, we can either use more assets or allow for trading at intermediate dates. The second possibility is more practical and it is essential, for example, to price American option, which allow for early exercise at any time during option life. In the limit this leads to a continuous time model and to the Black-Scholes framework. When the Black Scholes framework

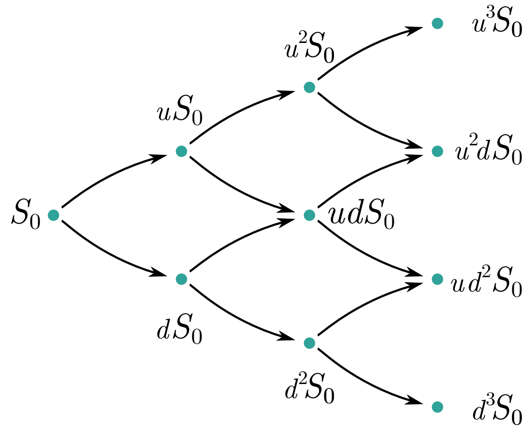

Figure 1.2.2: Recombining binomial lattice.

does not lead to an analytical solution, we must resort to some discretization approach, which can be sampling by Monte Carlo simulation, or setting up a grid and apply finite differences methods to solve the corresponding PDE. A multistage binomial lattice, like the one shown in Figure 1.2.2 is an alternative discretization approach; we could also consider non-recombinant trees, but recombination keeps computational effort to a manageable level.

A good way to simplify calculation is to adopt the convenient choice $u = 1/d$. This is not necessary, but in this way, an up step followed by a down step yields the same initial price:

$$S_0 ud = S_0 du = S_0.$$

As we may see form the figure, not only we have recombination, but the lattice uses a limited number of prices too. The selection of sensible values for u and d can be done with the aim of approximating the underlying continuous-time process.

1.2.2 The multiperiodal model

The binomial lattice should be a good approximation of the risk-neutral process

$$dS = rSdt + \sigma SdW.$$

Hence the parameters we need to set up the lattice should preserve some essential properties of the continuous-time model. This process is called *calibration*. Starting from S_t, after a small time interval Δt, the new random variable $S_{t+\Delta t}$ is such that

$$\ln\left(S_{t+\Delta t}\right) \sim N\left(\left(r - \sigma^2/2\right)\Delta t, \sigma^2 \Delta t\right).$$

Using properties of the log normal distribution we have

$$E\left[S_{t+\Delta t}/S_t\right] = e^{r\Delta t} \tag{1.2.3}$$

and

$$\operatorname{Var}\left[S_{t+\Delta t}/S_t\right] = e^{2r\Delta t}\left(e^{\sigma^2\Delta t} - 1\right). \tag{1.2.4}$$

A reasonable requirement on the discretized dynamics is that it should match these moments. Note that these are two conditions, but we have three parameters: p, u and d. So we have one degree of freedom, and we may choose $u = 1/d$. This is a convenient choice from a computational point of view, but is not the only possibility.

On the lattice, we have

$$E\left[S_{t+\Delta t}\right] = pu \cdot S_t + (1-p)\, d \cdot S_t,$$

which, together with (1.2.3), yields

$$p = \frac{e^{r\Delta t} - d}{u - d}. \tag{1.2.5}$$

Note that p is the risk-neutral probability, which does not depend on the true drift. To match, we see that, on the lattice,

$$\text{Var}\left[S_{t+\Delta t}\right] = S_t^2 + \left(pu^2 + (1-p)\, d^2\right) - S_t^2 e^{2r\Delta t}.$$

From (1.2.4) we also see

$$\text{Var}\left[S_{t+\Delta t}\right] = S_t^2 e^{2r\Delta t}\left(e^{\sigma^2\Delta t} - 1\right).$$

Using these two equations together, (1.2.5) and $u = 1/d$, we finally get

$$u = \frac{\left(1 - e^{2r\Delta t + \sigma^2\Delta t}\right) + \sqrt{\left(1 + e^{2r\Delta t + \sigma^2\Delta t}\right)^2 - 4e^{2r\Delta t}}}{2e^{r\Delta t}}.$$

Using a first-order expansion limited to the powers of order Δt, we may simplify the expression. Hence

$$u \approx 1 + \sigma\sqrt{\Delta t} + \frac{\sigma^2}{2}\Delta t.$$

But this expansion is the same expansion to the second order of $e^{\sigma\Delta t}$. We end up with the parametrization

$$u = e^{\sigma\sqrt{\Delta t}},\ d = e^{-\sigma\sqrt{\Delta t}},\ p = \frac{e^{r\Delta t} - d}{u - d},$$

which is known as CRR (Cox, Ross, and Rubinstein).

Assuming that the risk-free interest rate and volatility are constant in time, the parameters we have obtained apply to the entire lattice. To price an option, we should build (explicitly or implicitly) a lattice for the underlying asset prices, and then we should proceed backward in time. In fact, the option value is known at maturity, where it is given by the option payoff. Then we should apply equation (1.2.2) recursively, going backward one step at time, until we reach the initial node.

1.2.3 Simple binomial processes as diffusion approximations

Nelson and Ramaswamy in their paper [35] presents an approach to obtain simple binomial trees for several processes. A binomial approximation to a diffusion is defined as “computationally simple" if the number of nodes grows at most linearly in the number of time intervals. In the paper, it is shown how to construct computationally simple binomial processes that converge weakly to commonly employed diffusion in financial models. The convergence of the sequence of bond and European option prices from these processes to the corresponding values in the diffusion limit is also demonstrated.

1.2.3.1 Binomial diffusion approximations

Suppose we are given the stochastic differential equation

$$dy_t = \mu(y,t)\,dt + \sigma(y,t)\,dW_t \tag{1.2.6}$$

where $\{W_t, t \geq 0\}$ is a standard Brownian motion, $\mu(y,t)$ and $\sigma(y,t) \geq 0$ are the instantaneous drift and standard deviation of y_t, and y_0 is a constant. We wish to find a sequence of binomial processes that converges in distribution to (1.2.6). We then tackle the problem of constructing a sequence of binomial approximations, given a limit diffusion. To fix matters, take the interval $[0,T]$, and chop it into n equal pieces of length $h = T/n$. For each h consider a stochastic process $\{_h y_t\}$ on the interval $[0,T]$, which is constant between nodes and, at any given node, jumps up (down) some specified distance with probability q (resp. $1-q$). For example, if we set $q = 1/2$ and the up and down size equal to $\sqrt{h}$, it is well known that, as $n \to +\infty$, $\{_h y_t\}$ converges in distribution to a Brownian motion.

The probabilities of up or down jumps are specified as follows: define $q_h(y,hk), Y_h^+(y,hk)$ and $Y_h^-(y,hk)$ to be scalar valued functions defined in $\mathbb{R} \times [0,\infty)$ satisfying

$$0 \leq q_h(y,hk) \leq 1$$
$$-\infty < Y_h^-(y,hk) \leq Y_h^+(y,hk) < +\infty,$$

for all $y \in \mathbb{R}$ and all $k \in \{0,\ldots,n\}$. The stochastic process followed by $_h y_t$ is given by

$$\begin{aligned} {}_h y_0 &= y_0 \text{ for all } h, \\ {}_h y_t &= y_{kh} \text{ for all } kh \leq t < (k+1)h, \end{aligned} \tag{1.2.7}$$

$$\begin{aligned} P\left[{}_h y_{(k+1)h} = Y_h^+(y,hk) \,|hk, {}_h y_{kh}\right] &= q_h({}_h y_{kh}, bk), \\ P\left[{}_h y_{(k+1)h} = Y_h^-(y,hk) \,|hk, {}_h y_{kh}\right] &= 1 - q_h({}_h y_{kh}, bk), \\ P\left[{}_h y_{(k+1)h} = c\right] &= 0, \\ \text{for } c \neq Y_h^-(y,hk), \; c &\neq Y_h^+(y,hk) \end{aligned} \tag{1.2.8}$$

The stochastic process $_h y_t$ is a step function with initial value y_0 which jumps only at times $h, 2h, 3h, \ldots (n-1)h$. At each jump the process can make one of two possible moves: up to a value Y_b^+ or down to a value Y_h^-. Y_h^+, Y_h^- and q_h are allowed to depend on h, on the value of the process immediately before the jump ($_h y_{hk}$), and on the time index hk. By the statements in the previous equations, the process described is a Markov chain.

We apply a result from Stroock and Srinivasa Varadhan ([44]) which states conditions under which $\{_h y_t\}$ converges weakly when $h \to 0$ to the process in (1.2.6).

To use this result we need assumptions about both the limiting stochastic differential equation and the sequence of Markov chains defined above. The first two assumptions ensure that the limiting stochastic differential equation (1.2.6) is well behaved.

Assumption 1. *The functions $\mu(y,t)$ and $\sigma(y,t)$ are continuous, and $\sigma(y,t)$ is non-negative.*

Assumption 2. *With probability 1, a solution $\{y_t\}$ of the stochastic integral equation*

$$y_t = y_0 + \int_0^t \mu(y_s, s)\, ds + \int_0^t \sigma(y_s, s)\, dW_s,$$

exists for $0 < t < \infty$, and is distributionally unique.

Under Assumption 2, the distribution of the random process $\{y_t\}_{0 \leq t < T}$ is characterized by four things:

1. The starting point y_0
2. The continuity (with probability 1) of y_t as a random function of t
3. The drift function $\mu(y,t)$
4. The diffusion function $\sigma^2(y,t)$

If $\{{}_h y_t\}$ is convergent in distribution to $\{y_t\}$ when $h \to 0$, properties $1-4$ must be matched in the limit. Specifically, we require

1'. that ${}_h y_0 = y_0$, for all h

2'. that the jump size of ${}_h y_t$ become small at a sufficiently rapid rate as $h \to 0$

3'. that the drift of ${}_h y_t$ converges (in a sense to be made precise below) to (y,t)

4'. that the local variance of ${}_h y_t$ converges to $\sigma^2(y,t)$

Note that 1' is assured by (1.2.7). To ensure 2', we make the following assumption.

Assumption 3. *For all $\delta > 0$ and $T > 0$,*

$$\lim_{h \to 0} \sup_{\substack{|y| \leq \delta \\ 0 \leq t \leq T}} \left| Y_h^+(y,t) - y \right| = 0,$$
$$\lim_{h \to 0} \sup_{\substack{|y| \leq \delta \\ 0 \leq t \leq T}} \left| Y_h^-(y,t) - y \right| = 0.$$

For 3' and 4', define for any $h > 0$ the local drift $\mu_h(y,t)$ and the local second moment[1] $\sigma_h^2(y,t)$ of the binomial process described above by

$$\mu_h(y,t) = \left\{ q_h(y,t^*) \left[Y_h^+(y,t^*) - y \right] + (1 - q_h(y,t^*)) \left[Y_h^-(y,t^*) - y \right] \right\} / h$$
$$\sigma_h^2(y,t) = \left\{ q_h(y,t^*) \left[Y_h^+(y,t^*) - y \right]^2 + (1 - q_h(y,t^*)) \left[Y_h^-(y,t^*) - y \right]^2 \right\} / h$$

[1]This is not the local variance, because the moment is centered around y and not around the conditional mean. As $h \to 0$ however, the local variance and second moment approach the same limit.

with $t^* = h \cdot \lfloor t/h \rfloor$, where $\lfloor t/h \rfloor$ is the integer part of t/h. The next assumption requires that μ_h and σ_h^2 converge uniformly to μ and σ^2 on sets of the form $|y| \leq \delta$, $0 \leq t < T$.

Assumption 4. *For every $T > 0$ and every $\delta > 0$*

$$\lim_{h \to 0} \sup_{\substack{|y| \leq \delta \\ 0 \leq t \leq T}} |\mu_h(y,t) - \mu(y,t)| = 0,$$

and

$$\lim_{h \to 0} \sup_{\substack{|y| \leq \delta \\ 0 \leq t \leq T}} \left|\sigma_h^2(y,t) - \sigma^2(y,t)\right| = 0,$$

Theorem 1.20. *Under assumptions 1-4, $\{_h y_t\} \Rightarrow \{y_t\}$, where $\Rightarrow$ denotes weak convergence (i.e. convergence in distribution) and $\{y_t\}$ is the solution of (1.2.6)*

The intuition underlying the construction of a simple binomial sequence is uncomplicated. Suppose, following the suggestion in Cox and Rubinstein [18], we use the binomial jumps described by Figure (1.2.3) as the basic building block for a binomial tree, where

$$\begin{aligned} Y_h^+ &= y + \sqrt{h}\sigma(y,t) \\ Y_h^- &= y - \sqrt{h}\sigma(y,t) \\ q_h &= \frac{1}{2} + \sqrt{h}\mu(y,t) / (2\sigma(y,t)). \end{aligned}$$

In the previous equations, h is the time interval between jumps, and q_h is the probability of a jump to Y_b^+. The total displacement is $\sqrt{h}\left[-\sigma(y,t) + \sigma\left(Y_b^-, t+h\right)\right]$ if an up move follows a down move, and it is $\sqrt{h}\left[+\sigma(y,t) - \sigma\left(Y_b^+, t+h\right)\right]$ if a down move follows an up move. In general, these are not equal, so the branches of the binomial tree do not reconnect and the number of nodes doubles at each time step. However, whenever Assumptions 1-4 are satisfied by this binomial sequence (which is often the case), weak convergence will follow. But such a computationally complex tree is useless for purposes such as option pricing: after only 20 periods, the process could take more than a million different values, and after 40 periods, more than a trillion values. A computationally simple binomial representation would allow the process to take at most 21 and 41 values after 20 and 40 periods, respectively.

Note , however, that if $\sigma(y,t)$ is constant, then the displacements are equal-so computational simplicity is retained. This suggests that a transformation that purges the original stochastic differential equation (1.2.6) of conditional heteroskedasticity will permit us to construct a computationally simple tree.

1.2.3.2 Retaining computational simplicity:

To this end, consider a transformation $X(y,t)$, which is differentiable twice in y and once in t. We have, by Itô's lemma,

$$dX(y_t,t) = \left(\mu(y_t,t)\frac{\partial X(y_t,t)}{\partial y} + \frac{1}{2}\sigma^2(y_t,t)\frac{\partial^2 X(y_t,t)}{\partial y^2} + \frac{\partial X(y_t,t)}{\partial t}\right)dt + \left(\sigma(y_t,t)\frac{\partial X(y_t,t)}{\partial y}\right)dW_t. \tag{1.2.9}$$

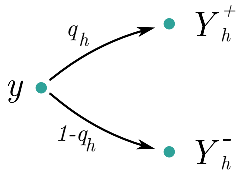

Figure 1.2.3: The simple binomial sequence.

Now choose $X(y,t)$ to satisfy

$$X(y,t) = \int \frac{dZ}{\sigma(Z,t)} \tag{1.2.10}$$

on the support of y. Then the term

$$\frac{\partial X(y,t)}{\partial y} \sigma(y,t)\, dW_t$$

in (1.2.9) becomes dW_t and the instantaneous volatility of the transformed process $x_t = X(y_t, t)$ is constant. In this case, we can develop a computationally simple binomial tree for x where the second moment of the local change in x is constant at every node. To arrive at the sequence of binomial processes on y, we transform from x back to y by defining

$$Y(x,t) = \{y : X(y,t) = x\}. \tag{1.2.11}$$

It is easy to see that $\partial Y / \partial x = \sigma(y,t)$ and, by Assumption 1, this means that $Y(x,t)$ is weakly monotone in x for a fixed t. Then we can use the transform in (1.2.11) to define a tree for y, so that

$$\begin{aligned} Y_h^+(x,t) &= Y\left(x + \sqrt{h}, t+h\right) \\ Y_h^-(x,t) &= Y\left(x - \sqrt{h}, t+h\right). \end{aligned}$$

Note that the tree for y has inherited the computational simplicity that the tree for x displays. Using the fact that

$$\frac{\partial Y(x,t)}{\partial x} = \sigma(Y(x,t),t),$$

a Taylor's series expansion of Y_h^+ and Y_h^- around $h = 0$ yields

$$\begin{aligned} Y_h^{\pm}(x,t) &= Y(x,t) \pm \sigma(Y(x,t),t)\sqrt{h} + O(h), \\ \sigma^2(Y(x,t),t) &= \sigma^2(Y(x,t),t) + O\left(\sqrt{h}\right). \end{aligned}$$

This shows that the local second moment of ${}_hy_t$ converges to the instantaneous variance $\sigma^2(y,t)$ as $h \to 0$. Finally, to get the local drift to match the drift of the limit diffusion, we need

$$\mu_h(y,t) \to \mu(y,t)$$

uniformly on $\{(y,t) : |y|, t < \delta\}$, for every $\delta > 0$. We tentatively choose

$$q_h = \frac{h\mu(Y(x,t),t) + Y(x,t) - Y_h^-(x,t)}{Y_h^+(x,t) - Y_h^-(x,t)} \tag{1.2.12}$$

which, if it is a legitimate probability (i.e., between 0 and 1) sets the local drift exactly equal to the drift of the limiting diffusion 1.2.6. This device - the use of a transform, its inverse, and the choice of the probability q_h - enables one to turn to construct a computationally simple binomial approximation. It turns out to be a useful device in many commonly employed diffusion in finance, where a transformation like (1.2.10) is readily available. A straightforward example of this transformation is for the Lognormal diffusion, where $\mu(y,t) = \mu y$ and $\sigma(y,t) = \sigma y$. The transformation is simple $X(y) = \sigma^{-1}\ln(y)$, and the inverse transformation is $Y(x) = e^{\sigma x}$. This was the transformation employed by Cox, Ross, and Rubinstein to obtain a computationally simple tree. Such transformations can be made for other diffusions, even if their drift and diffusion depend on t.

We must sometimes also allow x to jump up or down by a quantity greater than $\sqrt{h}$ in order to maintain the drift rate. Furthermore, the diffusion may have a boundary at 0 (or some other value). At such a boundary $\sigma(\cdot,t) = 0$ and the transformation may need to be modified.

1.2.3.3 Retaining computational simplicity: a general treatment

The principal complications that arise in implementing the strategy come from singularities in $\sigma(\cdot,\cdot)$; for example, $\sigma(y^*,t) = 0$ for some y^*. Such singularities are usually associated with boundaries on the support of the process, and often arise in financial economics; for example, with limited liability and in the absence of arbitrage, zero must be a lower boundary for stock prices and nominal interest rates.

There is a large variety of possible boundary behaviors so it is necessary to confine our attention to the cases likely to be most useful in finance. First, we consider the case in which $\sigma(\cdot,\cdot)$ has no singularities on $\mathbb{R} \times [0,\infty)$. Then, we consider the case in $\sigma(0,t) = 0$ and $\mu(0,t) \geq 0$, for all t, implying a lower boundary at zero on the support of the limiting diffusion.

Case 1. No singularities in $\sigma(y,t)$ As we did before, we define $X(y,t)$ along with x values corresponding to extreme values of y:

$$\begin{aligned} X(y,t) &= \int \frac{dZ}{\sigma(Z,t)}, \\ x^U(t) &= \lim_{y\to+\infty} X(y,t), \\ x^L(t) &= \lim_{y\to-\infty} X(y,t). \end{aligned} \tag{1.2.13}$$

The following assumption is convenient, and can be relaxed at the expense of simplicity.

Assumption 5. *The values $x^U(t)$ and $x^L(t)$ are constants.*

The definition of the inverse transform is now modified to read

$$y(x,t)=\begin{cases} y : X(y,t)=x, & \text{if } x^L < x < x^U \\ +\infty & \text{if } x^U \le x \\ -\infty & \text{if } x \le x^L. \end{cases} \tag{1.2.14}$$

We also set

$$q_h^*(x,t)=\max\left[0,\min\left[1,q_h(x,t)\right]\right]. \tag{1.2.15}$$

To verify Assumptions 1 and 2 for the current case, we employ Assumptions 6 and 7.

Assumption 6. *The functions $\mu(y,t)$ and $\sigma^2(y,t)$ are continuous everywhere. For every $R>0$ and every $T>0$ there is a number $\Delta_{T,R}>0$ such that*

$$\Delta_{T,R} \le \inf_{\substack{0\le t\le T\\ |y|\le R}} \sigma(y,t) \tag{1.2.16}$$

Relation (1.2.16) is a non-singularity assumption: it ensures that $\sigma(y,r)$ is bounded away from zero except at $t=\infty$ and/or $|y|=\infty$.

We must also ensure that the process for y does not explode to infinity in finite time.

Assumption 7. *Analytical solutions of (1.2.6) share the property that, for all T, $0<T<\infty$,*

$$\lim_{B\to\infty} P\left(\sup_{0\le t\le T}|y_t|>B\right)=0.$$

To state the following theorem, we need one last assumptions of regularity.

Assumption 8. *The first and second order partial derivatives*

$$\sigma_y,\sigma_t,\sigma_{yt},\sigma_{tt},Y_x,Y_{xx},Y_t,Y_{tt},Y_{xt}$$

are well defined and locally bounded for all $(y,t)\in\mathbb{R}\times[0,\infty)$.

Theorem 1.21. *Let Assumptions 5-8 hold. For $h>0$, define the x as a simple binomial tree with $\sigma_x=1$, with ${}_hx_0=X(y_0,0)$, and the transition for the x process given by*

$${}_hx_{h(k+1)}=\begin{cases} {}_hx_{hk}+\sqrt{h} \text{ with probability } q_b^*({}_hx_{hk},hk), \\ {}_hx_{hk}-\sqrt{h} \text{ with probability 1-} q_b^*({}_hx_{hk},hk). \end{cases}$$

Define the y-tree as a simple binomial tree whose state are obtained from those of the x-tree by inverse transformation. That is, for $h>0$, define ${}_hy_t=Y({}_hX_{hk},hk)$, for $hk\le t<h(k+1)$. By construction, $\{{}_hy_t\}$ is computationally simple. Then $\{{}_hy_t\}\Rightarrow\{y_t\}$ as $h\to 0$, where $\{y_t\}$ is a solution on (1.2.6).

For the proof of this theorem, see [35].

Case 2. A singularity at $y = 0$: $\sigma(0,t) = 0, \mu(0,t) \geq 0$. In this case the diffusion coefficient vanishes at the lower boundary (zero), but the drift rate might serve to return to the process above it. The lower limit for x is redefined as

$$x^L(t) = \lim_{y \to 0} X(y,t), \tag{1.2.17}$$

and the inverse transform (which is now a weakly monotone function of x) defined as

$$Y(x,t) = \begin{cases} y : X(y,t) = x & \text{if } x^L < x < x^U \\ +\infty & \text{if } x^U < x \\ 0 & \text{if } x \leq x^L. \end{cases} \tag{1.2.18}$$

As before we assume that x^L and x^U do not depend on t.

An important aspect of Case 2 relates to the step sizes: thus far they are (approximately) proportional to $\sigma(y,t)$. But if $\sigma(y,t)$ is very small near $y = 0$ and $\mu(y,t)$ is not small, we may need to take multiple jumps in this region in order to match the drift of the limit diffusion.

Choose $x^B > x^L$, and define the function $J_h^+(x,t)$ as

$$J_h^+(x,t) = \begin{cases} \text{the smallest, odd, positive, integer } j \text{such that} & \\ Y\left(x + j\sqrt{h}, t+h\right) - Y(x,t) \geq \mu(Y(x,t),t) \cdot h & \text{if } x < x^B \\ 1 & \text{if } x \geq x^B \end{cases}$$

$J_h^+(x,t)$ is the minimum number of upward jumps that keeps the jump probability p_h less than 1 without censoring; and it is odd so that the jump moves the process to an existing node on the tree. By permitting these multiple jumps in a restricted region near 0, we retain computational simplicity; at large values of y we disallow multiple upward jumps, because if J_h^+ is unbounded it might increase the number of nodes at a rate rapid enough to affect computational simplicity. Similarly, define $J_h^-(x,t)$ by

$$J_h^-(x,t) = \begin{cases} \text{the smallest, odd, positive, integer } j \text{such that} \\ \text{either (a) } Y(x,t) - Y\left(x - j\sqrt{h}, t+h\right) \leq \mu(Y(x,t),t) \cdot h \\ \text{or (b) } Y\left(x - j\sqrt{h}, t+h\right) = 0 \end{cases}$$

Here, $J_h^-(x,t)$ is the minimum number of downward jumps that either keeps the probability q_h positive (without censoring) or forces the down-state value for Y_h^- to zero. The transitions in the value for y are then restated as

$$Y_h^{\pm}(x,t) = Y\left(x \pm J_h^{\pm}(x,t) \cdot \sqrt{h}, t+h\right), \tag{1.2.19}$$

and we retain the definition of q_h^* given in (1.2.12) and (1.2.15).

Assumptions 7 and 8 need to be replaced by the following ones.

Assumption 9. *Let $\sigma(y,t)$ and $\mu(y,t)$ be continuous on $\mathbb{R}\times[0,\infty)$. There exists an increasing, non-negative function $\rho(u)$ from $[0,\infty)$ to $[0,\infty)$ such that*

$$\rho(u) > 0, \text{ for } u > 0,$$

$$\lim_{\varepsilon\to 0}\int_{\varepsilon}^{1}[\rho(u)]^{-2}\,du = \infty.$$

Further, for every $R > 0$ and $T > 0$, there exists a number $\Delta_{T,R} > 0$ such that

$$\sup_{\substack{|y^*|\le R, |y|\le R\\ 0\le t\le T}} |\sigma(y^*,t) - \sigma(y,t)| \le \Delta_{T,R}\,\rho(|y-y^*|),$$

$$\sup_{\substack{|y^*|\le R, |y|\le R\\ 0\le t\le T}} |\mu(y^*,t) - \mu(y,t)| \le \Delta_{T,R}\,|y-y^*|,$$

Assumption 10. *On every compact subset of $\{(y,t): 0<y<\infty, 0\le t<\infty\}$, $\sigma_y, \sigma_t, \sigma_{yt}$ and σ_{tt} exist and are bounded, and $\sigma(y,t)$ is bounded and bounded away from zero. There exists a $\Delta > 0$ such that for every $T > 0$*

$$\inf_{\substack{0\le t\le T\\ 0\le y\le \Delta}} \sigma_y(y,t) > 0.$$

Furthermore, Y_{xx}, Y_t, Y_{tt} and Y_{xt} exist for all $(y,t)\in[0,\infty)\times[0,\infty)$ and are bounded on bounded sets. For all $t\ge 0$, $\sigma(0,t)=0$ and $\mu(0,t)\ge 0$.

The theorem for Case 2 can now be stated.

Theorem 1.22. *Let assumptions 5,7,9 and 10 hold, and assume $y_0 > 0$. Define ${}_hx_{hk}$ and ${}_hy_t$ as in Theorem 1.21, replacing relations (1.2.13) and (1.2.14) with relations (1.2.17) and (1.2.18), and using (1.2.19) to define $Y_h^{\pm}$. Then $\{{}_hy_t\}\Rightarrow\{y_t\}$ as $h\to 0$; and if $x^B<\infty$, $\{{}_hy_t\}$ is computationally simple by construction. Further, 0 bounds the support of $\{{}_hy_t\}$ and y_t from below:*

$$P\left(\inf_{0\le t<\infty h} y_t < 0\right) = P\left(\inf_{0\le t<\infty} < 0\right) = 0$$

Theorems 1.21 and 1.22 show how to construct computationally simple approximations for diffusions encountered in many applications in finance. The proofs of these Theorems are available in [35].

1.2.3.4 The Hull and White model

The Hull and White case is a well known model in finance to describe interest rates dynamics: see Section 1.1.7.

It is well known that the short rate process r can be written as

$$r_t = \omega X_t + \beta(t),$$

where X is a stochastic process given by

$$dX_t = -kX_t dt + dW_t \; X_0 = 0,$$

and $\beta(t)$ is a deterministic function. The X_t process is an example of Ornstein-Uhlenbeck process,

$$dy_t = \beta(\alpha - y_t)\, dt + \sigma dW_t \tag{1.2.20}$$

and for those processes we explain the tree construction. We define a sequence $\{_h y_t\}$ of binomial approximations to (1.2.20) with common initial value y_0 and

$$Y_h^+(y,t) = y + \sigma\sqrt{h},$$
$$Y_h^-(y,t) = y - \sigma\sqrt{h},$$

and let

$$q_h = \begin{cases} \frac{1}{2} + \sqrt{h}\beta(\alpha - y)/(2\sigma) & \text{if } 0 \le 1/2 + \sqrt{h}\beta(\alpha - y)/(2\sigma) \le 1 \\ 0, & \text{if } 1/2 + \sqrt{h}\beta(\alpha - y)/(2\sigma) < 0 \\ 1 & \text{otherwise.} \end{cases}$$

1.2.3.5 The Heston model

The Heston model is one of the most used model to represent stochastic volatility: for more details see Section 1.1.6.

The SDE for the volatility in this model is

$$dv_t = k(\theta - v_t)\, dt + \omega\sqrt{v_t} dW, \; v_0 = \bar{v}_0,$$

with $k \ge 0$, $\mu \ge 0$, and the initial value $\bar{v}_0$ is a non negative constant. The necessary transformation is

$$X(v) = \int \frac{dZ}{\sigma\sqrt{Z}} = \frac{2\sqrt{v}}{\sigma},$$

with $x_0 = X(v_0)$. Zero is a lower boundary for v. As outlined in previous Section, we define the inverse transform

$$V(x) = \begin{cases} \sigma^2 x^2/4 & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Because the drift in (1.2.20) does not vanish as $r \to 0$, the value 0 is not an absorbing state for v unless either k or θ equals zero. This illustrates why it was necessary to introduce multiple jumps. We define

$$J_h^+(x) = \begin{cases} \text{the smallest, odd, positive, integer } j \text{such that} \\ 4hk\theta/\sigma^2 + x^2(1 - kh) < \left(x + j\sqrt{h}\right)^2 \end{cases}$$

$$J_h^-(x) = \begin{cases} \text{the smallest, odd, positive, integer } j \text{such that} \\ \text{either } 4hk\theta/\sigma^2 + x^2(1 - kh) \ge \left(x - j\sqrt{h}\right)^2 \\ \text{or } x - j\sqrt{h} \le 0, \end{cases}$$

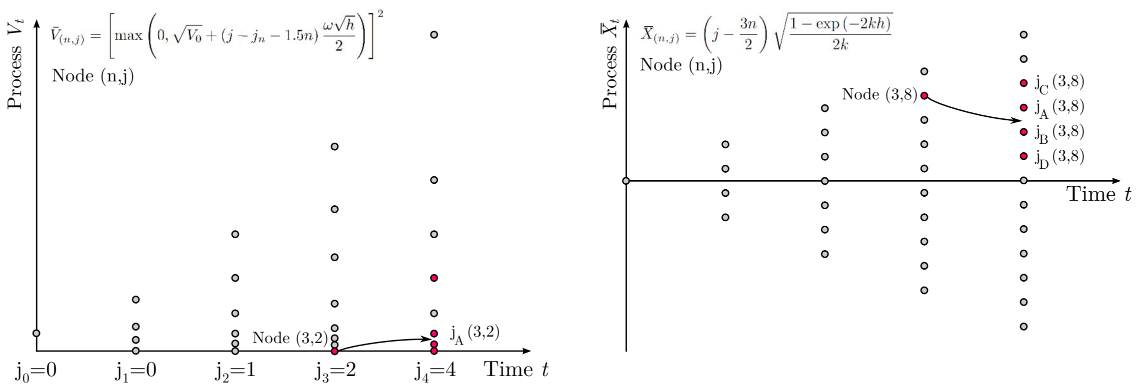

Figure 1.2.4: The trees for Heston and Hull-White models.

$$V_h^{\pm}(x) = V\left(x \pm J_h^{\pm} \cdot \sqrt{h}\right),$$

$$q_h(x) = \begin{cases} \left[\frac{hk(\theta - V(x)) + V(x) - V_h^-(x)}{V_h^+(x) - V_h^-(x)}\right] & \text{if } R_h^+(x) > 0 \\ 0 & \text{otherwise.} \end{cases}$$

We choose $J_h^{\pm}(x)$ in order to guarantee that $0 \leq q_h(x) \leq 1$, in such a way that the local drift converges to the diffusion limit.

1.2.4 Quadrinomial trees

The trees for the Heston model and BS HW model can be obtained from Appolloni et al. [4] or Nelson and Ramaswamy [35], as we described in the previous section. In this case, the trees are simple binary trees: the node values, and the transition probabilities are set in order to match an approximation of the first two moments of the process. This kind of tree perform well on short maturity, but the approximation errors accumulate on long maturities. Because of this error that accumulates, the convergence of the algorithm proved to be slow for long maturity options. Therefore, it was necessary to rethink the trees: the main aim was to set up trees which matched exactly some moments of the processes to be diffused. Here we present two trees (see Figure 1.2.4), one for stochastic volatility and one for stochastic interest rate. They are simple quadrinomial trees, and they are built to match the first 3 moments of the stochastic processes.

We suppose to fix a number $N > 0$, and we define $h = T/N$.

1.2.4.1 The General Case

Let Z be a Brownian motion, and let G be a Gaussian process, following

$$dG_t = a(G_t)\,dt + b\,dZ_t,$$

with variance that depends only by the time lapse, i.e. $G_{t+s} - G_s|\mathcal{F}_s \sim \mathcal{N}\left(\mu(t, G_s), \sigma^2(t)\right)$. We remark that $\mu(t, G_s)$ is the expectation and $\sigma^2(t)$ is the variance of the increment of the process:

$$\mu(t, G_s) = \mathbb{E}[G_{s+t} - G_s|\mathcal{F}_s] = \mathbb{E}\left[\int_s^{t+s} a(G_u)\,du + \int_s^{t+s} bdZ_u|\mathcal{F}_s\right] = \mathbb{E}\left[\int_s^{t+s} a(G_u)\,du|\mathcal{F}_s\right],$$

$$\sigma^2(t) = Var[G_{s+t} - G_s|\mathcal{F}_s] = \mathbb{E}\left[\left(\int_s^{t+s} a(G_u)\,du + \int_s^{t+s} bdZ_u\right)^2 |\mathcal{F}_s\right] - \mathbb{E}\left[\int_s^{t+s} a(G_u)\,du|\mathcal{F}_s\right]^2.$$

We show how to build a simple quadrinomial tree that can match the first three moments.

We define a quadrinomial tree. Let's fix a maturity T, and the number of steps N. Each node will be denoted by $G_{(n,j)}$ where n runs from 0 to N, and j from 0 to $3n$. Let $h = T/N$. The value of each node is

$$G_{(n,j)} = G_0 + (j - 1.5n)\sqrt{\sigma^2(h)}.$$

We remember the first three moments of the process G:

$$M_1 = \mathbb{E}[G_{t+h} - G_t|\mathcal{F}_t] = \mu(h, G_t),\; M_2 = \mathbb{E}\left[(G_{t+h} - G_t)^2 |\mathcal{F}_t\right] = \mu^2(h, G_t) + \sigma^2(h),$$

$$M_3 = \mathbb{E}\left[(G_{t+h} - G_t)^3 |\mathcal{F}_t\right] = \mu^3(h, G_t) + 3\mu(h, G_t)\,\sigma^2(h).$$

Let's fix a node $G_{(n,j)}$. To be brief, μ will denote $\mu\left(h, G_{(n,j)}\right)$ and σ will denote $\sqrt{\sigma^2(h)}$. We suppose that the expected value μ falls between the values of the nodes at time $(n+1)h$. This hypothesis can be obtained assuming that the time step h is small enough.

We define

$$j_A(n, j) = \text{ceil}\left[\frac{G_0 - \mu}{\sigma} + 1.5(n+1)\right],$$

i.e. the first node in the next time step level whose value is bigger than the mean of the process. This can be seen in Figure 1.2.4 (both sides): the arrow points out to the expected value of the process, and $j_A(n, j)$ is marked on the Figure. Let

$$j_B(n, j) = j_A(n, j) - 1,\; j_C(n, j) = j_A(n, j) + 1,\; j_B(n, j) = j_D(n, j) - 2.$$

To be brief we will only write j_A, j_B, j_C, j_D, and G_A will be $G_A = G_{(n+1, j_A)}$, and the same for the other letters: this is clear in Figure 1.2.4, on the right side.

We can now define a Markovian discrete time process $\hat{G}_n$, $n = 0, \dots, N$ with $\hat{G}_0 = G_{(0,0)}$ and we suppose that if $\hat{G}_n = G_{(n,j)}$, then it can move to G_A, G_B, G_C, G_D, according to the following probabilities

$$p_A = P\left[\hat{G}_{n+1} = G_A|\hat{G}_n = G_{(n,j)}\right] = \frac{(G_A - \mu)\left((G_A - \sigma - \mu)^2 + \sigma^2\right)}{2\sigma^3},$$

$$p_B = P\left[\hat{G}_{n+1} = G_B|\hat{G}_n = G_{(n,j)}\right] = \frac{(\mu - G_A + \sigma)\left((G_A - \mu)^2 + \sigma^2\right)}{2\sigma^3},$$

$$p_C = P\left[\hat{G}_{n+1} = G_C|\hat{G}_n = G_{(n,j)}\right] = \frac{(\mu - G_A + \sigma)\left((G_A - \sigma - \mu)^2 + 2\sigma^2\right)}{6\sigma^3},$$

$$p_D = P\left[\hat{G}_{n+1} = G_D | \hat{G}_n = G_{(n,j)}\right] = \frac{2\sigma^2(G_A - \mu) + (G_A - \mu)^3}{6\sigma^3}.$$

Since $G_A - \sigma < \mu \leq G_A$, we can easily show that these probabilities are well defined: all in $[0,1]$, their sum is equal to 1, and the first three moments of the variable $\left(\hat{G}_{n+1} | \hat{G}_n = G_{(n,j)}\right)$ are equal to the first three moments of the variable $\left(G_{t+h} | G_t = G_{(n,j)}\right)$.

Now, we approximate the process G by a discrete process $\bar{G}$ that is constant in each time lapse, and is defined as $\bar{G}_t = \hat{G}_{\lfloor t/N \rfloor}$. The weak convergence of this tree can be proved as in Nelson and Ramaswamy [35].

1.2.4.2 The Heston Model

The Heston process (3.3.1) for volatility has no constant variance and isn't Gaussian. We consider the process obtained by the square root:

$$d\sqrt{V_t} = \frac{4k\left(\theta - \sqrt{V_t}^2\right) - \omega^2}{8\sqrt{V_t}} dt + \frac{\omega}{2} dZ_t.$$

We approximate it with a Gaussian process with variance $\frac{\omega^2}{4}dt$. This approximation is helpful to define the grid of states-space for the Markov chain: inspired by [35], we define

$$j_n = \max\left(0, \text{floor}\left(1.5n - \frac{2\sqrt{V_0}}{\omega\sqrt{h}}\right)\right),$$

and we set

$$\bar{V}_{(n,j)} = \left(\max\left(0, \sqrt{V_0} + (j + j_n - 1.5n)\frac{\omega\sqrt{h}}{2}\right)\right)^2.$$

for $j = 0, \ldots, 3n - j_n$. The shift due to j_n helps to reject the many nodes with value equal to zero: if $j_n > 0$, then $\bar{V}_{(n,0)} = 0$ and $\bar{V}_{(n,1)} > 0$.

We would remark that the process $\bar{V}_t$ approximates V_t and not $\sqrt{V_t}$: the moments matching is done according to the moments of the process V_t.

We fix now the values of n and j. The discrete process $\bar{V}$ can jump from a node to another, as in a Markovian chain. We show now how to find the possible upcoming nodes.

The first three moments for the Heston process can be found in Alfonsi [2]:

$$\psi(h) = {}^{(1-e^{-kh})}/_{k}, \quad M_1 = \mathbb{E}\left[V_{t+h} | V_t = v\right] = ve^{-kh} + \theta k\psi(h),$$

$$M_2 = \mathbb{E}\left[(V_{t+h})^2 | V_t = v\right] = M_1^2 + \omega^2\psi(h)\left[\theta k\psi(h)/2 + ve^{-kh}\right],$$

$$M_3 = \mathbb{E}\left[(V_{t+h})^3 | V_t = v\right] = M_1M_2 + \omega^2\psi(h)\left[2v^2e^{-2kh} + \psi(h)\left(k\theta + \frac{\omega^2}{2}\right)\left(3ve^{-kh} + \theta k\psi(h)\right)\right].$$

Then, we can proceed as in the general case. Anyway, the grid we're using is based on an approximation: so the probabilities obtained solving the linear system may not be positive. If

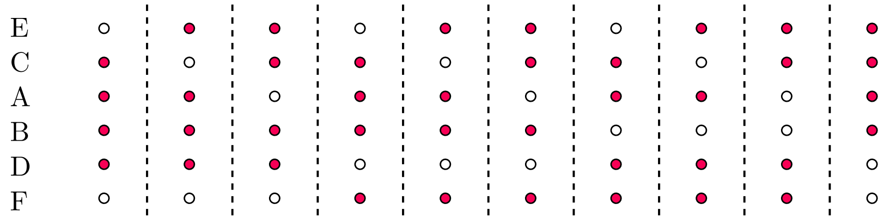

Figure 1.2.5: The possible combinations used to get positive probabilities in the Heston model tree. The red points correspond to the used nodes.

we get a negative transition probability for a given node, we try another combination of outcoming nodes, replacing one (or two) of the nodes A, B, C, D with one (or two) close to them. The nodes A or C may be replaced by a node E defined as the first node bigger than C, and the nodes B or D may be replaced with a node F, defined as the smallest before node D. This gives rise to 9 combinations to be tested. If the starting node is small and the node D verifies $j_D = j_n$ we could not do this last change because there would be no F node. In this case we allow the node D to be replaced by the node E: see Figure 1.2.5.

If these attempts don't give a positive result (negative probabilities), we give up trying to match the first three moments, and we are content to match an approximation of the first two as in [35], thus ensuring the weak convergence. In this case, we only use the nodes A, B, C, D: we define

$$p_{AB} = \frac{\mu - G_{n+1,j_B}}{G_{n+1,j_A} - G_{n+1,j_B}}, \quad p_{BA} = 1 - p_{AB},$$
$$p_{CD} = \frac{\mu - G_{n+1,j_D}}{G_{n+1,j_C} - G_{n+1,j_D}}, \quad p_{DC} = 1 - p_{CD},$$

and

$$p_A = \frac{5}{8} p_{AB}, \; p_B = \frac{5}{8} p_{BA}, \; p_C = \frac{3}{8} p_{CD}, \; p_D = \frac{3}{8} p_{DC}.$$

It is possible to show that the first moment of this variable is equal to M_1, and as $h \to 0$ the second moment approaches to M_2, ensuring the convergence, as proved in [35].

In all our numerical tests, this last option (matching only two moments) has never been necessary: changing the nodes, all moments were matched with positive probabilities.

1.2.4.3 The Hull-White Model

The process X in (3.3.2) is Gaussian. As shown in Ostrovski [39] the variables X_t and $\int_s^t X_y dy$ are bivariate normal distributed conditionally on X_s with well known mean and variance. We define

$$X_{(n,j)} = \left(j - \frac{3n}{2}\right) \sqrt{\frac{1 - \exp\left(-2kh\right)}{2k}}, \; n = 0, \ldots, N \text{ and } j = 0, \ldots, 3n.$$

Let's fix a node $X_{(n,j)}$. We define

$$H = \exp\left(-kh\right), \; K = \sqrt{\frac{1 - \exp\left(-2kh\right)}{2k}}, \; M_1 = X_{(n,j)} H,$$

$$j_A = \text{ceil}\left[\frac{M_1}{K} + \frac{3(n+1)}{2}\right],\ X_A = X_{(n+1,j_A)}.$$

The transition probabilities are given by

$$p_A = \tfrac{(X_A - M_1)}{2K^3}\left(K^2 + (K + M_1 - X_A)^2\right), \qquad p_B = \tfrac{(K+M_1-X_A)}{2K^3}\left(K^2 + (M_1 - X_A)^2\right),$$
$$p_C = \tfrac{(K+M_1-X_A)}{6K^3}\left(2K^2 + (K + M_1 - X_A)^2\right), \quad p_D = \tfrac{(X_A-M_1)}{6K^3}\left(2K^2 + (M_1 - X_A)^2\right).$$

1.3 Partial differential equations methods

This part is inspired by [1] and [7].

Numerical methods based on partial differential equations (PDEs) in applied finance are not very popular. Indeed, the models are therefore much more natural. Stochastic methods are also often simpler to implement than the algorithms used for solving the related PDEs. However, when it is possible to efficiently discretize the PDE (which is not always the case, the typical counter example being high-dimensional problems), these algorithms are usually much more efficient. Moreover, the solution to the partial differential equation gives more information. In the context of options pricing with constant parameters, one obtains for example the price of the option for all values of the maturity and for all spot prices, while the probabilistic formulation typically gives the value of the option for fixed maturity and fixed spot prices. In particular, this is useful for computing derivatives of the option's price (the so-called "greeks").

The PDEs obtained in Finance have several characteristics. First, they are posed on a bounded domain in time $[0, T]$, with typically a singular final condition at the maturity $t = T$, and very often in an unbounded domain in the spot variable, which leads to impose suitable "boundary conditions" at infinity to get well posed problems and to use appropriate numerical approximations (truncation to a bonded domain and artificial boundary conditions). These PDEs are usually of parabolic type, but often with degenerate diffusion. Because of operational constraints, the numerical methods for discretization of the PDE must be sufficiently fast and accurate to be useful in practice. These peculiarities of PDEs in Finance explain the need for up-to-date, and sometimes involved, numerical methods.

1.3.1 The Black and Scholes PDE for European options

We adopt the standard Black and Scholes model with a risky asset whose price at time t is S_t and a risk free bond B_t . The underlying security evolves in accordance with the Itô process

$$dS = \mu S dt + \sigma S dW,$$

while the bond evolves in accordance with

$$dB = rBdt$$

where r is the risk free rate.

The option value $\mathcal{V}$ at a given time depends on the underlying value . Then, by Itô's Lemma,

$$d\mathcal{V} = \left(\frac{\partial \mathcal{V}}{\partial t} + \mu S \frac{\partial \mathcal{V}}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \mathcal{V}}{\partial S^2}\right) dt + \left(\sigma S \frac{\partial \mathcal{V}}{\partial S}\right) dW.$$

We have written $S = S(t)$, $B = B(t)$, $\mathcal{V} = \mathcal{V}(t)$ and $dW = dW(t)$ for notational convenience. We also assume the portfolios are self financing, which implies that changes in portfolio value are due to changes in the value of the three instruments, and nothing else. Under this setup, any of the instruments can be replicated by forming a replicating portfolio of the other two instruments, using the correct weights.

We set up a self-financing portfolio Π that is comprised of one option and an amount Δ of the underlying stock, such that the portfolio is riskless, i.e. that is insensitive to changes in the price of the security. Hence the value of the portfolio at time t is $\Pi(t) = \mathcal{V}(t) + \Delta S(t)$. The self-financing assumption implies that $d\Pi = d\mathcal{V} + \Delta dS$ so we can write

$$\begin{aligned} d\Pi &= d\mathcal{V} + \Delta dS \\ &= \left(\frac{\partial \mathcal{V}}{\partial t} + \mu S \frac{\partial \mathcal{V}}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \mathcal{V}}{\partial S^2} + \Delta \mu S\right) dt + \left(\sigma S \frac{\partial \mathcal{V}}{\partial S} + \Delta \sigma S\right) dW. \end{aligned} \tag{1.3.1}$$

The portfolio must have two features. The first is that it must be riskless, which implies that the second term involving the Brownian motion dW is zero so that $\Delta = -\frac{\partial \mathcal{V}}{\partial S}$. Substituting for Δ in equation (1.3.1) implies that the portfolio follows the process

$$d\Pi = \left(\frac{\partial \mathcal{V}}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \mathcal{V}}{\partial S^2}\right) dt.$$

The second feature is that the portfolio must earn the risk free rate. This implies that the diffusion of the riskless portfolio is $d\Pi = r\Pi dt$. Hence we can write

$$\begin{gathered} d\Pi = r\Pi dt \\ \left(\frac{\partial \mathcal{V}}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \mathcal{V}}{\partial S^2}\right) dt = r\left(\mathcal{V} - \frac{\partial \mathcal{V}}{\partial S} S\right) dt \end{gathered}$$

Dropping the dt from both sides and re-arranging yields the Black Scholes PDE:

$$\frac{\partial \mathcal{V}}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \mathcal{V}}{\partial S^2} + rS\frac{\partial \mathcal{V}}{\partial S} - r\mathcal{V} = 0. \tag{1.3.2}$$

The portion of share to be held, Δ, is delta, also called the hedge ratio. The derivation stipulates that in order to hedge the single option, we need to hold Δ shares of the stock. This is the principle of delta hedging.

Hereinafter we treat the Black Scholes PDE (1.3.2). The payoff at maturity is usually described by a function ϕ that depend only by the terminal underlying value. This gives a terminal condition

$$\begin{cases} \frac{\partial \mathcal{V}}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \mathcal{V}}{\partial S^2} + rS\frac{\partial \mathcal{V}}{\partial S} - r\mathcal{V} = 0 \\ \mathcal{V}(T, S) = \phi(S) \end{cases}$$

This is a parabolic equation and can be solved by finite difference method.

1.3.2 Numerical schemes for Black Scholes PDE

We introduce now the finite difference method on the simple PDE (1.3.2).

We first concentrate on the discretization of (1.3.2) with respect to the variable S. The principle is to divide the interval $[0, S_{max}]$ into I intervals of length $\Delta S = S_{max}/I$ (where S_{max} has to be chosen large enough), and to approximate the derivatives by finite differences.

A possible semi-discretization of (1.3.2) if: for $i \in \{0, 1, \dots, I\}$,

$$\begin{cases} \frac{\partial \mathcal{V}_i}{\partial t} + \frac{1}{2}\sigma^2 S_i^2 \frac{\mathcal{V}_{i+1} - 2\mathcal{V}_i + \mathcal{V}_{i-1}}{\Delta S^2} + rS_i \frac{\mathcal{V}_{i+1} - \mathcal{V}_{i-1}}{2\Delta S} - r\mathcal{V}_i = 0 \\ \mathcal{V}_i(T) = \phi(S_i), \end{cases} \tag{1.3.3}$$

where $S_i = i\Delta S$ denotes the i-th discretization point, and $\mathcal{V}_i(t)$ is intended to be an approximation of $\mathcal{V}(t, S_i)$.

Now, (1.3.3) is a system of coupled ordinary differential equations (ODEs). The generalization to the case of a time and spot dependent r or σ is straightforward. Notice that for $S = 0$, P_0 can be solved independently (since $S_0 = 0$): $\mathcal{V}(t, 0) = \phi(0) \exp\left(-\int_t^T r ds\right)$.

In order to obtain a solution of the whole system of ODEs, one needs to define an appropriate boundary condition at $S = S_{max}$. Indeed (1.3.3) taken at $i = I$ involves $\mathcal{V}_{I+1}$ which is *a priori* not defined.

There are basically two methods to define an appropriate boundary condition at $S = S_{max}$. The first one consists of using some *a priori* knowledge on the values of $V(t, S)$ when S is large and making some approximations of $\mathcal{V}(t, S_{max})$. In this case the value of $\mathcal{V}_I$ is given as a data (this is a so-called Dirichlet boundary condition), and the unknowns are $(\mathcal{V}_i)_{0 \le i \le I-1}$. For example, in the case of a put option ($\phi(S) = (S - K)^-$) (resp. a call option with $\phi(S) = (S - K)^+$), it is known that $\lim_{S \to +\infty} \mathcal{V}(t, S) = 0$ (resp. $\mathcal{V}_I(t) = S_{max} - K \exp\left(-\int_t^T r ds\right)$). The error introduced by these artificial boundary conditions can be estimated.

Another method is based on some knowledge on the asymptotic behavior of the derivatives of $\mathcal{V}$. For example, in the case of the put option, one can use the so-called homogeneous Neumann boundary condition which writes $\partial\mathcal{V}/\partial S(t, S_{max}) = 0$ at the continuous level, and $\mathcal{V}_{I+1}(t) = \mathcal{V}_I(t)$ at the discrete level. In this case, the unknowns are $(\mathcal{V}_i)_{0 \le i \le I}$.

For both methods, S_{max} should chosen sufficiently large. In practice, the quality of the method may be assessed by measuring how sensitive the result is to the value of S_{max}.

Let us now consider the time discretization. Here again, the idea is to divide the time interval $[0, T]$ into N intervals of length $\Delta t = T/N$ and to replace the time derivative by a finite difference. Three numerical methods are classically used:

$$\text{(EE)} \begin{cases} \frac{\mathcal{V}_i^{n+1}-\mathcal{V}_i^n}{\Delta t} + \frac{\sigma^2 S_i^2}{2} \frac{\mathcal{V}_{i+1}^{n+1}-2\mathcal{V}_i^{n+1}+\mathcal{V}_{i-1}^{n+1}}{\Delta S^2} + rS_i \frac{\mathcal{V}_{i+1}^{n+1}-\mathcal{V}_{i-1}^{n+1}}{\Delta S} - r\mathcal{V}_i^{n+1} & = 0 \\ \mathcal{V}_i^N = \phi(S_i) \end{cases} \tag{1.3.4}$$

$$\text{(IE)} \begin{cases} \frac{\mathcal{V}_i^{n+1}-\mathcal{V}_i^n}{\Delta t} + \frac{\sigma^2 S_i^2}{2} \frac{\mathcal{V}_{i+1}^{n}-2\mathcal{V}_i^{n}+\mathcal{V}_{i-1}^{n}}{\Delta S^2} + rS_i \frac{\mathcal{V}_{i+1}^{n}-\mathcal{V}_{i-1}^{n}}{\Delta S} - r\mathcal{V}_i^{n} & = 0 \\ \mathcal{V}_i^N = \phi(S_i) \end{cases} \tag{1.3.5}$$

$$\text{(CN)} \begin{cases} \frac{\mathcal{V}_i^{n+1}-\mathcal{V}_i^n}{\Delta t} + \frac{1}{2}\left(\frac{\sigma^2 S_i^2}{2} \frac{\mathcal{V}_{i+1}^{n}-2\mathcal{V}_i^{n}+\mathcal{V}_{i-1}^{n}}{\Delta S^2} + rS_i \frac{\mathcal{V}_{i+1}^{n}-\mathcal{V}_{i-1}^{n}}{\Delta S} - r\mathcal{V}_i^{n}\right) + \\ \quad + \frac{1}{2}\left(\frac{\sigma^2 S_i^2}{2} \frac{\mathcal{V}_{i+1}^{n+1}-2\mathcal{V}_i^{n+1}+\mathcal{V}_{i-1}^{n+1}}{\Delta S^2} + rS_i \frac{\mathcal{V}_{i+1}^{n+1}-\mathcal{V}_{i-1}^{n+1}}{\Delta S} - r\mathcal{V}_i^{n+1}\right) & = 0 \\ \mathcal{V}_i^N = \phi(S_i). \end{cases} \tag{1.3.6}$$

The variable $\mathcal{V}_i^n$ is intended to be an approximation on $\mathcal{V}(t_n, S_i)$, with $t_n = n\Delta t$. Notice that using the discretization scheme (1.3.5) (the so-called Explicit Euler scheme), the values of $(\mathcal{V}_i^n)_{0\le i\le I}$ are explicitly obtained from the values of $\left(\mathcal{V}_i^{n+1}\right)_{0\le i\le I}$. On the contrary, in the two other schemes (1.3.5) (implicit Euler scheme) or (1.3.6) (Crank-Nicolson scheme), the values of $(\mathcal{V}_i^n)_{0\le i\le I}$ are obtained from the values of $\left(\mathcal{V}_i^{n+1}\right)_{0\le i\le I}$ through the resolution of a linear system, which is more demanding from the computational viewpoint.

Various numerical methods can be used for solving the linear system; here we don't describe them in details. Let us mention that basically, there exists two classes of methods: the direct methods which are based on Gaussian elimination, and the iterative methods which consist of computing the solution as the limit of as sequence of approximations and which requires matrix-vector multiplications. The method of choice depends on the characteristics of the problem.

1.3.3 Notions of stability and consistency

In order to analyze the convergence of the three discretization schemes (1.3.4), (1.3.5), and (1.3.6), and to understand the differences between these schemes, we need to introduce two important notions. The first is the *consistency*. A numerical method is said to be consistent if, when the exact solution is plugged into the numerical scheme, the error tends to zero when the discretization parameters tend to zero. In our context, it consists of replacing $\mathcal{V}_i^n$ in (1.3.4), (1.3.5), or (1.3.6) by $\mathcal{V}(t_n, S_i)$, and check that the remaining terms tend to zero when Δt and ΔS tend to zero. By using Taylor expansions, one can check that for (1.3.4) and (1.3.5) (respectively for (1.3.6)), the reminder terms are bounded from above by $C\left(\Delta t + \Delta S^2\right)$ (resp. $C\left(\Delta t^2 + \Delta S^2\right)$), where C denotes a constant which depends on some norms of the derivatives of $\mathcal{V}$. Therefore (1.3.4) and (1.3.5) (resp. (1.3.6)) are consistent discretization schemes of order 2 in the spot variable, and of order 1 (resp. 2) in time. The second important notion is the *stability*. A numerical scheme is said to be stable if the norm of the solution to the numerical scheme is bounded from above by a constant (independent of the discretization parameters) times the norm of the data (initial condition, boundary conditions, right-hand side). This property is clearly satisfied if the numerical method is convergent, *i.e.* if the numerical approximation converges to the solution of the PDE when the discretization parameters tend to zero. A general result states that, conversely, a consistent and stable discretization scheme is indeed convergent.

The estimate of convergence is given by the estimate of consistency error. For example, the error for the EI scheme is bounded from above by $C\left(\Delta t + \Delta S^2\right)$. Notice that the constant C in these estimates depends on the solution $\mathcal{V}$: for high-order schemes, one needs more regularity on $\mathcal{V}$. For example, for some parameters, it may happen that the results obtained with the CN scheme around $t = T$ are not better than those obtained with an order one scheme (IE or EE) since the solution is not sufficiently regular in time around $t = T$.

To give a precise meaning to all these results would require to specify the norms used to measure the errors. Let us simply mention that two norms are used in practice: the stability in L^∞-norm (the supremum of absolute values of the components) is related to a discrete maximum principle; and the stability in L^2-norm (the Euclidean of the vector) is related to an energy estimate on the variational formulation.

Remark 1.23 (Discrete maximum principle). The discrete maximum principle is the counterpart at the discrete level of the maximum principle at the continuous level. It states that if the data for numerical schemes are positive, then the solution is positive. Such schemes are by construction stable in L^∞-norm. The numerical methods based on binomial to trinomial trees can be interpreted as explicit finite difference methods to solve the PDE (1.3.3), which naturally satisfy a discrete maximum principle.

Let's discuss now the properties of the three discretization schemes. We already mentioned that they are all consistent. On the other hand, it can be shown that the explicit scheme (1.3.4) is stable under an additional assumption (a so-called CFL condition, see [16]) of the form $\Delta t \leq C\Delta S^2$, where C denotes a positive constant. The other two schemes (1.3.5) and (1.3.6) are unconditionally stable. In conclusion, with the explicit scheme, the values of $(\mathcal{V}_i^n)_{0\leq i\leq I}$ can be very rapidly obtained from the values of $\left(\mathcal{V}_i^{n+1}\right)_{0\leq i\leq I}$, but the time step must be sufficiently small with respect to the spot step to guarantee stability and hence convergence. On the other hand, the implicit schemes (1.3.5) and (1.3.6) require the resolution of a linear system at each time-step, but converge without any restriction on the time-step. In terms of computational costs, the balance is generally in favor of the implicit schemes, since the CFL condition appears to be very stringent in practice.

Remark 1.24. As S is the price of the underlying and this quantity is always positive, we can define its logarithm: we define $L = \ln(S)$. This is a process adapted to the filtration induced by S. We also write $\mathcal{P}(t, L) = \mathcal{V}(t, \exp(L))$, the value of the option at time t determined by the variable L. Then, by Itô's Lemma, we obtain that P is the solution of the following PDE with boundary conditions:

$$\begin{cases} \frac{\partial \mathcal{P}}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 \mathcal{P}}{\partial L^2} + \left(r - \frac{\sigma^2}{2}\right)\frac{\partial \mathcal{P}}{\partial L} - r\mathcal{P} = 0 \\ \mathcal{P}(T, L) = \phi(\exp(L)) \end{cases}$$

This change of variables, let us get rid of the dependency in S of the advection and diffusion terms in (1.3.3). It isn't better to discretize the PDE after this change of variable, since it corresponds to take a grid refined near $S = 0$, which is useless in this case. What actually matters is to refine the grid around the singularity of $\mathcal{V}$.

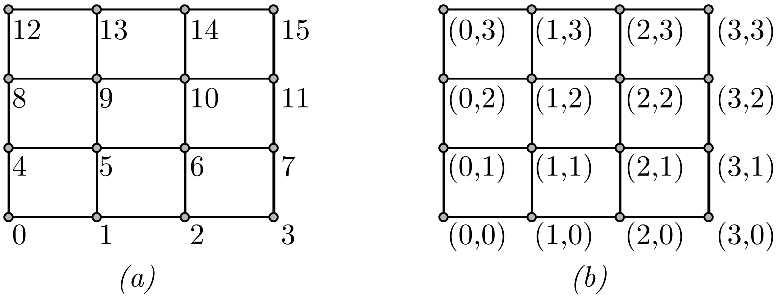

Figure 1.3.1: A rectangular domain Ω discretized by the finite difference method. In case (a) the grid points are numbered using lexicographic, while in case (b) grid points are numbered using two indices.

1.3.4 PDE ADI methods

Let's consider pricing in more complex model, such as the Heston model (see 1.1.6) or the Black-Scholes Hull-White model (see 1.1.7). Using an argument similar to the one used for the Black and Scholes model, it is possible to prove that the value of the option $\mathcal{V}_t$ is the solution of a PDE defined in the domain $\mathbb{R}^+ \times \mathbb{R}^d$ with $d = 2$ or $d \geq 2$ in more complex models.

For example, the PDE for the Heston model is

$$\frac{\partial \mathcal{V}}{\partial t} + \frac{vS^2}{2}\frac{\partial^2 \mathcal{V}}{\partial S^2} + \frac{\omega^2 v}{2}\frac{\partial^2 \mathcal{V}}{\partial v^2} + rS\frac{\partial \mathcal{V}}{\partial S} + \rho\omega S v\frac{\partial^2 \mathcal{V}}{\partial S \partial v} + k\left(\theta - v\right)\frac{\partial \mathcal{V}}{\partial v} - r\mathcal{V} = 0$$

while in the Black-Scholes Hull-White model it is

$$\frac{\partial \mathcal{V}}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 \mathcal{V}}{\partial S^2} + \frac{\omega^2}{2}\frac{\partial^2 \mathcal{V}}{\partial r^2} + rS\frac{\partial \mathcal{V}}{\partial S} + \rho\omega S\sigma\frac{\partial^2 \mathcal{V}}{\partial S \partial r} + k\left(\theta_t - r\right)\frac{\partial \mathcal{V}}{\partial r} - r\mathcal{V} = 0$$

These PDE are particularly stiff because the increase of dimension creates problems in applying implicit algorithms. In fact, an implicit algorithm would lead to a band matrix. A way to solve such problems is ADI algorithm.

The Alternating Direction Implicit (ADI) method was first used by Peaceman and Rachfors (see [40]) for solving parabolic PDEs in 1950s. Since then, it has been widely used in many applications.

ADI algorithm

The solution process of the ADI algorithm can be best explained using the model equation

$$\begin{aligned} \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \quad & (x,y,t) \in \Omega \times [0,T], \qquad (1.3.7) \\ u(x,y,t) = 0 \quad & (x,y) \in \partial\Omega,\ t > 0 \\ u(x,y,0) = u_0(x,y) \quad & (x,y) \in \Omega, \end{aligned}$$

where Ω is a unit square as shown in Figure 1.3.1, $\partial\Omega$ is the boundary of Ω. If the central finite difference scheme with five-point stencil is used to discretize the spatial derivatives in equations (1.3.7), we will have, at point (i,j), the following equation:

$$\frac{\partial u_{i,j}}{\partial t} = \frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{h^2} + \frac{u_{i,j-1} - 2u_{i,j} + u_{i,j+1}}{h^2} \tag{1.3.8}$$

$$i = 1, 2, \ldots, N_I,\ j = 1, 2, \ldots, N_J.$$

Instead of treating both terms in the right side of equation (1.3.8) implicit, we can treat one term implicit and the other term explicit, which gives rise to the following equations

$$\frac{u_{i,j}^{n+\frac{1}{2}} - u_{i,j}^{n}}{\frac{1}{2}\Delta t} = \frac{u_{i-1,j}^{n+\frac{1}{2}} - 2u_{i,j}^{n+\frac{1}{2}} + u_{i+1,j}^{n+\frac{1}{2}}}{h^2} + \frac{u_{i,j-1}^{n} - 2u_{i,j}^{n} + u_{i,j+1}^{n}}{h^2}, \tag{1.3.9}$$

and

$$\frac{u_{i,j}^{n+1} - u_{i,j}^{n-\frac{1}{2}}}{\frac{1}{2}\Delta t} = \frac{u_{i-1,j}^{n+\frac{1}{2}} - 2u_{i,j}^{n+\frac{1}{2}} + u_{i+1,j}^{n+\frac{1}{2}}}{h^2} + \frac{u_{i,j-1}^{n+1} - 2u_{i,j}^{n+1} + u_{i,j+1}^{n+1}}{h^2}. \tag{1.3.10}$$

If we assemble all equations and consider the explicit relation between equations on different vertical lines, we will have

$$\left(I_{N_I} - \frac{\Delta t}{2}A\right) u_{*,j}^{n+\frac{1}{2}} = E u_{*,j-1}^{n} + D u_{*,j}^{n} + F u_{*,j+1}^{n}, \tag{1.3.11}$$

$$j = 1, 2, \ldots, N_J,$$

where I_{N_I} is the identity matrix, and

$$\begin{aligned} u_{*,j}^{n} &= \left\{u_{1,j}^{n}, u_{2,j}^{n}, \ldots, u_{N_I,j}^{n}\right\}^T,\ j = 1, 2, \ldots, N_J \\ u_{*,0}^{n} &= u_{*,N_J+1}^{n} = 0, \\ u_{0,j}^{n} &= u_{N_I+1,j}^{n} = 0, \end{aligned}$$

$$A = \frac{1}{h^2}\left[\underbrace{\begin{matrix} -2 & 1 & & \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & -2 \end{matrix}}_{N_I}\right],$$

$$D = \frac{1}{h^2}\begin{bmatrix} h^2 - \Delta t & & & \\ & \ddots & & \\ & & \ddots & \\ & & & h^2 - \Delta t \end{bmatrix},$$

$$E = F = I_{N_I}.$$

There are actually N_J independent equation systems in equation (1.3.11) with tri-diagonal coefficient matrices.

After finishing the first step described in equation (1.3.9), we need to continue to the second step given by equation (1.3.10). If we still use the same ordering as was used for writing equation (1.3.11), then the assembled matrix equation system corresponding to equation (1.3.10) is

$$\left(I_{N_J} - \frac{\Delta t}{2} A\right) u_{i,*}^{n+1} = E u_{i-1,*}^{n+\frac{1}{2}} + D u_{i,*}^{n+\frac{1}{2}} + F u_{i+1,*}^{n+\frac{1}{2}}, \tag{1.3.12}$$

$$i = 1, 2, \ldots, N_I.$$

The matrices in (1.3.12) are the same as in (1.3.11), except than their order is $N_J \times N_J$ rather than $N_I \times N_I$. A complete ADI step can now be given as the following algorithm.

Algorithm 1.25 (ADI).

1. *Solve the N_J independent linear equation systems with tridiagonal coefficient matrices in equation (1.3.11) to get $u_{*,j}^{n+\frac{1}{2}}, j = 1, \ldots, N_J$.*
2. *Reorder the grid points.*
3. *Solve the N_I independent linear equation systems with tridiagonal coefficient matrices in equation (1.3.12) to get $u_{i,*}^{n+1}, j = 1, \ldots, N_I$.*

1.4 The Hybrid methods

The Hybrid methods were introduced in Briani et al. [11]. They are called "Hybrid" because the both use trees methods and PDE or MC methods. Here we present the two methods using quadrinomial trees presented in section 1.2.4.

1.4.1 The Hybrid Monte Carlo Method

This method is a simple and efficient way to produce MC scenarios for different models. This method is called "hybrid" because it combines trees and MC methods. First, a simple tree needs to be built: this can be done according to Appolloni et al. [4] and also [35], or as we did in Section 1.2.4. Then, using a vector of Bernoulli random variables, we move from the root through the tree, describing the scenario for the volatility or the interest rate. The values of the underlying at each time step can be easily obtained using an Euler scheme.

The generations of the volatility process and of the interest rate process behaves in a similar way: we start from the node $(0, 0)$ of the tree and according to a discrete random variable and to the node probabilities, we move to the next node and so on. Let R be a discrete random variable that can assume value A, B, C, D with probabilities p_A, p_B, p_C, p_D: sampling such a variable at each node, we get the values of the process at each time step.

We distinguish two cases for the two models.

1.4.1.1 The Heston Model

We approximate the couple (S_t, v_t) in $[0,T]$ by a discrete process $\left(\bar{S}_{k\Delta t}, \bar{v}_{k\Delta t}\right)_{k=0,\ldots,T/\Delta t}$, with $\left(\bar{S}_0, \bar{v}_0\right) = (S_0, v_0)$. For each scenario, we generate the volatility.

Let $N \sim \mathcal{N}(0,1)$ and $B \sim \mathcal{B}(0.5)$. We deduce the value of $\bar{S}_{t+\Delta t}$ by

$$\bar{S}_{t+\Delta t} = \begin{cases} \bar{S}_t \exp\Big[\big(r - \frac{\rho}{\sigma}k\theta\big)\Delta t + \big(\frac{\rho}{\sigma}k - \frac{1}{2}\big)\Big(\frac{\bar{v}_{t+\Delta t}+\bar{v}_t}{2}\Big)\Delta t + \frac{\rho}{\sigma}\left(\bar{v}_{t+\Delta t} - \bar{v}_t\right) + \sqrt{\left(1-\rho^2\right)\Delta t \bar{v}_t}N\Big] & \text{if } B = 0, \\ \bar{S}_t \exp\Big[\big(r - \frac{\rho}{\sigma}k\theta\big)\Delta t + \big(\frac{\rho}{\sigma}k - \frac{1}{2}\big)\Big(\frac{\bar{v}_{t+\Delta t}+\bar{v}_t}{2}\Big)\Delta t + \frac{\rho}{\sigma}\left(\bar{v}_{t+\Delta t} - \bar{v}_t\right) + \sqrt{\left(1-\rho^2\right)\Delta t \bar{v}_{t+\Delta t}}N\Big] & \text{if } B = 1. \end{cases}$$

According to (3.3.1), we use the normal variable N to generate the Gaussian increment of S, and the Bernoulli variable B to split the operator associated to the Heston process. This scheme (without splitting) appears in Briani et al. [11] and the splitting method appears in Alfonsi [2].

1.4.1.2 The Black-Scholes Hull-White Model

We approximate the couple (S_t, X_t) in $[0,T]$ by a discrete process $\left(\bar{S}_{k\Delta t}, \bar{X}_{k\Delta t}\right)_{k=0,\ldots,T/\Delta t}$, with $\left(\bar{S}_0, \bar{X}_0\right) = (S_0, 0)$, and we deduce the interest rate by $\bar{r}_t = \omega \bar{X}_t + \beta(t)$. Let $N \sim \mathcal{N}(0,1)$. We deduce the value of $\bar{S}_{t+\Delta t}$ by

$$\bar{S}_{t+\Delta t} = \bar{S}_t \exp\left[\left(\frac{\bar{r}_{t\Delta t} + \bar{r}_t}{2} - \frac{\sigma^2}{2}\right)\Delta t + \sigma\left(\left(\bar{X}_{t+\Delta t} + \bar{X}_t\left(k\Delta t - 1\right)\right)\rho + \sqrt{\Delta t}\bar{\rho}N\right)\right].$$

1.4.2 The Hybrid PDE Method

The Hybrid PDE Method is a new approach based both on tree and finite difference methods. It can be used to perform pricing in several model such as the Heston model or the Black-Scholes Hull-White model.

Roughly speaking, our method approximates the CIR type volatility process (or the Ornstein-Uhlenbeck process) through a tree approach, which turns out to be very robust and reliable. And at each step, we make use of a suitable transformation of the asset price process allowing one to take care of a new diffusion process with null correlation w.r.t. the volatility process. Then, by taking into account the conditional behavior with respect to the evolution of the volatility process, we consider a finite difference method to deal with the evolution of the (transformed) underlying asset price process. We stress that jumps may be allowed in the dynamics for the underlying asset prices process, but this shall be the subject of a further work.

1.4.2.1 The Heston Model

Starting from the model in (3.3.1), we call $\bar{\rho} = \sqrt{1-\rho^2}$ and we write $Z_t^S = \rho Z_t^V + \bar{\rho}\bar{Z}_t^S$, where $\bar{Z}^S$ is a Brownian motion uncorrelated with Z^V. Then,

$$\begin{cases} dS_t = rS_t dt + \sqrt{v_t}S_t\left(\rho dZ_t^V + \bar{\rho} d\bar{Z}_t^S\right) & v_0 = \bar{v}_0, \\ dv_t = k\left(\theta - v_t\right)dt + \omega\sqrt{v_t}dZ_t^V & S_0 = \bar{S}_0, \end{cases} \quad d\left\langle \bar{Z}_t^S, Z_t^V \right\rangle = 0,$$

we define the process

$$Y_t^E = \ln(S_t) - \frac{\rho}{\omega} v_t,\ Y_0^E = \ln(S_0) - \frac{\rho}{\omega} v_0,$$
$$S_t = \exp\left(Y_t^E + \frac{\rho}{\omega} V_t\right). \tag{1.4.1}$$

Then

$$dY_t^E = \left(r - \frac{v_t}{2} - \frac{\rho}{\omega} k(\theta - v_t)\right) dt + \bar{\rho}\sqrt{v_t} d\bar{Z}_t^S.$$

This process Y_t^E is important because it's a process uncorrelated with V_t, and we introduced it as in [10]. We are going to use it to define a PDE to be solved along the tree.

1.4.2.2 The Black-Scholes Hull-White Model

Starting from the model (3.3.2), we call $\bar{\rho} = \sqrt{1-\rho^2}$ and we write $Z_t^S = \rho Z_t^r + \bar{\rho}\bar{Z}_t^S$, where $\bar{Z}^S$ is a Brownian motion uncorrelated with Z^r. Then,

$$\begin{cases} dS_t = r_t S_t dt + \sigma S_t \left(\rho dZ_t^r + \bar{\rho} d\bar{Z}_t^S\right) & S_0 = \bar{S}_0, \\ dX_t = -kX_t dt + dZ_t^r & X_0 = 0, \quad d\left\langle \bar{Z}_t^S, Z_t^r \right\rangle = 0, \\ r_t = \omega X_t + \beta(t), \end{cases}$$

we define the process

$$Y_t^U = \ln(S_t) - \rho\sigma X_t,\ Y_0^U = \ln(S_0),$$
$$S_t = \exp\left(Y_t^U + \rho\sigma X_t\right). \tag{1.4.2}$$

Then

$$dY_t^U = \left(r_t - \frac{\sigma^2}{2} + \sigma\rho k X_t\right) dt + \sigma\bar{\rho} d\bar{Z}_t^S.$$

This process Y_t^U is important because it's a process uncorrelated with X_t, and we introduced it as in [10]. We are going to use it to define a PDE to be solved along the tree.

1.4.2.3 Algorithm structure

The structures for this algorithm consist in a tree and a PDE solver. As described in Briani et al. [10], [11], we use a tree to diffuse the volatility (or the interest rate) along the life of the product, and we solve backward a 1D PDE freezing at each node of the tree the volatility (or the interest rate). The tree is built according to Section 1.2.4 (quadrinomial tree, matching the first three moments of the process), and the PDE is solved with a finite difference approach. We have to solve the PDE between event times, and at each event time we apply the changes to the states to reproduce the effects of the events.

We remark that we solve the PDEs doing a single time step that requires only a linear complexity because we have to solve a linear system with tridiagonal matrix. The computational cost is low as observed in [10] and [11]. We observe that X_t and V_t processes are mean reverting. Thanks to the way the trees are built, there are many nodes in the trees that cannot be visited by the approximating Markov chain. Therefore their probability $p_{n,j}$ to be visited is worth 0 and they have no impact on the values at the root of the tree. There is no reason to do any operation for those nodes. So, to save time, we do the standard step (solve backward the four PDEs and mix up the vectors according to the transition probabilities) only for those nodes having $p_{n,j} > 0$. This curtailing technique reduces the computational time, and the convergence of the method is preserved. A similar approach is used in [3].

Chapter 2

Variable Annuities

2.1 Introduction

In this Chapter we describe the main features of Variable Annuities. These products will be our object of interest.

2.1.1 What is a Variable Annuity?

The term Variable Annuity (hereinafter, we will abbreviate it with *VA*) is used to refer to a wide range of life insurance products, whose benefits can be protected against investment and mortality risks by selecting one or more guarantees out of a broad set of possible arrangements. Despite the unique characteristics can change, there are some common to all of them: a VA is a long-term, tax-deferred investment, designed for obtaining a post-retirement income.

Formally, a VA is a contract between a policy holder (hereinafter, we will abbreviate it with *PH*) and a insurance company, under which the insurer agrees to make periodic payments to the PH beginning either immediately or at some future date. The holder buys a VA contract by making either a single purchase payment (lump sum) or a series of purchase payments.

Variable Annuities were introduced in the 1970s in the United States and in the first years of 2000 they became popular also in Europe (expecially Germany, UK, and France) and known in Italy (see [8]). The cause of the success of these policies is that they offer a range of investment options. The value of the investment will change depending on the performance of the investment options chosen. The investment options for a VA are typically mutual funds that invest in stocks, bonds, money market instruments, or some combination of the three.

Variable Annuities are designed to be long-term investments, to meet retirement and other long-range goals. Variable annuities also involve investment risks, just as mutual funds do.

2.1.2 Differences between V.A. and other instruments

Although Variable Annuities are typically invested in mutual funds, they differ from mutual funds in several important ways.

First, VAs may let the PH receive periodic payments for the rest of his life (or the life of any other person designated). This feature offers protection against the possibility that, after retire, the PH will outlive his assets.

Second, VAs may have a death benefit. If the PH dies before the insurer has started making payments, the beneficiary is guaranteed to receive a specified amount—typically at least the amount of the purchase payments. The policy beneficiary will also get a benefit from this feature if, at the time of the death of the holder, his account value is less than the guaranteed amount.

Third, VAs are tax-deferred. That means the PH pays no taxes on the income and investment gains from his annuity until he withdraws his money. He may also transfer his money from one investment option to another within a VA without paying tax at the time of the transfer. When the PH takes his money out of the VA, however, he will be taxed on the earnings at ordinary income tax rates (for example, according to Italian Law, the financial return, equal to the difference between the amount paid and the premiums paid, shall be subject to the application of a substitute tax on income, at the time of payment of the benefit, according to what provided by the D.L. August 13, 2011 n. 138, converted into Law 148 of 14 September 2011). For other information see [47].

2.2 How Variable Annuities work

A VA has usually two phases: an accumulation phase and a payout phase.

2.2.1 The accumulation phase

During the accumulation phase, the PH makes purchase payments, which he can allocate to a number of investment options. For example, he could designate 40% of his purchase payments to a bond fund, 40% to a U.S. stock fund, and 20% to an international stock fund. The money he has allocated to each mutual fund investment option will increase or decrease over time, depending on the fund's performance. In addition, VA often allow the PH to allocate part of his purchase payments to a fixed account. A fixed account, unlike a mutual fund, pays a fixed rate of interest. The insurance company may reset this interest rate periodically, but it will usually provide a guaranteed minimum (e.g., 3% per year).

During the accumulation phase, the PH can typically transfer his money from one investment option to another without paying tax on his investment income and his accumulation phase, he may have to pay "surrender charges," which are discussed below.

2.2.2 The payout phase

At the beginning of the payout phase, the PH may receive his purchase payments plus investment income and gains (if any) as a lump-sum payment, or he may choose to receive them as a stream of payments at regular intervals (generally monthly). If he chooses to receive a stream of payments, he may have a number of choices of how long the payments will last. Under most annuity contracts, he can choose to have his annuity payments last for a period that he sets (such as 20 years) or for an indefinite period (such as his lifetime or the lifetime of him and his spouse or other beneficiary). During the payout phase, holder's annuity contract may permit him to choose between receiving payments that are fixed in amount or payments that vary based on the performance of the mutual fund investment options.

The amount of each periodic payment will depend, in part, on the time period that the PH selects for receiving payments. Some annuities do not allow to withdraw money from the account once the holder has started receiving regular annuity payments. In addition, some annuity contracts are structured as immediate annuities, which means that there is no accumulation phase and the holder will start receiving annuity payments right after he purchases the annuity.

2.3 The Death Benefit and other features

A common feature of VA is the death benefit. If the holder dies, a person he selects as a beneficiary (such as his spouse or child) will receive the greater of: (i) all the money in the holder's account, or (ii) some guaranteed minimum (such as all purchase payments minus prior withdrawals). This second case is known as GMDB (Guaranteed Minimum Death Benefit). Some VA allow the PH to choose a "stepped-up" death benefit. Under this feature, the guaranteed minimum death benefit may be based on a greater amount than purchase payments minus withdrawals. For example, the guaranteed minimum might be the account value as of a specified date, which may be greater than purchase payments minus withdrawals if the underlying investment options have performed well. The purpose of a stepped-up death benefit is to "lock in" the investment performance and prevent a later decline in the value of the account from eroding the amount that the holder expects to leave to his heirs. This feature carries a charge, however, which will reduce account value.

Variable annuities sometimes offer other optional features, which also have extra charges. The PH pays for each benefit provided by his VA. These charges are usually independent and tied to the relative benefit.

The following are the most common optional features (for more details see [5]:

- GMAB (*Guaranteed Minimum Accumulation Benefit*): this features provides to guarantees the minimum amount received by the annuitant after the accumulation period, protecting the value of the annuity and the annuitant from market fluctuations. The GMAB will be used only if the market value of the annuity is below the minimum guaranteed value.

- GMIB (*Guaranteed Minimum Income Benefit*): a common feature that guarantees a particular minimum level of annuity payments, even if the holder does not have enough money in his account (perhaps because of investment losses) to support that level of payments. When the annuity has been annuitized (start the paying of the annuities), this specific option guarantees that the annuitant will receive a minimum value's worth of payments.

- GMWB (*Guaranteed Minimum Withdrawal Benefit*): this specific option gives annuitants the ability to protect their retirement investments against downside market risk by allowing the annuitant the right to withdraw a maximum percentage of their entire investment each year until the initial investment amount has been recouped. The GMWB is the real novelty of Variable Annuities in respect of traditional life insurance contracts.

These three options are also called GMxBs guarantees (namely, Guaranteed Minimum Benefits of type 'x'). Another VA adds to the previous three:

- GLWB (*Guaranteed Lifelong Withdrawal Benefit*): this option is similar to GMWB, but this policy has no fixed maturity. The annuitants has the right to perform periodic withdrawals, with a minimal guaranteed withdrawal, for all his life. A priori, there are no limits on guaranteed withdrawals, and on the total guaranteed withdrawal. Usually, a death benefit is always included.

These contracts may include other features such as long-term care insurance (LTC), which pays for home health care or nursing home care if the PH becomes seriously ill. The most common forms of guarantees associated with the growth of the benefit base are:

- Roll-ups: this is the simplest form of return guarantee. A roll-up provides guaranteed appreciation of the benefits base at a specific interest rate. The guarantee may accrue on a simple or compound interest basis. A 0 per cent roll-up is the same as a return-of-principal guarantee.

- Ratchets: also called a "high watermark." With a ratchet, the benefits base is set equal to the highest of all values of the underlying funds throughout the accumulation phase, evaluated at a pre-defined time interval (e.g. annually). At various frequencies the existing benefits base is compared to the account value, and if the account value is higher, the benefits base is "ratcheted" up to the new level.

- Resets: resets are triggered at the discretion of the PH. They involve a comparison of the current account value to the original account value, and the benefits base is reset to the higher level. Other policy provisions such as a waiting period may be reset as well.

- Some VA offer guaranteed appreciation of the benefits base that combines one or more of the above forms of guarantees. For example, a common combination guarantee is the maximum of a roll-up and a ratchet.

Example 2.1. A man owns a VA that offers a death benefit equal to the greater of account value or total purchase payments minus withdrawals. He has made purchase payments totaling € 50000. In addition, he has withdrawn € 5000 from his account. Because of these withdrawals and investment losses, his account value is currently € 40000. If he dies, his designated beneficiary will receive € 45000 (the € 50000 in purchase payments he put in, minus € 5000 in withdrawals).

2.4 Variable Annuities charges

A person who invests in a VA, pays several charges. These charges will reduce the value of his account and the return on his investment. Often, they will include the following.

Guarantee charges

These charges are used by the insurance company to cover the guarantees of the policy. They are a fixed percentage of the account value, and are usually withdraws continuously. These charges are active for the whole product life and are fixed at the beginning; finding the fair value of these fees, consists in pricing the product.

Example 2.2. The guarantee charges of a VA are 2%. If the initial account value is € 100, and the linked found increases of 5%, then the final value of the account value is € $100 \cdot 1.05 \cdot (1 - 0.02) =$ € 102.9.

Surrender charges

If the PH withdraws money from a VA within a certain period after a purchase payment (typically within six to eight years, but sometimes as long as ten years), the insurance company usually will assess a "surrender" charge, which is a type of sales charge. This charge is used to pay a commission to his financial professional for selling the VA. Generally, the surrender charge is a percentage of the amount withdrawn, and declines gradually over a period of several years, known as the "surrender period." For example, a 7% charge might apply in the first year after a purchase payment, 6% in the second year, 5% in the third year, and so on until the eighth year, when the surrender charge no longer applies.

Often, contracts will allow the holder to withdraw part of his account value each year (10% or 15% of his account value, for example) without paying a surrender charge.

Example 2.3. A man purchases a VA contract with a € 10000 purchase payment. The contract has a schedule of surrender charges, beginning with a 7% charge in the first year, and declining by 1% each year. In addition, the holder is allowed to withdraw 10% of his contract value each year free of surrender charges. In the first year, he decides to withdraw € 5000, or one-half of his contract value of € 10000 (assuming that his contract value has not increased or decreased because of investment performance). In this case, he could withdraw € 1000 (10% of contract value) free of surrender charges, but he would pay surrender charge of 7%, or € 280, on the other € 4000 withdrawn.

Mortality and expense risk charge

This charge is equal to a certain percentage of the account value, typically in the range of 1.25% per year. This charge compensates the insurance company for insurance risks it assumes under the annuity contract. Profit from the mortality and expense risk charge is sometimes used to pay the insurer's costs of selling the VA, such as a commission paid to the financial professional for selling the VA.

Example 2.4. A VA has a mortality and expense risk charge at an annual rate of 1.25% of account value. The average account value during the year is € 20000 so the holder will pay € 250 in mortality and expense risk charges that year.

Administrative fees

The insurer may deduct charges to cover record-keeping and other administrative expenses. This may be charged as a flat account maintenance fee (perhaps € 25 or € 30 per year) or as a percentage of his account value (typically in the range of 0.15% per year).

Example 2.5. A VA charges administrative fees at an annual rate of 0.15% of account value. The average account value during the year is € 50000. The PH will pay € 75 in administrative fees.

Fees and charges for other features

Special features offered by some VAs, such as a stepped-up death benefit, a guaranteed minimum income benefit, or long-term care insurance, often carry additional fees and charges.

Other charges, such as initial sales loads, or fees for transferring part of the account from one investment option to another, may also apply.

2.5 Bonus credits

Some insurance companies offer VA contracts with "bonus credit" features. These contracts promise to add a bonus to the contract value based on a specified percentage (typically ranging from 1% to 5%) of purchase payments.

Example 2.6. A man purchases a VA contract that offers a bonus credit of 3% on each purchase payment. He makes a purchase payment of € 20000. The insurance company issuing the contract adds a bonus of € 600 to the account.

Variable annuities with bonus credits may carry a downside; higher expenses can outweigh the benefit of the bonus credit offered.

Frequently, insurers will charge the holder for bonus credits in one or more of the following ways:

Higher surrender charges

Surrender charges may be higher for a VA that pays a bonus credit than for a similar contract with no bonus credit.

Longer surrender periods

The PH purchases payments may be subject to surrender charges for a longer period than they would be under a similar contract with no bonus credit.

Higher mortality and expense risk charges and other charges

Higher annual mortality and expense risk charges may be deducted for a VA that pays the holder a bonus credit. Although the difference may seem small, over time it can add up. In addition, some contracts may impose a separate fee specifically to pay for the bonus credit.

2.6 A few real examples

2.6.1 Italian market

We present in this section a few examples of VA available in the Italian market (see [49]).

These financial products are entering slowly in the Italian market through some companies such as AXA, with the product of the *Accumulator line*, and Assicurazioni Generali, with *Generali Active*.

The selection *Accumulator* of AXA offers a minimum return of 25% in 10 years or 13% in five years in return for payment of an initial prize fund of at least € 2500. If the PH keeps the investment after the expiry of the contract, it turns into a traditional unit-linked (so, by that date, there is no guarantee of capital) in the event of early redemption in the first four years instead the holder pays a penalty of 1%. So, this policy offers a GMAB optional feature.

Generali Active Savings, instead, consists of a scheduled program of recurring single premiums, by installments, of the annual minimum of € 600. This product provides at the end of the accumulation phase (minimum 15, maximum 25 years) a capital guaranteed total capitalization for each installment paid, the premium invested at 2% per year for the period between the commencement of single installment and the expiration of plan premiums, or the date of death, in case of death before the expiry of the payment of premiums. So, that policy offers a GMAB optional feature as AXA one.

2.6.2 Japanese market

An emblematic example of VA is the Yen VA++ sold by AXA in the Japanese market. It's a VA with GMIB and GMDB, with both roll-up (up to 10 times) and yearly ratchet.

There exist two versions of this product: a 25 years TC (time certain) and a whole life (WL). The premium is paid as a unique solution (single premium). The PH can lapse anytime with no surrender charges.

The ratchets take place yearly, every anniversary date throughout the contract. The initial income benefit of the ratchet is equal to the initial gross premium, and each year during the deferral period, it is updated as follows

$$IB\_ratchet_t = \max\left(IB\_ratchet_{t-1}, AV_t\right) \tag{2.6.1}$$

For the annuity payment period the benefit is updated in two different ways for the two product types. For the TC product we continue to use formula (2.6.1), while for the WL product we have

$$IB\_ratchet_t = \begin{cases} \max\left(IB\_ratchet_{t-1}, AV_t + \sum_{s<t} Payout_s\right) & \text{if } AV_t > 0 \\ IB\_ratchet_{t-1} & \text{if } AV_t = 0 \end{cases}.$$

The roll up benefit is calculated each year at the contract anniversaries, up to 10 years and only during the deferral period. The roll up rate is equal to 1.5% for the TC, and 2.5% for the WL. The initial roll up benefit is equal to the initial gross premium, and at each anniversary we have

$$IB\_rullup_t = Initial\_premium \cdot (1 + t \cdot rollup\_rate) .$$

The GMIB is calculated as the max between the ratchet benefit and the roll-up benefit

$$IB_t = \max (IB\_ratchet_t, IB\_rollup_t)$$

At each anniversary, if the PH is still alive, he receives a sum equal to

$$Payout_t = IB_t \times IBRate$$

where the $IBRate$ is worth 3% for the WL, and $1/ (contrac\_period - deferral\_period)$ for the TC.

The last of the deferral period is chosen by the PH at the beginning of the contract with some duration limits.

The death benefit is calculated as

$$DB_t = \max \left(AV_t, IB_t - \sum_{s<t} Payout_s \right) .$$

There are several charges: at the entry a 5% is calculated on the single premium paid. Everyday throughout the contract life, mortality, expense, and fund management charges are calculated as a percentage of the AV, and withdrawn.

In figure (2.6.1) there is an example of the development of a YEN VA++ WL contract.

2.7 Structure

Unlike the with-profit or participating business, reference funds backing VA are not required to replicate the guarantees selected by the PH, as these are hedged by specific assets. Therefore, reference fund managers have more flexibility in catching investment opportunities.

The peculiarity of VA, and their distinctive compared to European products, is that the guarantee is external to the fund and does not affect the asset allocation. For this reason, the VA products, regardless of factors specific to individual products and/or individual guarantees, may be defined multidimensional guaranteed products as opposed to guaranteed Italians products, who can be considered one-dimensional.

Guarantees of performance associated with unit-linked products of the first generation are normally made through the annexation of a guaranteed fund to the family of funds available for the product concerned. The financial guarantee (of performance, capital etc ...) is then "included" in specific funds with assets specifically chosen to cope with the guarantee offered by the funds themselves. A typical unit-linked (guaranteed) Italian could be represented by a one-dimensional scheme. In the case, for example of a product with 5 funds 2 which guaranteed it would be:

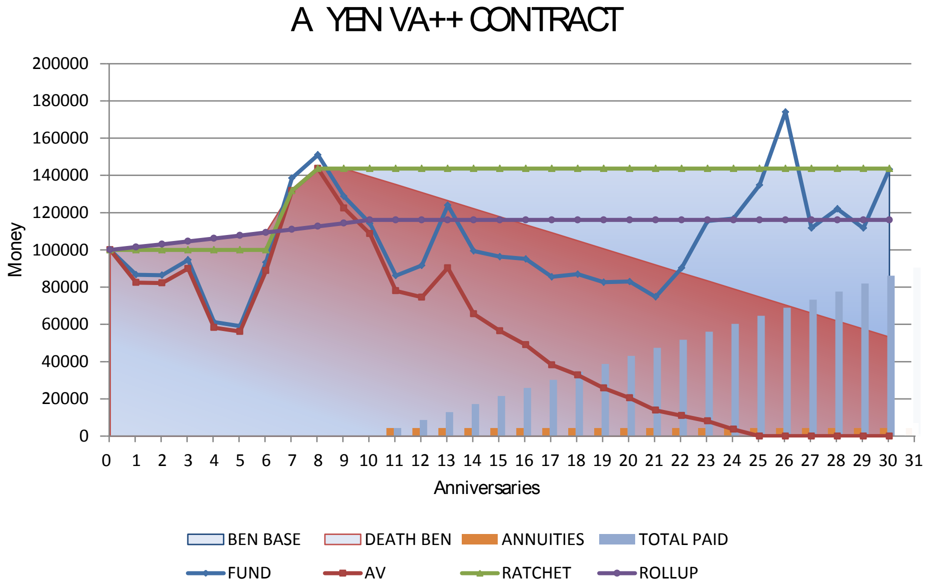

Figure 2.6.1: An example of the development of a VA contract.

Fund / Warranty	Percentage
Not guaranteed Fund #1	0%
Not guaranteed Fund #2	25%
Not guaranteed Fund #3	15%
Guaranteed Fund #1	40%
Guaranteed Fund #2	20%

The return guarantees, associated with VAs, are real options associated contract, at the option of the contracting party, in each of the funds available for the product. VA can be represented by a double entry table. If, for example, it consists of 3 funds and 3 funds guarantees, it would be:

Fund	Percentage	Warranty #1	Warranty #2	Warranty #3
Fund #1	30%			
Fund #2	20%			
Fund #3	50%			

At each fund (not guaranteed) and at each benefit is associated a cost: the total cost is theoretically equal, for each fund selected by each insurance position, to the sum of the cost of the fund and of the guarantees associated with it.

2.8 Variable annuity risk management

There are generally three buckets of risk that exist with almost all life insurance products, VA included. These are:

- insurance risk
- market risk
- behavioral or utilization risk.

For VA specifically:

- Longevity risk is the primary insurance risk due to the nature of the income guarantees that are offered; some mortality risk exists due to the nature of the death benefit guarantees that are offered.
- Equity risk and interest rate risk are the primary market risks due to 1) the underlying equity and fixed-income investments that drive the PH's account value performance and 2) the long-term nature of the income guarantees. In addition, some credit risk also is present in the fixed-income investments.
- Persistence risk and benefit utilization risk are the primary behavioral or utilization risks due to the nature of the product structure which generally has the insurer receiving revenue over time and insurance claims being paid well into the future.

Primary "Lines of Defense"

Insurers use a number of lines of defense to manage the above buckets of risks. These are:

- product design and prudence in assumptions
- risk pooling ("law of large numbers")
- natural hedges and a diverse balance sheet
- asset liability management (ALM) and reinsurance
- stress scenario analysis for single and combined shocks and the appropriate provision and management of economic risk capital.

These lines of defense are employed to varying degrees in an insurer's risk management strategy, depending on the nature of the risk and the availability and effectiveness of each method. Not all of the lines of defense listed are used with all risks or types of insurance. For instance, reinsurance is generally not used as a primary risk management strategy due to the current limited availability of reinsurance for VA guaranteed benefits.

2.9 Pricing

As can be easily understood, many risks affect the Variable Annuity business; mortality/longevity, financial, policyholder behavior risks are the most relevant. It is not fully known how these risks interact, especially when referring to the post-retirement phase.

Among the different phases of the risk management process, the pricing and hedging of guarantees, i.e. of the relevant financial options, should be a major concern for the insurer when designing the contract. Appropriate evaluation techniques need to be developed in order to account on one hand for the interaction between financial and mortality/longevity issues, on the other for the PH behavior.

Chapter 3

Pricing and Hedging GLWB in the Heston and in the Black-Scholes with Stochastic Interest Rate Models

3.1 Introduction

This Chapter presents the results about the research paper [23]. We consider a Guaranteed Lifelong Withdrawal Benefit (GLWB) annuity. We restrict our attention to a simplified form of a GLWB which is initiated by making a lump sum payment to an insurance company. This lump sum is then invested in risky assets, usually a mutual fund. The *benefit base*, or guarantee account balance, is initially set to the amount of the lump sum payment. The holder of the policy (hereinafter, we will abbreviate it with *PH*) is entitled to withdraw a fixed fraction of the benefit base for life, even if the actual investment in the risky asset (*account value*) declines to zero. Upon the death of the PH, his (her) heirs receives the remaining amount in the risky asset account. Typically, these contracts have ratchet provisions (step-ups), that periodically increase the benefit base if the risky asset investment has increased to a value larger than the guarantee account value, and roll up provisions, that periodically increase the benefit base according to a deterministic function. In addition, the benefit base may also be increased if the PH doesn't withdraw in a given year (bonus). Finally, the PH may withdraw more than the contractually specified amount, including complete surrender of the contract, upon payment of a penalty. Complete surrender here means that the PH withdraws the entire amount remaining in the investment account, and the contract terminates. In most cases, this penalty for full or partial surrender declines to zero after five to seven years.

The hedging costs for this guarantee are offset by deducting a proportional fee from the risky asset account. From an insurance point of view, these products are treated as financial ones: the products are hedged as if they were pure financial products, and the mortality risk is hedged using the law of large numbers. Therefore, it is very important for insurance companies to be able to price quickly these products. Moreover these products have long maturities that could last almost 60 years. The Black-Scholes model, with its constant interest rate and volatility, seems to be unsuitable for these products: that's why we present our pricing methods in two frameworks, modeling stochastic volatility (Heston model [26]) and stochastic interest rate (Hull-White model [28]) .

There have been several recent articles on pricing GLWBs. In particular, we would remember the Forsyth and Vetzal's work [21]: they used a PDE approach in a multi regimes model to price GLWBs contracts. This approach proved to be very fast and accurate, and we used it as a reference for our work. Concerning the use of stochastic volatility, Kling et al. [30] used a Monte Carlo approach to price products. We have made reference also to Bacinello et al. [5]: variable annuities (including GLWBs) are priced using a Monte Carlo approach. The PH's behavior is assumed to be semi-Static, i.e. the holder withdraws at the contract rate or surrenders the contract.

In this Chapter, we price GLWBs guarantees, and we find the no-arbitrage fee, in the Heston model and the Black-Scholes with stochastic interest rate model (*BS HW model*). First, we treat a Static withdrawal strategy: the PH withdraws at the contract rate. Then, taking the point of view of the worst case for the hedger, we price the guarantees assuming that the PH follows an optimal withdrawal strategy. We also used these methods to calculate the Greeks for hedging and Risk Management. Moreover we performed a mortality shock useful in Risk Management framework. For this purpose we present four numerical methods: a hybrid tree-finite difference method and a Hybrid Monte Carlo method (both introduced by Briani et al. [10]), an ADI finite difference scheme (Haentjens and Hout [25]), and a Standard Monte Carlo method with Longstaff-Schwartz least squares regression (Longstaff and Schwartz [33]).

We use the term *no-arbitrage fee* in the sense that this is the fee which is required to maintain a replicating portfolio. A description of the replicating portfolio for these types of guarantees is given in Chen et al. [14] and Belanger et al. [6].

The main results of this Chapter are the following ones:

- We formulate the determination of the no-arbitrage fee (i.e. the cost of maintaining a replicating hedging portfolio) in the Heston model and in the BS HW model using different pricing methods;

- We present the effects of stochastic volatility and stochastic interest rate on pricing and calculation of Greeks, and the sensitivity of the GLWB fee to various modeling parameters;

- We use different numerical methods to price the GLWB contract;

- We present numerical examples which show the convergence of these methods.

The Chapter is organized as follows: in Section 2, we describe the main features of the contract such as mortality, withdrawals, and ratchets. In Section 3, we provide a brief review of the stochastic models used afterward. In Section 4, we present the numerical methods, and how to implement them to solve the GLWB contract pricing problem. In Section 5 we perform numerical tests in order to show their behavior and we study the sensitivity of the no-arbitrage fee to economic, contractual and longevity assumptions. Finally, in Section 6, we present the conclusions.

3.2 The GLWB contract

In the following, we will refer to the contract described in the paper of Forsyth [21], with some variations useful to compare our results with other works. We make a brief summary of the main features of the contract.

3.2.1 Mortality

We price the products in a risk-neutral measure, therefore in the following we assume that mortality risk is diversifiable (Milevsky and Salisbury, [34]). When this assumption is not justified, then the risk-neutral value of the contract can be adjusted using an actuarial premium principle (Gaillardetz and Lakhmiri, [22]). Hereinafter, the time variable will be denoted by the letter t, and we assume that the contract starts at $t = 0$.

First we suppose that no policy holder can live longer than a given age. This age will be denoted by τ (usually $\tau = 122$). The age of the PH at the beginning will be denoted by a_0 (usually $a_0 = 65$). So, the maturity of the contract is $T = \tau - a_0$ (usually $T = 57$): when the time variable t reaches T all PHs are dead, and the contract is worth zero.

The effects of the mortality on the contract are described using two functions:

- $\mathcal{M} : [0, T] \to \mathbb{R}$ is the probability density that describes the random variable M associated to the death year of the PH. The fraction of the original owners who die in $[t, t + dt]$ is equal to $\mathcal{M}(t)\,dt$.
- $\mathcal{R} : [0, T] \to \mathbb{R}$ is the fraction of the original owners who are still alive at time t

$$\mathcal{R}(t) = 1 - \int_0^t \mathcal{M}(s)\,ds.$$

We remark that $\mathcal{R}(0) = 1$ and $\mathcal{R}(T) = 0$. For seek of simplicity, we assume $\mathcal{M}$ to be constant between contract anniversaries: if $t \in [k, k+1[$, $k \in \mathbb{N}$ then $\mathcal{M}(t) = \mathcal{M}(k)$.

3.2.2 Contract state parameters

At time $t = 0$ the policy holder pays with lump sum the gross premium GP to the insurance company. This may be reduced by some initial fees, giving a net premium P. The premium P is invested in a fund, whose price is denoted by the variable S_t. The state parameters of the contract are:

- Account value: A_t, $A_0 = P$.
- Base benefit: B_t, $B_0 = GP$.

Both these two variables are initially set equal to the gross premium or to the premium.

We suppose that the acquisition charges are equal to $GP - P$ aren't used for hedging purposes, but only to cover entry costs for management control. We suppose that there is a set of discrete times t_i, which we term *event times*. At these times, withdrawals, ratchets, and bonuses may occur. Normally, event times are annually or quarterly. We first consider the evolution of the value of the guarantee excluding these event times t_i.

3.2.3 Evolution of the contract between event times.

Let $t \in]t_i, t_{i+1}[\subseteq [0, T]$. As we said before, S_t denotes the underlying fund driving the account value. The dynamics of S_t will be described in the next Section. The account value A_t follows the same dynamics of S_t with the exception of the fact that some fees may be subtracted continuously:

$$dA_t = \frac{A_t}{S_t} dS_t - \alpha_{tot} A_t dt. \tag{3.2.1}$$

We suppose that the total annual fees are charged to the PH and withdrawn continuously from the investment account A_t. These fees include the mutual fund management fees α_m and the fee charged to fund the guarantee (also known as the rider) α_g, so that

$$\alpha_{tot} = \alpha_m + \alpha_g.$$

The only portion used by the insurance company to hedge the contract is that coming from α_g: the other part of the fees has to be considered as a outgoing money flow as PH's withdrawals are.

Continuously withdrawn fees are typical of the contract described by Forsyth. Fees may also be withdrawn at the end of each policy year t_i: this is what Kling et al. do in [30]. In this second case

$$dA_t = \frac{A_t}{S_t} dS_t. \tag{3.2.2}$$

When the PH dies, a death benefit usually equal to A_t, is paid out to the heirs of the PH. According to formulation of the contract, this death benefit may be paid immediately or at the upcoming event time. If it is paid immediately, the contract stops immediately and the account value and the benefit base becomes equal to zero; otherwise the contract goes on up to the next event time as if nothing has happened and then it ends.

3.2.4 Event times

An event time is a sequence of operations under the contract, which occur at fixed dates, usually at each anniversary of the signing of the contract. The times these events take place are denoted by $t_i = \Delta t \cdot i$ and usually $\Delta t = 1$. Let's define $I = T/\Delta t$; then, i runs in $\{0, \dots I\}$.

When an event time occurs, we assume that the following events happen in this order:

1. Withdrawal of the fees by the insurance company (if it is not time continuous);
2. If the PH died, payment of the death benefits;
3. If the PH is still alive, he (she) is entitled to withdraw a certain amount of money;
4. If provided by the contract, a ratchet may increase the benefit base B_t.

We denote with $\left(A_{t_i}^-, B_{t_i}^-, t_i\right)$ the state variables just before an event time that occurs at time t_i and with $\left(A_{t_i}^{k+}, B_{t_i}^{k+}, t_i\right)$ the state variables just after the update due to the i-th point of the previous numbered list.

3.2.4.1 Fees

Fees may be withdrawn continuously by the account value, as supposed in Forsyth and Vetzal in [21]. In this case, between two event times, the account value changes as prescribed by equation (4.2.2), and nothing special happens:

$$\left(A_{t_i}^{1+}, B_{t_i}^{1+}, t_i\right) = \left(A_{t_i}^{-}, B_{t_i}^{-}, t_i\right).$$

Otherwise, fees may be withdrawn at the end of the period, as supposed in Kling et al. [30]. In this case, between two event times, the account value changes as prescribed by equation (3.2.2), and at the event time the state parameters become

$$\left(A_{t_i}^{1+}, B_{t_i}^{1+}, t_i\right) = \left(A_{t_i}^{-} e^{-\alpha_{tot}\Delta t}, B_{t_i}^{-}, t_i\right).$$

It is important to be able to deduce the management fees F_t^{man} withdrawn by the account value because they are not used to hedge the contract and therefore they have to be considered as an outgoing money flow. If these fees are withdrawn continuously, we can calculate them observing that their dynamics between two event times is

$$dF_t^{man} = \alpha_m A_t dt + r_t dt.$$

This ODE has the following solution

$$F_t^{man} = \int_0^t e^{\int_s^t r_u du} \alpha_m A_s ds.$$

and can be used in a Monte Carlo approach.

If the fees are withdrawn at the end of the period, we can calculate management fees as a fraction of the total fees withdrawn:

$$F_{t_i}^{tot} = F_{t_{i-1}}^{tot} + A_{t_i}^{0}\left(1 - e^{-\alpha_{tot}\Delta t}\right),$$

$$F_{t_i}^{man} = F_{t_{i-1}}^{man} + \frac{\alpha_{man}}{\alpha_{tot}}\left(F_{t_i}^{tot} - F_{t_{i-1}}^{tot}\right).$$

3.2.4.2 Death Benefit

If the PH died at a given time $\bar{t} \in]t_{i-1}, t_i[$, his (her) heirs will obtain a death benefit that is usually equal to the account value. If the contract provides that the death benefit is paid immediately, then the death benefit $DB_{\bar{t}}$ is paid in $\bar{t}$ and is equal to $A_{\bar{t}}$. Otherwise, if the DB is paid at the next event time, then $DB_{t_i} = A_{t_i}^{1+}$ and the contract is concluded (after the DB payment it's worthless):

$$\left(A_{t_i}^{2+}, B_{t_i}^{2+}, t_i\right) = (0, 0, t_i).$$

3.2.4.3 Withdrawal, bonus, surrender event

According to the contract, if the PH is still alive at event time t_i, then he (she) is entitled to withdraw a certain amount W_{t_i} from his (her) policy, also if the account value is equal to 0. This guaranteed amount is given by

$$W_{t_i} = G\Delta t \cdot B_{t_i}^{2+},$$

where G is a constant defined by the contract. In a Static framework, the PH is supposed to withdraw exactly this guaranteed amount. Otherwise, in a optimization framework, he (she) may withdraw a fraction γ_i of the guaranteed withdrawn:

$$W_{t_i} = \gamma_i G\Delta t \cdot B_{t_i}^{2+}.$$

- The case $\gamma_i = 0$ corresponds to no withdrawal. In this case, the contract may provide a bonus (b_{t_i} is specified by the contract):
$$\left(A_{t_i}^{3+}, B_{t_i}^{3+}, t_i\right) = \left(A_{t_i}^{2+}, B_{t_i}^{2+}\left(1 + b_{t_i}\right), t_i\right).$$
- If $0 < \gamma_i \leq 1$ the PH withdraws at a lower rate than the standard rate, and the new state variables are
$$\left(A_{t_i}^{3+}, B_{t_i}^{3+}, t_i\right) = \left(\max\left(0, A_{t_i}^{2+} - W_{t_i}\right), B_{t_i}^{2+}, t_i\right).$$
- A third case is possible: the PH may want to withdraw more than the maximum admitted. In this case we suppose $\gamma_i \in]1, 2]$, where the case $\gamma_i = 2$ corresponds to a total surrender. We define
$$A' = \max\left(0, A_{t_i}^{2+} - G\Delta t \cdot B_{t_i}^{2+}\right).$$
The withdrawn amount is
$$W_{t_i} = G\Delta t \cdot B_{t_i}^{2+} + \left(\gamma_i - 1\right) A' \left(1 - \kappa_{t_i}\right).$$
where $\kappa_{t_i} \in [0, 1]$ is a penalty for withdrawal above the contract amount. The new state variables are
$$\begin{aligned}\left(A_{t_i}^{3+}, B_{t_i}^{3+}, t_i\right) &= \left(\max\left(0, A_{t_i}^{2+} - G\Delta t \cdot B_{t_i}^{2+} - \left(\gamma_i - 1\right) A'\right), \left(2 - \gamma_i\right) B_{t_i}^{2+}, t_i\right) \\ &= \left(\left(2 - \gamma_i\right) A', \left(2 - \gamma_i\right) B_{t_i}^{2+}, t_i\right).\end{aligned}$$

3.2.4.4 Ratchet

If the contract species a ratchet (step-up) feature, then the value of the benefit base B is increased if the investment account has increased. The guarantee account B can never decrease, unless the contract is partially or fully surrendered:

$$\left(A_{t_i}^{4+}, B_{t_i}^{4+}, t_i\right) = \left(A_{t_i}^{3+}, \max\left(B_{t_i}^{3+}, A_{t_i}^{3+}\right), t_i\right).$$

Another feature that may be included in the contract is roll-up: for seek of simplicity we won't treat this mechanism.

3.2.5 Similarity reduction

An important property of GLWB contract is the fact that these contract behave good under scaling transformations. If $\mathcal{V}(A,B,t)$ denotes the value of a contract, it is possible to prove that for any scalar $\eta > 0$

$$\eta\mathcal{V}(A,B,t) = \mathcal{V}(\eta A,\eta B,t). \tag{3.2.3}$$

Then, we just have to treat the case $B = \hat{B}$ for a fixed $\hat{B}$ (for example $\hat{B} = P$), and then, choosing $\eta = \hat{B}/B$, we can obtain

$$\mathcal{V}(A,B,t) = \frac{B}{\hat{B}}\mathcal{V}\left(\frac{\hat{B}}{B}A,\hat{B},t\right),$$

which means that we can solve the pricing problem only for a single representative value of B. This effectively reduces the problem dimension. The similarity reduction (4.2.5) was also exploited from Shah et Bertsimas in [43]. We can observe how the reduction similarity works both in the case of a contract that does not contain mechanisms for increasing the base benefits (ratchet), both for contracts with these properties.

3.3 The stochastic models of the fund S

To understand the different impacts of stochastic volatility and stochastic interest rate over such a long maturity contract, we price the GLWB VA according to two models: the Heston model, which provides stochastic volatility, and the Black-Scholes Hull-White model, which provide stochastic interest rate. As we said before, the process S represents the underlying fund driving the account value A_t of the product.

3.3.1 The Heston model

The Heston model [26] is one of the most known and used models in finance to describe the evolution of the volatility of an underlying asset and the underlying asset itself. In order to fix the notation, we report its dynamics:

$$\begin{cases} dS_t = rS_t dt + \sqrt{v_t}S_t dZ_t^S & S_0 = \bar{S}_0, \\ dv_t = k(\theta - v_t)dt + \omega\sqrt{v_t}dZ_t^v & v_0 = \bar{v}_0, \end{cases} \tag{3.3.1}$$

where Z^S and Z^v are Brownian motions, and $d\langle Z_t^S, Z_t^v\rangle = \rho dt$.

3.3.2 The Black-Scholes Hull-White model

The Hull-White model [28] is one of historically most important interest rate models, which is nowadays often used for risk-management purposes. The important advantage of the HW model is the existence of closed formulas to calculate the prices of bonds, caplets and swaptions. In order to fix the notation, we report the dynamics of the BS HW model:

$$\begin{cases} dS_t = r_t S_t dt + \sigma S_t dZ_t^S & S_0 = \bar{S}_0, \\ dr_t = k(\theta_t - r_t)dt + \omega dZ_t^r & r_0 = \bar{r}_0, \end{cases} \tag{3.3.2}$$

where Z^S and Z^r are Brownian motions, and $d\left\langle Z_t^S, Z_t^r\right\rangle = \rho dt$.

3.4 Numerical methods of pricing

In this Section we describe the four pricing methods: a Hybrid Monte Carlo method, a Standard Monte Carlo method, a Hybrid PDE method, and an ADI PDE method.

We remember that our aim is to find the fair value for α_g: it's the charge that makes the initial value of the policy equal to the initial gross premium. To achieve this target, we price the policy (with one of the following procedures) and then we use the secant method to approach the correct value for α_g. Therefore, the main goal is to be able to find the initial value for a given value of α_g: $\mathcal{V}(A_0, B_0, 0)(\alpha_g)$.

We remark that we want to calculate the value of the policy from the point of view of the insurance company: the management fees are treated as a outgoing cash flows, and if we assume that the policy holder follows a withdrawal strategy, we consider the worst one for the insurance company.

3.4.1 The Hybrid Monte Carlo method

The value of a GLWB policy can be calculated through a Monte Carlo set of simulations. This procedure is based on two steps: generation of a scenario (a sampling of the underlying values along the life of the product), and projection of the product into the scenario. According to the way we obtain the scenarios, we distinguish two Monte Carlo models: Hybrid MC (HMC) and Standard MC (SMC).

The Hybrid MC method has been explained in Section 1.4.

3.4.1.1 Scenario generation

The generations is done according to 1.4.1.1 and 1.4.1.2

3.4.1.2 Projection

Once we have generated the scenarios, we project the policy into them: it means we calculate the initial value of the contract as the sum of discounted cash flows. This calculation depends on whether we take an optimized strategy or not. Let $\mathcal{V}(A, B, t)$ be the value of a policy at time t, having account value equal to A and base benefit equal to B. From now on, we fix a specific scenario. Let $V(A, B, t)$ be the value of a policy in that scenario at time t, having account value equal to A and base benefit equal to B.

Constant withdrawal In this case the strategy of the PH is fixed: in each event time $\gamma_i = 1$ (for completeness we continue to write γ_i). A simple way to calculate the value of the policy is calculating forward the cash flows, conditioning on the death time. As in Holz et al. [27], we have:

$$V(A_0, B_0, 0) = \sum_{i=0}^{I} \mathcal{M}(t_i) \left(\sum_{k=0}^{i} e^{-\int_0^{t_k} r_s ds} W_{t_k} + e^{-\int_0^{t_i} r_s ds} A_{t_i}^{1+} \right).$$

Anyway, we developed another approach, useful for the optimal withdrawal case. First we calculate the values $\left(A_{t_i}^{4+}, B_{t_i}^{4+}, t_i\right)$ for all t_i neglecting the effect of mortality (equivalently, assuming the PH to die at the end), with a forward approach:

$$A_{t_i}^{4+} = \max\left(0, A_{t_{i-1}}^{4+}\frac{S_{t_i}}{S_{t_{i-1}}}e^{-\alpha_{tot}\Delta t} - \gamma_{t_i} G\Delta t B_{t_{i-1}}^{4+}\right),$$

$$B_{t_i}^{4+} = \begin{cases} \max\left(B_{t_{i-1}}^{4+}, A_{t_i}^{4+}\right) & \text{if ratchet,} \\ B_{t_{i-1}}^{2+} & \text{otherwise.} \end{cases}$$

Then, we proceed backwards, calculating the value of the contract at each time t_i just before the withdrawal. The value of the contract at time t_i can be written as the discounted value at time t_{i+1} plus the discounted value of the cash flows relating the period $\left[t_i^{4+}, t_{i+1}^{4+}\right]$. The final condition on the value of the contract is

$$V\left(A_T^{4+}, B_T^{4+}, T\right) = 0,$$

because all PHs are dead and all benefits have been paid. Then

$$V\left(A_{t_i}^{4+}, B_{t_i}^{4+}, t_i\right) = e^{-\int_{t_i}^{t_{i+1}} r_s ds}\left[V\left(A_{t_{i+1}}^{4+}, B_{t_{i+1}}^{4+}, t_{i+1}\right) + \mathcal{R}\left(t_{i+1}\right) W_{t_{i+1}}\right] + DB + MF,$$

where DB and MF stands for the discounted value in t_i of the death benefit and management fees paid in $\left[t_i^{4+}, t_{i+1}^{4+}\right]$. We distinguish four cases depending on how the management fees and the death benefit are paid. The proof of the following formulas is available in the Appendix.

CASE 1: DB paid at the end, fees withdrawn at the end

$$DB = \mathcal{M}\left(t_i\right) e^{-\int_{t_i}^{t_{i+1}} r_s ds} A_{t_i}^{4+}\frac{S_{t_{i+1}}}{S_{t_i}} e^{-\alpha_{tot}\Delta t},$$

$$MF = \mathcal{R}\left(t_i\right) e^{-\int_{t_i}^{t_{i+1}} r_s ds} A_{t_i}^{4+}\frac{S_{t_{i+1}}}{S_{t_i}}\left(1 - e^{-\alpha_{tot}\Delta t}\right)\frac{\alpha_m}{\alpha_{tot}}.$$

CASE 2: DB paid at the end, fees withdrawn continuously

$$DB = \mathcal{M}\left(t_i\right) e^{-\int_{t_i}^{t_{i+1}} r_s ds} A_{t_i}^{4+}\frac{S_{t_{i+1}}}{S_{t_i}} e^{-\alpha_{tot}\Delta t},$$

$$MF = \mathcal{R}\left(t_i\right)\alpha_m\frac{A_{t_i}^{4+}}{S_{t_i}}\int_{t_i}^{t_{i+1}} e^{-\int_{t_i}^{t} r_s ds} S_t e^{-\alpha_{tot}(t-t_i)} dt.$$

CASE 3: DB paid immediately, fees withdrawn at the end

$$DB = \mathcal{M}\left(t_i\right)\frac{A_{t_i}^{4+}}{S_{t_i}}\int_{t_i}^{t_{i+1}} e^{-\int_{t_i}^{t} r_s ds} S_t e^{-\alpha_{tot}(t-t_i)} dt,$$

$$MF = \mathcal{M}\left(t_i\right)\frac{\alpha_m}{\alpha_{tot}}\frac{A_{t_i}^{4+}}{S_{t_i}}\int_{t_i}^{t_{i+1}} S_t\left(1 - e^{-\alpha_{tot}(t-t_i)}\right) e^{-\int_{t_i}^{t} r_u du} dt + \\ + \mathcal{R}\left(t_{i+1}\right) e^{-\int_{t_i}^{t_{i+1}} r_s ds} A_{t_i}^{4+}\frac{S_{t_{i+1}}}{S_{t_i}}\left(1 - e^{-\alpha_{tot}\Delta t}\right)\frac{\alpha_m}{\alpha_{tot}}.$$

CASE 4: DB paid immediately, fees withdrawn continuously

$$DB = \mathcal{M}(t_i) \frac{A_{t_i}^{4+}}{S_{t_i}} \int_{t_i}^{t_{i+1}} e^{-\int_{t_i}^{t} r_s ds} S_t e^{-\alpha_{tot}(t-t_i)} dt,$$

$$MF = \mathcal{M}(t_i) \alpha_m \frac{A_{t_i}^{4+}}{S_{t_i}} \int_{t_i}^{t_{i+1}} S_t e^{-\alpha_{tot}(t-t_i)} e^{-\int_{t_i}^{t} r_u du} (t_{i+1} - t)\, dt + \\ + \mathcal{R}(t_{i+1}) \alpha_m \frac{A_{t_i}^{4+}}{S_{t_i}} \int_{t_i}^{t_{i+1}} e^{-\int_{t_i}^{t} r_s ds} S_t e^{-\alpha_{tot}(t-t_i)} dt.$$

Proceeding in this way, it is possible to calculate $V\left(A_0^{4+}, B_0^{4+}, 0\right)$. The initial value of the policy is

$$V\left(A_0^-, B_0^-, 0\right) = V\left(A_0^{4+}, B_0^{4+}, 0\right),$$

if the first withdrawal takes place at time $t = t_1$, or

$$V\left(A_0^-, B_0^-, 0\right) = V\left(A_0^{4+}, B_0^{4+}, 0\right) + \gamma_0 G \Delta t P$$

if the first withdrawal takes place at time $t = 0$. Then we simply have to calculate the average of $V\left(A_0^-, B_0^-, 0\right)$ among the simulated scenarios to approximate $\mathcal{V}\left(A_0^-, B_0^-, 0\right)$.

Optimal withdrawal In this case we suppose that at each event time t_i the PH can withdraw a fraction γ_i of the regular amount. To price in this case, we suppose that the PH chooses the value of γ that causes the worst hedging case for the insurance company. As we did before, we denote $\mathcal{V}(A, B, t)$ the expected value at time t of a generic policy whose state parameters are A, B :

$$\mathcal{V}(A, B, t) = \mathbb{E}\left[V(A, B, t)\right].$$

So, we suppose that the PH chooses γ_i such that

$$\gamma_i = \underset{\gamma \in [0,2]}{\operatorname{argmax}} \left[W_{t_i} + \mathcal{V}\left(A^{4+}, B^{4+}, t\right)\right].$$

This expected value can be calculated with a Longstaff-Schwartz approach:

1. Simulate N random scenarios and price the policy into these scenarios using random values for γ_i.

2. For $t = T$ to $t = 0$:

   (a) Approximate the function $\mathcal{V}(A, B, t)$ using the least squares projection into a space of functions (usually polynomials).

   (b) For each scenario find the optimal value of γ_t .

   (c) Recalculate the upcoming state variables from $s = t$ to $s = T$ assuming that the PH chooses the best value for γ.

3. Calculate the average of the initial value $V(A_0, B_0, 0)$ over all the scenarios.

The approximation of the function $\mathcal{V}(A, B, t)$ can be improved by the similarity reduction property.

3.4.2 Standard Monte Carlo method

The Monte Carlo method is very similar to the Hybrid Monte Carlo one. The only different thing, is the way we produce the random scenarios. The projection phase is the same as in Hybrid Monte Carlo.

3.4.2.1 Scenario generation

We distinguish two cases for the two models.

The Heston model The generation of the scenarios (underlying and volatility) in this case has been done using a third order scheme described in Alfonsi [2]. For more details, see Section 1.1.6.2.

The Black-Scholes Hull-White model The generation of the scenarios (underlying and interest rate) in this case has been done using an exact scheme described in Ostrovski [39], with a few changes in order to incorporate the correlation between underlying and interest rate. For more details, see Section 1.1.7

3.4.3 PDE Hybrid method

The Hybrid PDE approach is different from the previous ones. In fact it's a PDE pricing method and it's based on Briani et al. [10], [11] both for Heston and Hull-White case. Using a tree to diffuse the volatility or the interest rate, we freeze these values between two tree-levels; we mix the option values associated to the upcoming four nodes and then we solve one PDE (for each tree node) using the mixed data as starting values. For more details, see Section 1.4.2

We can resume the pricing methods in three features: model, algorithm structure and pricing.

3.4.3.1 The Heston model

Starting from the model in (3.3.1), we call $\bar{\rho} = \sqrt{1-\rho^2}$ and we write $Z_t^S = \rho Z_t^v + \bar{\rho}\bar{Z}_t^S$, where $\bar{Z}^S$ is a Brownian motion uncorrelated with Z^v. Then,

$$\begin{cases} dS_t = rS_t dt + \sqrt{v_t} S_t \left(\rho dZ_t^v + \bar{\rho} d\bar{Z}_t^S\right) & S_0 = \bar{S}_0, \\ dv_t = k\left(\theta - v_t\right) dt + \omega\sqrt{v_t} dZ_t^v & V_0 = \bar{V}_0, \end{cases} \quad d\left\langle \bar{Z}_t^S, Z_t^v \right\rangle = 0,$$

we define the process

$$E_t = \ln\left(A_t\right) - \frac{\rho}{\omega} v_t, \ E_0 = \ln\left(A_0\right) - \frac{\rho}{\omega} v_0,$$

$$A_t = \exp\left(E_t + \frac{\rho}{\omega} v_t\right). \tag{3.4.1}$$

Then

$$dE_t = \left(r - \frac{v_t}{2} - \frac{\rho}{\omega} k\left(\theta - v_t\right) - \alpha_{tot}\right) dt + \bar{\rho}\sqrt{v_t} d\bar{Z}_t^S,$$

if fees are taken continuously, otherwise

$$dE_t = \left(r - \frac{v_t}{2} - \frac{\rho}{\omega}k(\theta - v_t)\right)dt + \bar{\rho}\sqrt{v_t}d\bar{Z}_t^S.$$

This process E_t is important because it's a process uncorrelated with V_t, and we introduced it as in [10]. We are going to use it to define a PDE to be solved along the tree.

3.4.3.2 The Black-Scholes Hull-White model

Starting from the model **(3.3.2)**, we call $\bar{\rho} = \sqrt{1-\rho^2}$ and we write $Z_t^S = \rho Z_t^r + \bar{\rho}\bar{Z}_t^S$, where $\bar{Z}^S$ is a Brownian motion uncorrelated with Z^r. Then,

$$\begin{cases} dS_t = r_t S_t dt + \sigma S_t\left(\rho dZ_t^r + \bar{\rho}d\bar{Z}_t^S\right) & S_0 = \bar{S}_0, \\ dX_t = -kX_t dt + dZ_t^r & X_0 = 0, \quad d\left\langle \bar{Z}_t^S, Z_t^r\right\rangle = 0, \\ r_t = \omega X_t + \beta(t), \end{cases}$$

we define the process

$$U_t = \ln(A_t) - \rho\sigma X_t, \; U_0 = \ln(A_0),$$
$$A_t = \exp(U_t + \rho\sigma X_t). \tag{3.4.2}$$

Then

$$dU_t = \left(r_t - \frac{\sigma^2}{2} + \sigma\rho k X_t - \alpha_{tot}\right)dt + \sigma\bar{\rho}d\bar{Z}_t^S,$$

if fees are taken continuously, otherwise

$$dU_t = \left(r_t - \frac{\sigma^2}{2} + \sigma\rho k X_t\right)dt + \sigma\bar{\rho}d\bar{Z}_t^S.$$

This process U_t is important because it's a process uncorrelated with X_t, and we introduced it as in [10]. We are going to use it to define a PDE to be solved along the tree.

3.4.3.3 Algorithm structure

The structures for this algorithm consists in a tree and a PDE solver. As described in Briani et al. [10],[11], we use a tree to diffuse the volatility (or the interest rate) along the life of the product, and we solve backward four 1D PDEs freezing at each node of the tree the volatility (or the interest rate) and using different initial data. The tree is built according to Section 1.2.4 (quadrinomial tree, matching the first three moments of the process), and the PDEs are solved with a finite difference approach. We have to solve the PDEs between two event times, and at each event time we apply the changes to the states to reproduce the effects of the events.

We remark that we solve the PDEs doing a single time step that requires only a linear complexity because we have to solve a linear system with tridiagonal matrix. The computational cost is low as observed in [10] and [11]. We observe that X_t and V_t processes are mean reverting. Thanks to the way the trees are built, there are many nodes in the trees that cannot be visited by the approximating Markov chain. Therefore their probability $p_{n,j}$ to be visited is worth 0 and they have no impact on the values at the root of the tree. There is no reason to do any operation for those nodes. So, to save time, we do the standard step (solve backward the four PDEs and mix up the vectors according to the transition probabilities) only for those nodes having $p_{n,j} > 0$. This curtailing technique reduces the computational time, and the convergence of the method is preserved. A similar approach is used in [3].

3.4.3.4 Pricing

The PDE we have to solve at each node is essentially the same as in Forsyth and Vetzal [21]. We distinguish four cases as we did in Monte Carlo case. We denote with $\mathcal{V}(A,B,t)$ the value of a contract at time t whose account value is worth A and whose base benefit is worth B. Consequently, we define

$$\mathcal{V}^{He}(E,B,t) = \mathcal{V}\left(\exp\left(E + \frac{\rho}{\omega}V_t\right), B, t\right),$$

and

$$\mathcal{V}^{HW}(U,B,t) = \mathcal{V}\left(\exp\left(U + \rho\sigma X_t\right) B, t\right).$$

The variables $\bar{r}$, $\bar{X}$ and $\bar{V}$ will denote the frozen values of r_t, X_t and V_t. We solve the transformed PDEs between two event times for each node of the tree four times: one for each of the possible next nodes, using the initial data corresponding to these nodes. To reduce the run time, we do this only for active nodes ($p_{n,j} > 0$): this cutting technique dramatically reduced calculation times without compromising the quality of results. Then, using the inverse transformations (3.4.1) and (3.4.2), we apply the event times actions. In the next few paragraphs, we are going to write 2 PDEs: one for the Heston model, and one for the BS HW model.

CASE 1: DB paid at the end, fees withdrawn at the end

The terminal condition is

$$\mathcal{V}(A,B,T) = \mathcal{R}(T-\Delta t)\, A\left(1 - \left(1 - e^{-\alpha_{tot}\Delta t}\right)\frac{\alpha_g}{\alpha_{tot}}\right).$$

The associated PDEs are

$$\mathcal{V}_t^{He} + \frac{\bar{\rho}^2\bar{V}}{2}\mathcal{V}_{EE}^{He} + \left(r - \frac{\bar{v}}{2} - \frac{\rho}{\omega}k(\theta - \bar{v})\right)\mathcal{V}_E^{He} - r\mathcal{V}^{He} = 0\,, \qquad \text{(He 1)}$$

$$\mathcal{V}_t^{HW} + \frac{\bar{\rho}^2\sigma^2}{2}\mathcal{V}_{UU}^{HW} + \left(\bar{r} - \frac{\sigma^2}{2} + \sigma\rho k\bar{X}\right)\mathcal{V}_U^{HW} - \bar{r}\mathcal{V}^{HW} = 0\,. \qquad \text{(HW 1)}$$

For $t_i = T-1$ to $t_i = 0$ we have to:

1. Solve the PDE backward from t_{i+1} to t_i;
2. Calculate the value of $\mathcal{V}$ from the value of $\mathcal{V}^{He}$ or $\mathcal{V}^{HW}$;
3. In case of ratchet $\mathcal{V}\left(A,B,t_i^{3+}\right) = \mathcal{V}\left(A, \max(A,B), t_i^{4+}\right)$;

4. Withdrawal:

   (a) if $\gamma_{t_i} = 0$:
   $$\mathcal{V}\left(A, B, t_i^{2+}\right) = \mathcal{V}\left(A, B\left(1 + b_{t_i}\right), t_i^{3+}\right);$$

   (b) if $\gamma_{t_i} \in [0, 1]$:
   $$\mathcal{V}\left(A, B, t_i^{2+}\right) = \mathcal{V}\left(\max\left(0, A - \gamma_{t_i} G\Delta t B\right), B, t_i^{3+}\right) + \mathcal{R}\left(t_i\right)\gamma_{t_i} G\Delta t B;$$

   (c) if $\gamma_{t_i} \in]1, 2]$:
   $$\begin{aligned}\mathcal{V}\left(A, B, t_i^{2+}\right) = \mathcal{V}\left(\max\left(0, A - G\Delta t B\right)\left(2 - \gamma_{t_i}\right), B\left(2 - \gamma_{t_i}\right), t_i^{3+}\right) + \\ + \mathcal{R}\left(t_i\right)\left(G\Delta t B + \left(\gamma_{t_i} - 1\right)\max\left(0, A - G\Delta t B\right)\left(1 - \kappa_{t_i}\right)\right);\end{aligned}$$

5. Death benefit: $\mathcal{V}\left(A, B, t_i^{1+}\right) = \mathcal{V}\left(A, B, t_i^{2+}\right) + \left(\mathcal{R}\left(t_{i-1}\right) - \mathcal{R}\left(t_i\right)\right)A$;

6. Fees: $\mathcal{V}\left(A, B, t_i^{-}\right) = \mathcal{V}\left(Ae^{-\alpha_{tot}\Delta t}, B, t_i^{1+}\right) + \mathcal{R}\left(t_{i-1}\right)\frac{\alpha_m}{\alpha_{tot}}A\left(1 - e^{-\alpha_{tot}\Delta t}\right)$;

7. Calculate the value of $\mathcal{V}^{He}$ or $\mathcal{V}^{HW}$ from the value of $\mathcal{V}$.

CASE 2: DB paid at the end, fees withdrawn continuously

The differences between this case and the case 1 are the following ones. The terminal condition is
$$\mathcal{V}\left(A, B, T\right) = \mathcal{R}\left(T - \Delta t\right)A.$$

The associated PDEs are

$$\mathcal{V}_t^{He} + \frac{\bar{\rho}^2\bar{V}}{2}\mathcal{V}_{EE}^{He} + \left(r - \frac{\bar{v}}{2} - \frac{\rho}{\omega}k\left(\theta - \bar{v}\right) - \alpha_{tot}\right)\mathcal{V}_E^{He} - r\mathcal{V}^{He} + \alpha_m\mathcal{R}\left(t\right)\exp\left(E_t + \frac{\rho}{\omega}\bar{v}\right) = 0\ , \quad \text{(He 2)}$$

$$\mathcal{V}_t^{HW} + \frac{\bar{\rho}^2\sigma^2}{2}\mathcal{V}_{UU}^{HW} + \left(\bar{r} - \frac{\sigma^2}{2} + \sigma\rho k\bar{X} - \alpha_{tot}\right)\mathcal{V}_U^{HW} - \bar{r}\mathcal{V}^{HW} + \alpha_m\mathcal{R}\left(t\right)\exp\left(U_t + \rho\sigma\bar{X}\right) = 0\ . \quad \text{(HW 2)}$$

Point 6 (fees step) becomes
$$\mathcal{V}\left(A, B, t_i^{-}\right) = \mathcal{V}\left(A, B, t_i^{1+}\right).$$

CASE 3: DB paid immediately, fees withdrawn at the end

The differences between this case and the case 1 are the following ones. The terminal condition is
$$\mathcal{V}\left(A, B, T\right) = 0.$$

The associated PDEs are

$$\mathcal{V}_t^{He} + \frac{\bar{\rho}^2\bar{v}}{2}\mathcal{V}_{EE}^{He} + \left(r - \frac{\bar{v}}{2} - \frac{\rho}{\omega}k\left(\theta - \bar{v}\right)\right)\mathcal{V}_E^{He} - r\mathcal{V}^{He} + \mathcal{M}\left(t_i\right)\exp\left(E_t + \frac{\rho}{\omega}\bar{V}\right)\left(1 - \left(1 - e^{-\alpha_{tot}(t - t_i)}\right)\frac{\alpha_g}{\alpha_{tot}}\right) = 0\ , \quad \text{(He 3)}$$

$$\mathcal{V}_t^{HW}+\frac{\bar{\rho}^2\sigma^2}{2}\mathcal{V}_{UU}^{HW}+\left(\bar{r}-\frac{\sigma^2}{2}+\sigma\rho k\bar{X}\right)\mathcal{V}_U^{HW}-\bar{r}\mathcal{V}^{HW}+\mathcal{M}\left(t_i\right)\exp\left(U_t+\rho\sigma\bar{X}\right)\left(1-\left(1-e^{-\alpha_{tot}(t-t_i)}\right)\frac{\alpha_g}{\alpha_{tot}}\right)=0\,. \tag{HW 3}$$

Point 5 (death benefit step) and 6 (fees step) become:

- Death benefit: $\mathcal{V}\left(A,B,t_i^{1+}\right)=\mathcal{V}\left(A,B,t_i^{2+}\right)$.
- Fees: $\mathcal{V}\left(A,B,t_i^{-}\right)=\mathcal{V}\left(Ae^{-\alpha_{tot}\Delta t},B,t_i^{1+}\right)+\mathcal{R}\left(t_i\right)\frac{\alpha_m}{\alpha_{tot}}A\left(1-e^{-\alpha_{tot}\Delta t}\right)$.

CASE 4: DB paid immediately, fees withdrawn continuously

The differences between this case and the case 1 are the following ones. The terminal condition is

$$\mathcal{V}\left(A,B,T\right)=0.$$

The associated PDEs are

$$\mathcal{V}_t^{He}+\frac{\bar{\rho}^2\bar{v}}{2}\mathcal{V}_{EE}^{He}+\left(r-\frac{\bar{v}}{2}-\frac{\rho}{\omega}k\left(\theta-\bar{v}\right)-\alpha_{tot}\right)\mathcal{V}_E^{He}-r\mathcal{V}^{He}+\exp\left(E_t+\frac{\rho}{\omega}\bar{v}\right)\left(\alpha_m\mathcal{R}\left(t\right)+\mathcal{M}\left(t_i\right)\right)=0\,, \tag{He 4}$$

$$\mathcal{V}_t^{HW}+\frac{\bar{\rho}^2\sigma^2}{2}\mathcal{V}_{UU}^{HW}+\left(\bar{r}-\frac{\sigma^2}{2}+\sigma\rho k\bar{X}-\alpha_{tot}\right)\mathcal{V}_U^{HW}-\bar{r}\mathcal{V}^{HW}+\exp\left(U_t+\rho\sigma\bar{X}\right)\left(\alpha_m\mathcal{R}\left(t\right)+\mathcal{M}\left(t_i\right)\right)=0\,. \tag{HW 4}$$

Point 5 (death benefit step) and 6 (fees step) become

$$\mathcal{V}\left(A,B,t_i^{-}\right)=\mathcal{V}\left(A,B,t_i^{1+}\right)=\mathcal{V}\left(A,B,t_i^{2+}\right).$$

This concludes the Static withdrawal case. In the optimal withdrawal case, we suppose the PH to change the value of γ_i used in step n. 4 (withdrawal step). He (she) will choose the value of $\gamma_i\in[0,2]$ in order to maximizes the value of $\mathcal{V}\left(A,B,t_i^{2+}\right)$. This maximization can be done using a grid of values for γ_i and choosing at each time the best value.

3.4.4 PDE ADI method

Consider the asset price process given by the system of stochastic differential equations described in Section 4.2. We describe the ADI method only in the case 2, but the other cases can be easily adapted. Moreover, we have chosen to not use the transformed PDE described in Section 4.4.3.4, but the classical version of PDEs for the Black-Scholes, Heston and Black-Scholes Hull-White model. The associated PDEs are

$$\mathcal{V}_t^{He}+\frac{VA^2}{2}\mathcal{V}_{AA}^{He}+\frac{\omega^2v}{2}\mathcal{V}_{vv}^{He}+\left(r-\alpha_{tot}\right)A\mathcal{V}_A^{He}+\rho\omega Av\mathcal{V}_{Av}^{He}+k\left(\theta-v\right)\mathcal{V}_v^{He}-r\mathcal{V}^{He}+\alpha_m\mathcal{R}\left(t\right)A=0 \tag{He 2b}$$

$$\mathcal{V}_t^{HW}+\frac{\sigma^2 A^2}{2}\mathcal{V}_{AA}^{HW}+\frac{\omega^2}{2}\mathcal{V}_{rr}^{HW}+(r-\alpha_{tot})\,A\mathcal{V}_A^{HW}+\rho\omega A\sigma\mathcal{V}_{Ar}^{HW}+k\,(\theta_t-r)\,\mathcal{V}_r^{HW}-r\mathcal{V}^{HW}+\alpha_m\mathcal{R}\,(t)\,A=0 \tag{HW 2b}$$

Because of the long maturity, solving a two-dimensional PDE is a very costly and slow method. The idea is to use splitting schemes of ADI (alternating directional implicit) type. In this Chapter, we only present the Douglas scheme, but various scheme are available in the literature. In order to solve the PDE, we should address many numerical difficulties. The first one is the mesh and we have chosen to use the meshes described in [25] with the parameters

$$A_{left}=0.8S_0 \quad A_{right}=1.2S_0 \quad A_{max}=100S_0 \quad \text{and } d_1=S_0/20,$$

for the mesh of variable A,

$$R_{\max}=10R_0, \quad c=R_0 \quad \text{and } d_2=R_{\max}/400$$

for the mesh of variable r in the Black-Scholes Hull-White model, and

$$v_{\max}=MIN(MAX(100v_0,1),5) \quad \text{and } d_3=v_{\max}/500.$$

for the mesh of variable v in the Heston model. The second difficulty is the choice of the splitting scheme. We have chosen the Douglas scheme with parameter $\theta=1/2$ because it is the easiest to implement, but of course some higher order schemes (in time) would be more optimal. The last difficulty, but not the least, is the choice of boundary conditions. Since there is no closed form solutions for the GLWB product, it is difficult to make the right choice for the boundary conditions. Moreover the boundary conditions have a big impact on the solution, because of the long maturity. Usually the choice of boundary conditions have no importance. Indeed since there are many points in the mesh between S_0 and S_{max}, the system of equations does not connect these distant values in a hard way. For example, choosing homogeneous Neumann, non-homogeneous Neumann or Dirichlet conditions can lead to very closed prices. But, because of the long maturity, the prices are really impacted by a bad choice of boundary conditions. Actually the system (for one year) will be solved many times and it will connect every points in the mesh in a very intricate manner. We now describe the boundary conditions in the Heston and Hull-White models. The choice of homogeneous Neumann conditions is usually done because it simplifies the system to solve (exactly it simplifies the finite difference scheme at the boundary). In the context of GLWB, the boundary conditions for the Black-Scholes Hull-White model will be given by:

$$\begin{aligned}
&\frac{\partial\mathcal{V}_t^{HW}}{\partial s}(A,r,t)=0, && \text{if } A=0 \text{ or } A=A_{\max},\\
&\frac{\partial\mathcal{V}_t^{HW}}{\partial r}(A,r,t)=0, && \text{if } r=\pm R_{\max},
\end{aligned}$$

on the mesh $[0,A_{\max}]\times[-R_{\max},R_{\max}]$, and the boundary condition for the Heston model will be given by:

$$\begin{aligned}
&\frac{\partial\mathcal{V}_t^{He}}{\partial s}(A,v,t)=0, && \text{if } A=0 \text{ or } A=A_{\max},\\
&\frac{\partial\mathcal{V}_t^{He}}{\partial v}(A,v,t)=0, && \text{if } v=v_{\max},
\end{aligned}$$

on the mesh $[0, A_{\max}] \times [0, v_{\max}]$, and with no condition at $v = 0$ since it is an outflow boundary.

3.5 Numerical results

In this Section we compare the numerical methods introduced in Section 4.4: Hybrid Monte Carlo (*HMC*), Standard Monte Carlo (*SMC*), Hybrid PDE (*HPDE*), and ADI PDE (*APDE*). In particular we compare pricing and Greeks computation in *Static Case* 4.5.2 and *Dynamic Case* 4.5.3.

We chose the parameters of the methods according to 4 configurations (*A, B, C, D*), with an increasing number of steps and so that the calculation time for the various methods in each configuration were close. The 4 configurations parameters are reported in Table 4.1 with the notation (time steps per year × space steps × vol steps) for the ADI PDE method, (time steps per year × space steps) for the Hybrid PDE method and (time steps per year × number of simulations) for the MC's one. In Monte Carlo for Dynamic case, we also add the degree of the approximating polynomial. These values had been chosen to achieve approximately these run times: (A) 30 s, (B) 120 s, (C) 480 s, (D) 1900 s. To reduce the run time we did the secant iterations using an increasing number of time steps for all the methods: the values in Table 4.1 are those used for the last 3 iterations.

We use the Standard MC both as a pricing method, both as a benchmark (BM). About the benchmark, in the Static case we used 10^7 independent runs. In the Dynamic case we used 10^6 independent runs, arranged in 10 sub runs; in each sub runs the expected value has been approximated by a 6 order polynomial. At each event time, the policy holder can chose between $\gamma = 0$, $\gamma = 1$ and $\gamma = 2$.

The search for the fair α_g value has been driven by the secant method. The initial values for this method were $\alpha_g = 0$ bp and $\alpha_g = 200$ bp.

To achieve Delta calculation in Monte Carlo methods we used a 1‰ shock in Static case and 1% in Dynamic case.

We used the DAV 2004R mortality Table, 65 year old German male (see [21] for the Table). It contains the probabilities that a person aged t will die within the next year. It's easy to get the function $\mathcal{M}$ from these probabilities.

3.5.1 Static case

In the Static case we suppose the PH to withdrawal exactly at the guaranteed rate: $\gamma_t = 1$.

The Static tests 1 and 2 are inspired by [21]: in their article, Forsyth and Vetzal price a GLWB contract in a Static framework, under the Black Scholes model with $r = 0.04$ and $\sigma = 0.15$. The contract parameters are reported in the Table 3.2; the contract type corresponds to *case 2* in Section 4.4.1.2 and 4.4.3.4.

	BS HW Static				Heston Static			
	HMC	SMC	HPDE	APDE	HMC	SMC	HPDE	APDE
A	$5\times 1.3\cdot 10^5$	$1\times 2.7\cdot 10^5$	30×400	$18\times 180\times 36$	$5\times 8.6\cdot 10^4$	$5\times 7.4\cdot 10^4$	35×400	$26\times 260\times 13$
B	$10\times 2.3\cdot 10^5$	$1\times 9.8\cdot 10^5$	60×600	$27\times 270\times 54$	$10\times 1.6\cdot 10^5$	$10\times 1.4\cdot 10^5$	70×600	$40\times 400\times 20$
C	$20\times 5.4\cdot 10^5$	$1\times 4.9\cdot 10^6$	100×1000	$40\times 400\times 80$	$20\times 3.8\cdot 10^5$	$20\times 3.5\cdot 10^5$	100×1000	$64\times 640\times 32$
D	$40\times 1.0\cdot 10^6$	$1\times 2.0\cdot 10^7$	200×2000	$62\times 620\times 124$	$40\times 7.3\cdot 10^5$	$40\times 7.5\cdot 10^5$	200×2000	$104\times 1040\times 52$
	BS HW Dynamic				Heston Dynamic			
	HMC	SMC	HPDE	APDE	HMC	SMC	HPDE	APDE
A	$5\times 3.3\cdot 10^3\times 2$	$5\times 3.2\cdot 10^3\times 2$	30×400	$16\times 160\times 32$	$5\times 3.2\cdot 10^3\times 2$	$5\times 3.2\cdot 10^3\times 2$	35×400	$22\times 220\times 11$
B	$10\times 1.6\cdot 10^4\times 3$	$5\times 1.6\cdot 10^4\times 3$	60×600	$24\times 240\times 48$	$10\times 1.5\cdot 10^4\times 3$	$10\times 1.5\cdot 10^4\times 3$	70×600	$36\times 360\times 18$
C	$20\times 5.2\cdot 10^4\times 4$	$5\times 5.3\cdot 10^4\times 4$	100×1000	$38\times 380\times 76$	$20\times 4.9\cdot 10^4\times 4$	$20\times 4.9\cdot 10^4\times 4$	100×1000	$60\times 600\times 30$
D	$40\times 1.4\cdot 10^5\times 5$	$5\times 1.6\cdot 10^5\times 5$	200×2000	$60\times 600\times 120$	$40\times 1.3\cdot 10^5\times 5$	$40\times 1.3\cdot 10^5\times 5$	200×2000	$100\times 1000\times 50$

Table 3.1: Configuration parameters for the BS HW model and for the Heston model, Static and Dynamic.

Initial age of PH	65	Gr. premium	100	DB payment	next anniv.
G	0.05	Initial fees	0	Ratchet	Off/On (annual)
Withdrawal rate	1 per Y	α_m	0	Strategy	Static ($\gamma = 1$)
First withdrawal	1^{st} anniv.	Fees taken	cont.ly		

Table 3.2: The contract parameters for Static tests (except Test 2B).

They treated two cases: no ratchet, and annual ratchet. In the first case they get $\alpha_g = 35.51$ bp and in the second case $\alpha_g = 64.92$ bp. In Test 1 and Test 2 we introduce respectively stochastic interest rate and stochastic volatility to analyze the impact of these model developments on the fair guarantee fee. The parameters for interest rate and volatility have been chosen to be plausible.

To compare our results in the Heston model with Kling's ones in [30] we performed test 2B. In this case, product parameters are reported in Table 3.3, and correspond to *case 1* in Section 4.4.1.2 and 4.4.3.4.

Initial age of PH	65	Gr. premium	100	DB payment	next anniv.
G	4.90%, 4.19% if ratchet	Initial fees	4%	Ratchet	Off/On (annual)
Withdrawal rate	1 per Y	α_m	151 bp	Strategy	Static ($\gamma = 1$)
First withdrawal	1^{st} anniv.	Fees taken	at the end		

Table 3.3: The contract parameters for Test 2B-Static.

3.5.1.1 Test 1-Static: the Black-Scholes Hull-White model

In this test we want to price a product according to BS HW model. We use the same corresponding parameters as in test [21]. Model parameters are shown in Table 3.4. Results are available in Table 3.5.

All four methods behaved well and in the configuration D, gave results consistent with the benchmark. HPDE proved to be the best: all configurations gave results consistent with the benchmark. Then APDE and SMC, and HMC gave good results too. SMC performed a little better than HMC: the first method simulates the underlying value and the interest rate exactly and so it is enough to simulate the values at each event time. HMC matches the first three moments of the BS HW r process, but doesn't reproduce exactly its law: therefore it is right to increase the number of steps per year. So, for a given run time, we can simulate less scenarios in HMC than SMC: effectively, the confidence interval of HMC is larger than SMC one. Moreover, SMC over performed the benchmark when using configuration D. Particularly, the correlation between underlying and interest rate has a fundamental role, and its impact can be bigger than impact of the ratchet: for example, case no ratchet with $\rho = 0.5$ gave a higher price than case ratchet with $\rho = -0.5$ (111 bp vs 84 bp).

3.5.1.2 Test 2-Static: the Heston model

In this test we want to price a product according to the Heston model. Model parameters are shown in the Table 3.6. Results are shown in Table 3.7.

In this Test, MC methods had more problems; the values of PDE methods are close to the benchmark, while values from MC methods were far, but compatibles with the benchmark (the value of BM is inside MC confidence interval). Probably, in this case, the benchmark is not very accurate: this is due to the fact we used SMC to calculate it. If we compare the two MC approaches, in this case, they both use a third order approximation and than they become equivalent: HMC proved to be faster than SMC when using few time steps (we could exploit +16% simulations in configuration A), while SMC proved to be slightly faster in high time steps simulations, because of more time needed to build the volatility tree (−3% simulations in configuration D). HPDE showed to be very stable (case no ratchet, $\rho = -0.5$, α_g didn't change through configurations B-D), but APDE behaved well to (monotone convergence). In the Heston model, correlation has a less important role than in BS HW case: among the different values of ρ, the value of α_g changes less then 5 bp in no-ratchet case, and less than 1.5 bp in ratchet case.

3.5.1.3 Test 2B-Static: the Heston model

In this test we want to obtain the results shown in [30], where the contracts are priced with MC techniques. The values given in [30] are 150 bp for both cases (no ratchet and ratchet). Model parameters are given in Table 3.8. Results are available in Table 4.24.

In this Test, all methods gave the same results, but not the same results as in Kling et al. [30]. One possibility is that we have misinterpreted some of the contractual specifications in Kling's paper, leading to some subtle differences in the contracts that we are considering as compared to theirs, and these discrepancies result in different fees. Another potential explanation is that a Monte Carlo method was used to determine the fee by Kling et al.; this may have introduced a significant error when calculating the fee unless a very large number of simulations was used. They didn't report a confidence interval for their results, so it's hard to understand the cause of the gap. Moreover we can observe that, in this Test, our two MC methods gave larger confidence intervals than Test 2-Static: probably, the parameters used for Test 2B-Static shape a harder pricing problem than the previous test, and more simulations should be performed to obtain the same quality results. Also in this case, HPDE proved to be the most stable method.

3.5.1.4 Test 3-Static: Hedging

To reduce financial risks, insurance companies have to hedge the sold VA: to accomplish this target they must calculate the Greeks of products.

In this test we want to show how the different methods can be used to calculate the main Greeks. This can be done through finite differences for small shocks on the variables. In general, the PDE methods are ahead w.r.t. MC methods: the price is computed through finite differences and so the price under shock is already computed. For MC methods this is quite harder because the pricing has to be repeated changing the inputs.

To start, we calculate the underlying greek Delta, for the products of Test 1-Static and Test 2-Static. As in this case we don't want to compute the fair fee α_g, we fix it arbitrarily. We choose two values for each model: one for no ratchet case, and one for ratchet case. The values chosen are such as to cover the costs of the insurer regardless of the correlation, and may be plausible on a real case. Results are available in Table 4.10 (all values in Table must be multiplied by 10^{-4}).

In this Test, we got very accurate results with all method. Anyway, HPDE proved to be the best: it's the more stable and accurate. We remark that despite fair fee changes a lot when changing the correlation parameter ρ, the value of Delta changes much less. Delta calculation proved to be harder in the Heston model case than in the BS HW model case.

3.5.1.5 Test 4-Static: Risk Management

Mortality and longevity risks are unhedgeable risks. Usually, the Risk Management Team has to calculate the financial reserve taking into account these risks. Usually extreme scenarios are chosen and policies are priced according to them. In this Test we analyzed how the different pricing methods behaved under mortality shocks: the mortality probabilities have been increased by 10% except the last one who's equal to 1. To be brief, we simply report the fair fee for D case. Results are available in Table 3.11.

In this Test, we got results similar to Test 1-Static and Test 2-Static, and mortality shocks didn't affect the convergence quality of the four methods. We observe that mortality shocks reduce the value of α_g (about minus 5 bp) and this means that an increase in mortality shouldn't be a source of losses for the insurer. Consequently, insurers should pay attention to longevity risk.

S_0	r	*curve*	k	ω	ρ	σ
100	0.04	*flat*	1.0	0.2	*variable*	0.15

Table 3.4: The model parameters about Test 1-Static.

ρ		*no ratchet*					*annual ratchet*				
		HMC	SMC	HPDE	APDE	BM	HMC	SMC	HPDE	APDE	BM
−0.5	*A*	45.99 ± 1.06	45.58 ± 0.73	45.72	47.35	45.81	84.85 ± 1.23	84.58 ± 0.85	84.56	88.20	84.71
	B	45.52 ± 0.79	45.31 ± 0.38	45.71	46.09		84.35 ± 0.91	84.15 ± 0.44	84.60	85.66	
	C	45.58 ± 0.52	45.71 ± 0.17	45.69	45.99	±0.12	84.31 ± 0.60	84.65 ± 0.20	84.63	85.36	±0.14
	D	45.85 ± 0.38	45.71 ± 0.08	45.72	45.81		84.68 ± 0.44	84.63 ± 0.10	84.64	84.94	
0	*A*	82.29 ± 1.39	81.40 ± 0.95	81.87	82.91	81.88	157.77±1.65	156.77±1.14	156.36	161.04	157.09
	B	82.43 ± 1.04	81.53 ± 0.50	81.92	81.75		157.68±1.23	156.55±0.59	156.46	157.91	
	C	81.62 ± 0.68	81.77 ± 0.22	81.80	81.89	±0.16	156.50±0.80	157.05±0.27	156.87	157.70	±0.19
	D	81.99 ± 0.50	81.83 ± 0.11	81.79	81.81		157.16±0.59	157.07±0.13	156.96	157.27	
+0.5	*A*	111.75±1.76	110.30±1.20	111.14	109.23	111.05	224.19±2.15	222.14±1.48	221.78	227.14	222.83
	B	112.73±1.32	110.85±0.63	111.07	108.93		224.59±1.60	222.26±0.77	222.32	223.44	
	C	110.89±0.86	111.08±0.28	111.05	109.93	±0.20	222.18±1.05	222.94±0.35	222.52	223.36	±0.24
	D	111.29±0.63	111.11±0.14	111.02	110.42		222.97±0.77	222.94±0.17	222.67	222.96	

	HMC	SMC	HPDE	APDE
A	30 s	30 s	30 s	28 s
B	119 s	120 s	128 s	184 s
C	472 s	478 s	395 s	461 s
D	1866 s	1896 s	1903 s	1800 s

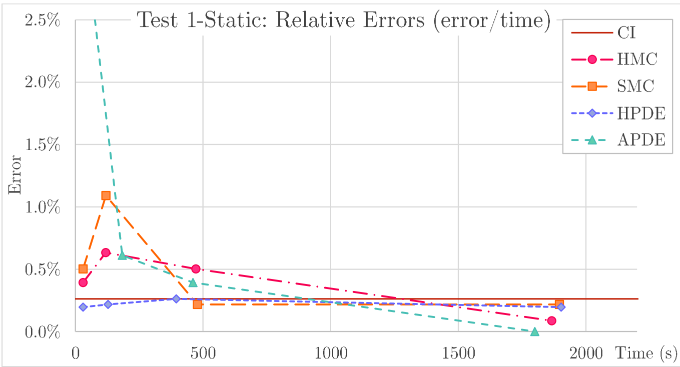

Table 3.5: Test 1-Static. In the first Table, the fair fee for the Black-Scholes Hull-White model, with no ratchet or annual ratchet. In the second Table the run times for the no-ratchet case ($\rho = -0.5$). Finally, the plot of relative error (w.r.t. BM value) for the four methods in the case $\rho = -0.5$ with no ratchet. The parameters used for this test are available in Table 3.2 and in Table 3.4.

S_0	V_0	θ	k	ω	ρ	r
100	0.15^2	0.15^2	1.0	0.2	*variable*	0.04

Table 3.6: The model parameters about Test 2-Static.

ρ		*no ratchet* HMC	SMC	HPDE	APDE	BM	*annual ratchet* HMC	SMC	HPDE	APDE	BM
-0.5	A	36.77 ± 1.25	36.17 ± 1.36	37.00	37.41	37.16	61.47 ± 1.23	60.90 ± 1.35	61.51	62.30	61.84
	B	36.74 ± 0.92	36.40 ± 0.99	37.01	37.26		61.29 ± 0.90	60.85 ± 0.97	61.59	62.06	
	C	36.79 ± 0.59	36.94 ± 0.62	37.01	37.11	± 0.12	61.36 ± 0.58	61.56 ± 0.61	61.63	61.80	± 0.11
	D	37.47 ± 0.43	37.33 ± 0.42	37.01	37.06		62.15 ± 0.42	61.95 ± 0.42	61.66	61.77	
0	A	35.67 ± 1.56	34.02 ± 1.61	35.18	35.52	35.22	63.22 ± 1.60	61.63 ± 1.65	62.56	63.43	62.64
	B	34.53 ± 1.13	34.48 ± 1.22	35.18	35.39		61.88 ± 1.17	61.64 ± 1.25	62.55	63.11	
	C	35.05 ± 0.74	35.05 ± 0.77	35.15	35.24	± 0.15	62.39 ± 0.76	62.37 ± 0.79	62.59	62.78	± 0.11
	D	35.28 ± 0.54	35.47 ± 0.53	35.15	35.19		62.47 ± 0.54	62.97 ± 0.55	62.59	62.68	
$+0.5$	A	33.70 ± 2.02	32.43 ± 2.14	32.58	32.76	32.63	61.44 ± 2.06	63.26 ± 2.31	62.84	63.99	62.97
	B	31.45 ± 1.43	32.26 ± 1.64	32.58	32.77		62.58 ± 1.61	62.41 ± 1.72	62.90	63.62	
	C	32.63 ± 0.96	32.94 ± 0.99	32.54	32.68	± 0.19	63.53 ± 1.02	62.94 ± 1.06	62.88	63.22	± 0.20
	D	32.31 ± 0.69	33.00 ± 0.72	32.52	32.65		62.55 ± 0.83	62.43 ± 0.86	62.89	63.09	

	HMC	SMC	HPDE	APDE
A	30 s	30 s	32 s	30 s
B	122 s	119 s	131 s	114 s
C	477 s	476 s	410 s	491 s
D	1915 s	1907 s	1755 s	1933 s

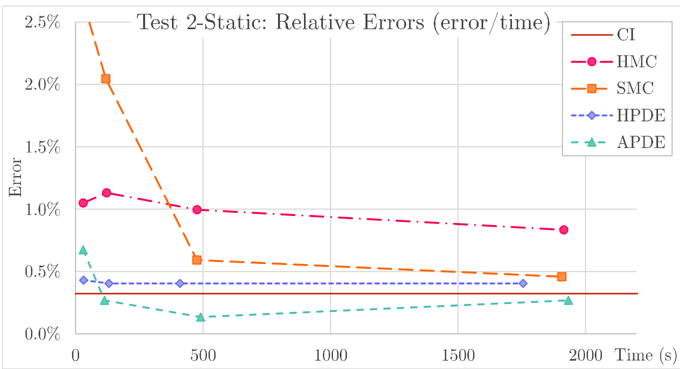

Table 3.7: Test 2-Static. In the first Table, the fair fee for the Heston model, with no ratchet or annual ratchet. In the second Table the run times for the no-ratchet case ($\rho = -0.5$). Finally, the plot of relative error (w.r.t. BM value) for the four methods in the case $\rho = -0.5$ with no ratchet. The parameters used for this test are available in Table 3.2 and in Table 3.6.

S_0	V_0	θ	k	ω	ρ	r
100	0.22^2	0.22^2	4.75	0.55	−0.569	0.04

Table 3.8: The model parameters about Test 2B-Static.

	No ratchet					*Annual ratchet*				
	HMC	SMC	HPDE	APDE	BM	HMC	SMC	HPDE	APDE	BM
A	138.54 ± 2.70	141.86 ± 2.70	130.83	137.13	131.11	125.40 ± 2.46	128.33 ± 2.77	117.08	124.19	117.56
B	132.57 ± 1.98	137.16 ± 2.18	130.80	135.76		119.33 ± 1.80	123.68 ± 2.00	117.18	122.78	
C	131.10 ± 1.28	135.79 ± 1.32	130.80	133.85	±0.80	117.74 ± 1.17	124.49 ± 1.20	117.23	120.09	±0.71
D	130.22 ± 0.92	132.17 ± 0.90	130.82	133.02		116.98 ± 0.84	118.72 ± 0.82	117.19	119.62	

Table 3.9: Test 2B-Static. Fair fee for the Heston model, with no ratchet or annual ratchet. The parameters used for this test are available in Table 3.3 and in Table 3.8.

	ρ		*no ratchet* ($\alpha_g = 150$ *bp*)					*ratchet* ($\alpha_g = 250$ *bp*)				
			HMC	SMC	HPDE	APDE	BM	HMC	SMC	HPDE	APDE	BM
Black-Scholes Hull-White	−0.5	*A*	6055 ± 12	6060 ± 8	6058	6034	6059	7123 ± 12	7119 ± 8	7118	7128	7121
		B	6043 ± 9	6054 ± 4	6058	6050		7108 ± 9	7115 ± 4	7119	7126	
		C	6059 ± 6	6057 ± 2	6058	6052	±1	7117 ± 6	7119 ± 2	7120	7120	±1
		D	6059 ± 4	6057 ± 1	6058	6055		7120 ± 4	7119 ± 1	7120	7120	
	0	*A*	6057 ± 13	6057 ± 9	6060	6026	6059	7390 ± 13	7392 ± 9	7380	7394	7393
		B	6052 ± 10	6057 ± 5	6059	6044		7382 ± 10	7389 ± 5	7387	7391	
		C	6058 ± 6	6057 ± 2	6058	6050	±1	7389 ± 6	7392 ± 2	7390	7389	±1
		D	6059 ± 5	6058 ± 1	6058	6055		7391 ± 5	7392 ± 1	7391	7390	
	+0.5	*A*	6095 ± 13	6093 ± 9	6100	6002	6097	7647 ± 14	7651 ± 9	7636	7649	7650
		B	6101 ± 10	6097 ± 5	6098	6041		7646 ± 10	7648 ± 5	7643	7644	
		C	6098 ± 7	6096 ± 2	6097	6063	±1	7647 ± 7	7651 ± 2	7647	7644	±2
		D	6099 ± 5	6097 ± 1	6097	6080		7650 ± 5	7650 ± 1	7649	7646	

	ρ		*no ratchet* ($\alpha_g = 50$ *bp*)					*ratchet* ($\alpha_g = 100$ *bp*)				
			HMC	SMC	HPDE	APDE	BM	HMC	SMC	HPDE	APDE	BM
Heston	−0.5	*A*	7870 ± 20	7856 ± 23	7875	7867	7875	8509 ± 14	8499 ± 15	8502	8512	8509
		B	7873 ± 15	7868 ± 16	7875	7873		8506 ± 10	8503 ± 11	8506	8516	
		C	7874 ± 9	7877 ± 10	7875	7874	±1	.8505 ± 7	8511 ± 7	8507	8512	±1
		D	7888 ± 7	7880 ± 7	7875	7872		8513 ± 5	8513 ± 5	8508	8506	
	0	*A*	7803 ± 23	7181 ± 25	7797	7786	7897	8405 ± 16	8390 ± 17	8395	8400	8398
		B	7792 ± 16	7790 ± 18	7797	7794		8398 ± 12	8391 ± 13	8397	8405	
		C	7796 ± 11	7801 ± 11	7797	7797	±2	8399 ± 8	8398 ± 8	8397	8401	±2
		D	7803 ± 8	7803 ± 8	7797	7795		8402 ± 6	8403 ± 6	8398	8395	
	+0.5	*A*	7730 ± 31	7719 ± 31	7718	7699	7718	8268 ± 22	8292 ± 22	8281	8283	8282
		B	7703 ± 20	7717 ± 22	7717	7712		8292 ± 16	8281 ± 16	8282	8290	
		C	7718 ± 13	7726 ± 14	7717	7717	±3	8283 ± 10	8484 ± 10	8282	8286	±2
		D	7714 ± 9	7723 ± 11	7717	7715		8278 ± 7	8287 ± 7	8282	8279	

Table 3.10: Test 3-Static. Delta calculation for the Black-Scholes Hull-White model and the Heston model, with no ratchet or annual ratchet (these value must be multiplied by 10^{-4}). The parameters used for this test are available in Table 3.2, in Table 3.4 and in Table 3.6.

		No ratchet					Ratchet				
		HMC	SMC	HPDE	APDE	BM	HMC	SMC	HPDE	APDE	BM
BS HW	−0.5	41.89	41.75	41.75	41.85	41.84	75.96	75.91	75.92	76.20	76.00
		±0.37	±0.08			±0.12	±0.42	±0.09			±0.13
	0	76.73	76.21	76.18	76.22	76.28	143.12	143.03	142.95	143.23	143.06
		±0.48	±0.11			±0.15	±0.56	±0.12			±0.18
	+0.5	104.65	104.48	104.42	103.89	104.41	204.49	204.45	204.24	204.49	204.37
		±0.61	±0.13			±0.19	±0.73	±0.16			±0.23

		No ratchet					Ratchet				
		HMC	SMC	HPDE	APDE	BM	HMC	SMC	HPDE	APDE	BM
Heston	−0.5	34.48	34.19	33.88	33.92	33.88	56.03	56.83	55.54	55.61	55.56
		±0.43	±0.42			±0.11	±0.42	±0.41			±0.11
	0	32.03	32.23	31.90	31.94	31.88	55.89	55.18	55.80	55.87	55.85
		±0.53	±0.53			±0.14	±0.55	±0.54			±0.15
	+0.5	28.97	29.67	29.25	29.32	29.26	55.11	55.93	55.40	55.53	55.39
		±0.68	±0.71			±0.19	±0.72	±0.72			±0.20

Table 3.11: Test 4-Static. Impact of +10% mortality shocks of fair fee. The parameters used for this test are available in Table 3.2, in Table 3.4 and in Table 3.6.

3.5.2 Dynamic case

In the Dynamic case, the policy holder is supposed to choose the worst strategy from an hedger point of view, changing the value of γ_t. The PH can withdraw more ($1 \le \gamma_t \le 2$) or less ($0 \le \gamma_t \le 1$) than the standard rate (see 3.2.4.3 for more details).

In this pricing framework, we refer to the prices in Forsyth and Vetzal [21]: in their article, the authors price a GLWB contract in a Static framework, under the Black Scholes model with $r = 0.04$ and $\sigma = 0.15$. The contract parameters are reported in the Table 3.12 (Table 6.7 in [21]). They treated two cases: no ratchet, and ratchet every 3 years; both of them corresponds to *case 4* in Section 4.4.1.2 and 4.4.3.4. In the first case they got $\alpha_g = 63.1$ bp and in the second case $\alpha_g = 70.7$ bp. In Test 1 and 2 we introduce respectively stochastic interest rate and stochastic volatility to analyze the impact of these model developments on the fair guarantee fee. The parameters for interest rate and volatility models are the same as the Static case.

Here a brief summary of the numerical results for this Section.

3.5.2.1 Test 1-Dynamic: the Black-Scholes Hull-White model

Test 1-Dynamic is the Dynamic case of Test 1-Static. Model parameters are shown in Table 3.4. Results are available in Table 3.13.

In this test PDE methods proved to be much more efficient than MC ones. In fact MC ones use Longstaff-Schwartz method to find the optimal withdrawal: this method needs a lot of scenarios to approximate through the least squares approach the value of the policy for a given set of variable, and the regression is time demanding. Then, working at fixed time, we could perform fewer scenarios than Static case (around 10%), while PDE methods used almost the same parameters as in Static case. Moreover the regression problem proved to be be hard: sometimes, excluding the value $\gamma = 0$ among the possible values that the PH can chose (therefore excluding no withdrawal case), we got higher values for α_g. This means that the regression isn't very accurate, and sometimes we fail to find the optimal withdrawal: that's why, using MC methods we usually find smaller value for α_g than the right value. In particular, we excluded the value $\gamma = 0$ while using configurations A and B. We would remark that also the benchmark is affected by these computation problems and in case no-ratchet with $\rho = -0.5$ we got a smaller value for benchmark than PDE methods (around 261bp vs 266 bp). Another thing to remark is that MC methods behaved better while ratchets were considered: maybe in this case in it easier to find the best strategy. The two MC methods proved to be equivalent:

Initial age of PH	65	Gr. premium	100	DB payment	cont.ly	Strategy	Dynamic
G	0.05	Initial fees	0	Ratchet	Off/On	Bonus	5%
Withdrawal rate	1 per Y	α_m	0	Ratchet rate	every 3 Ys	$\kappa(t)$	see tab below
First withdrawal	1^{st} anniv.	Fees taken	cont.ly				

$\kappa(t)$	$0 \le t \le 1$	$1 < t \le 2$	$2 < t \le 3$	$3 < t \le 4$	$4 < t \le 5$	$t > 5$
	5%	4%	3%	2%	1%	0%

Table 3.12: The contract parameters used in the Dynamic case.

the differences in scenarios generation run-time are negligible because most of the time is spent in finding the best withdrawal. Both APDE and HPDE methods gave good and stable results, but HPDE performed better in case A.

3.5.2.2 Test 2-Dynamic: the Heston model

Test 2-Dynamic is the Dynamic case of Test 2-Static. Model parameters are shown in Table 3.6. Results are available in Table 3.14.

In this test things are similar to Test 1-Dynamic, but the optimization problem seemed to be easier than in Test 1-Dynamic: MC methods converged better, especially when using high level configurations. PDE methods behaved good as usual, and HPDE method proved to be a bit better than APDE method. The two MC methods proved to be equivalent. We note that, in the Heston model case, Dynamic strategy increases the value of α_g less than in BS HW case: probably, playing on interest rate lets the PH to gain more than playing on volatility.

3.5.2.3 Test 3-Dynamic: Hedging

Test 3-Dynamic is the Dynamic case of Test 3-Static. Results are available in Table 4.20.

In this test we got good results with all methods, but MC methods proved to be inaccurate while using configurations A and B. The range of possible values for Delta increased with regard to Test 3-Static. The two MC methods proved to be equivalent.

3.5.2.4 Test 4-Dynamic: Risk Management

Test 4-Dynamic is the Dynamic case of Test 4-Static. Results are available in Table 3.16.

In this test, we got similar results with regard to Test 4-Static: the fees reduced a little (around 20 bp in the BS HW model case and around 6 bp in the Heston model case).

In Figure 3.5.1 we present, as an example, the optimal strategy at time $t = 1$ in two different cases. We can see how it is worth to lapse when the account value reaches high values, and especially when the interest rate is high or the volatility is low. It's more difficult to understand when do no withdrawal: there must be a convenient mix of all the variables.

ρ		*no ratchet*					*ratchet*				
		HMC	SMC	HPDE	APDE	BM	HMC	SMC	HPDE	APDE	BM
-0.5	*A*	79.71 ± 9.68	66.5 ± 10.1	85.50	81.92	84.35	98.5 ± 16.1	90.24 ± 10.5	96.57	92.80	96.37
	B	77.21 ± 5.10	72.78 ± 4.52	85.34	84.78		92.31 ± 5.12	92.69 ± 5.01	96.34	95.75	
	C	79.86 ± 2.73	80.29 ± 2.60	85.27	84.86	± 0.66	94.70 ± 2.77	93.79 ± 2.67	96.25	95.78	± 0.64
	D	81.66 ± 1.58	81.58 ± 1.46	85.23	84.54		93.66 ± 1.62	94.78 ± 1.62	96.19	95.40	
0	*A*	162.6 ± 18.5	148.4 ± 13.2	172.55	167.86	169.05	182.3 ± 13.2	179.5 ± 14.3	186.44	181.96	186.53
	B	155.43 ± 7.70	156.52 ± 6.95	172.60	171.48		182.42 ± 6.16	181.35 ± 6.53	186.48	185.33	
	C	161.53 ± 4.39	164.88 ± 4.41	172.57	171.61	± 0.90	184.21 ± 3.70	183.84 ± 3.58	186.54	185.45	± 0.86
	D	164.51 ± 2.53	163.15 ± 2.23	172.58	171.15		183.87 ± 2.21	183.27 ± 2.15	186.55	185.00	
$+0.5$	*A*	256.6 ± 15.6	238.6 ± 23.3	265.33	261.22	261.29	273.9 ± 24.8	262.7 ± 24.9	272.18	268.10	274.02
	B	246.0 ± 10.4	248.5 ± 10.9	266.66	265.71		268.37 ± 8.39	269.82 ± 9.11	273.24	272.33	
	C	253.70 ± 5.96	253.33 ± 5.38	266.93	265.94	± 1.23	271.90 ± 5.51	271.60 ± 4.49	273.67	272.46	± 1.20
	D	259.00 ± 3.38	254.06 ± 3.11	267.29	265.38		272.24 ± 2.95	270.35 ± 2.86	273.99	271.94	

	HMC	SMC	HPDE	APDE
A	30 s	31 s	32 s	30 s
B	119 s	122 s	127 s	120 s
C	482 s	487 s	463 s	466 s
D	1911 s	1942 s	1732 s	1815 s

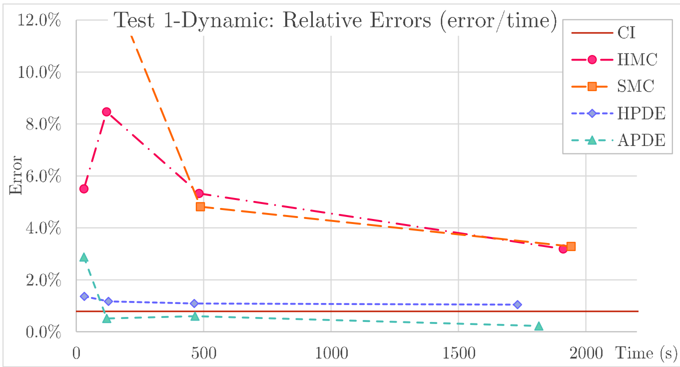

Table 3.13: Test 1-Dynamic. In the first Table, the fair fee for the Black-Scholes Hull-White model, with no ratchet or annual ratchet. In the second Table the run times for the no-ratchet case ($\rho = -0.5$). Finally, the plot of relative error (w.r.t. BM value) for the four methods in the case $\rho = -0.5$ with no ratchet. The parameters used for this test are available in Table 3.12 and in Table 3.4.

ρ		*no ratchet*					*ratchet*				
		HMC	SMC	HPDE	APDE	BM	HMC	SMC	HPDE	APDE	BM
-0.5	*A*	56.96 ± 6.01	62.56 ± 8.78	64.57	64.89	65.38	66.72 ± 8.37	72.06 ± 6.17	71.23	71.58	72.03
	B	61.68 ± 3.05	57.20 ± 3.96	64.72	64.64		72.53 ± 3.56	68.45 ± 3.36	71.37	71.30	
	C	64.68 ± 2.02	64.03 ± 2.05	64.76	64.42	±0.45	71.65 ± 1.92	71.95 ± 1.99	71.43	71.04	±0.45
	D	63.85 ± 1.28	64.67 ± 1.28	64.81	64.35		70.39 ± 1.22	71.33 ± 1.25	71.50	70.96	
0	*A*	58.83 ± 11.0	58.68 ± 19.68	61.92	61.92	62.32	65.95 ± 9.04	79.17 ± 10.36	68.97	68.98	69.54
	B	58.95 ± 4.01	54.34 ± 4.08	61.91	61.67		69.14 ± 4.53	65.30 ± 4.25	68.95	68.69	
	C	58.26 ± 2.25	57.76 ± 2.34	61.88	61.43	±0.56	68.89 ± 2.43	68.50 ± 2.44	68.94	68.41	±0.58
	D	59.16 ± 1.43	59.70 ± 1.47	61.87	61.35		66.76 ± 1.53	68.67 ± 1.57	68.93	68.33	
$+0.5$	*A*	52.66 ± 15.29	61.70 ± 13.04	57.25	57.50	56.59	63.67 ± 9.76	83.75 ± 13.01	64.60	64.79	65.42
	B	57.26 ± 5.06	51.95 ± 6.09	57.33	57.26		68.21 ± 6.30	60.31 ± 5.94	64.66	64.53	
	C	52.23 ± 3.01	51.47 ± 3.50	57.36	57.01	±0.73	63.78 ± 3.13	64.06 ± 3.25	64.68	64.25	±0.74
	D	52.60 ± 1.85	52.24 ± 1.83	57.39	56.94		62.98 ± 1.87	61.86 ± 1.97	64.71	64.16	

	HMC	SMC	HPDE	APDE
A	30 s	31 s	33 s	28 s
B	119 s	122 s	126 s	107 s
C	481 s	493 s	418 s	460 s
D	1903 s	1844 s	1690 s	1896 s

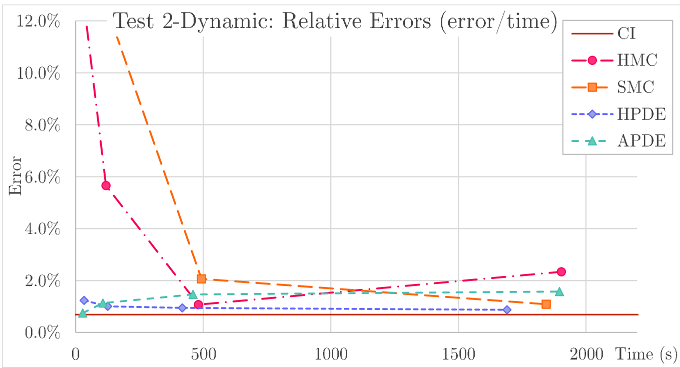

Table 3.14: Test 2-Dynamic. In the first Table, the fair fee for the Heston model, with no ratchet or annual ratchet. In the second Table the run times for the no-ratchet case ($\rho = -0.5$). Finally, the plot of relative error (w.r.t. BM value) for the four methods in the case $\rho = -0.5$ with no ratchet. The parameters used for this test are available in Table 3.12 and in Table 3.6.

	ρ		*no ratchet* ($\alpha_g = 300$ *bp*)					*ratchet* ($\alpha_g = 350$ *bp*)				
			HMC	SMC	HPDE	APDE	BM	HMC	SMC	HPDE	APDE	BM
Black-Scholes Hull-White	−0.5	*A*	8513 ± 573	7951 ± 557	8078	8097	8078	8959 ± 542	8279 ± 551	8148	8200	8157
		B	8220 ± 347	8304 ± 305	8081	8091		7922 ± 324	8104 ± 330	8151	8180	
		C	8091 ± 172	8146 ± 161	8082	8073	±42	8155 ± 173	8120 ± 162	8152	8163	±38
		D	8089 ± 108	8105 ± 104	8082	8093		8164 ± 101	8174 ± 99	8152	8164	
	0	*A*	7898 ± 550	7697 ± 489	7516	7531	7485	7733 ± 509	7379 ± 667	7538	7539	7517
		B	7685 ± 257	7389 ± 269	7517	7529		7640 ± 259	7417 ± 249	7527	7539	
		C	7488 ± 139	7443 ± 150	7517	7514	±31	7433 ± 145	7604 ± 137	7528	7530	±29
		D	7428 ± 90	7489 ± 83	7517	7518		7523 ± 87	7525 ± 84	7528	7533	
	+0.5	*A*	7444 ± 470	7612 ± 491	7333	7342	7324	7569 ± 421	7292 ± 500	7304	7314	7309
		B	7257 ± 242	7368 ± 209	7337	7350		7386 ± 207	7413 ± 192	7308	7322	
		C	7306 ± 116	7201 ± 124	7339	7336	±28	7469 ± 116	7293 ± 112	7309	7307	±26
		D	7270 ± 78	7302 ± 78	7340	7339		7291 ± 70	7298 ± 68	7310	7310	

	ρ		*no ratchet* ($\alpha_g = 75$ *bp*)					*ratchet* ($\alpha_g = 100$ *bp*)				
			HMC	SMC	HPDE	APDE	BM	HMC	SMC	HPDE	APDE	BM
Heston	−0.5	*A*	8181 ± 472	7794 ± 524	8436	8429	8432	8349 ± 293	8374 ± 324	8477	8474	8481
		B	8440 ± 250	8383 ± 223	8436	8436		8304 ± 174	8527 ± 154	8479	8480	
		C	8405 ± 87	8426 ± 87	8437	8437	±19	8535 ± 80	8499 ± 76	8479	8480	±14
		D	8437 ± 50	8472 ± 53	8437	8438		8516 ± 44	8501 ± 43	8480	8480	
	0	*A*	8756 ± 626	8304 ± 586	8329	8319	8297	8751 ± 758	8440 ± 420	8341	8332	8351
		B	8080 ± 283	8345 ± 394	8330	8327		8466 ± 251	8184 ± 217	8341	8339	
		C	8313 ± 148	8137 ± 185	8330	8329	±29	8303 ± 106	8398 ± 101	8341	8340	±18
		D	8330 ± 75	8238 ± 79	8330	8331		8343 ± 52	8283 ± 66	8341	8441	
	+0.5	*A*	7308 ± 1145	7453 ± 914	8217	8205	8242	8244 ± 736	8192 ± 522	8191	8180	8206
		B	8238 ± 623	7919 ± 416	8218	8215		8150 ± 276	8213 ± 229	8191	8189	
		C	8143 ± 145	7874 ± 241	8218	8217	±46	8216 ± 133	8123 ± 178	8192	8190	±22
		D	8144 ± 119	8131 ± 108	8218	8219		8123 ± 66	8195 ± 77	8192	8191	

Table 3.15: Test 3-Dynamic. Delta calculation for the Black-Scholes Hull-White model and the Heston model, with no ratchet or ratchet (once every 3 years); these value must be multiplied by 10^{-4}. The parameters used for this test are available in Table 3.12, in Table 3.4 and in Table 3.6.

		No ratchet					*Ratchet*				
		HMC	SMC	HPDE	APDE	BM	HMC	SMC	HPDE	APDE	BM
BS HW	−0.5	76.64	75.74	76.85	76.19	76.16	85.10	85.47	86.59	85.84	86.97
		±1.49	±1.37			±0.60	±1.53	±1.51			±0.61
	0	149.80	149.34	155.47	154.11	155.97	168.14	167.94	169.78	168.26	170.32
		±2.38	±2.14			±0.96	±2.14	±2.00			±0.96
	+0.5	236.49	236.33	242.08	240.22	242.1	248.40	246.49	250.08	248.12	250.24
		±2.84	±2.79			±1.2	±2.68	±2.48			±0.53

		No ratchet					*Ratchet*				
		HMC	SMC	HPDE	APDE	BM	HMC	SMC	HPDE	APDE	BM
Heston	−0.5	57.74	58.77	58.78	58.32	58.56	63.73	64.50	64.74	64.22	64.59
		±1.26	±1.27			±0.46	±1.23	±1.27			±0.46
	0	53.08	53.77	55.57	55.07	55.76	59.86	61.17	61.79	61.22	62.11
		±1.45	±1.46			±0.58	±1.51	±1.59			±0.59
	+0.5	46.33	46.19	50.93	50.50	50.66	55.67	54.79	57.29	56.76	59.01
		±1.87	±1.77			±0.70	±1.85	±1.94			±0.76

Table 3.16: Test 4-Dynamic. Impact of +10% mortality shocks of fair fee. The parameters used for this test are available in Table 3.12, in Table 3.4 and in Table 3.6.

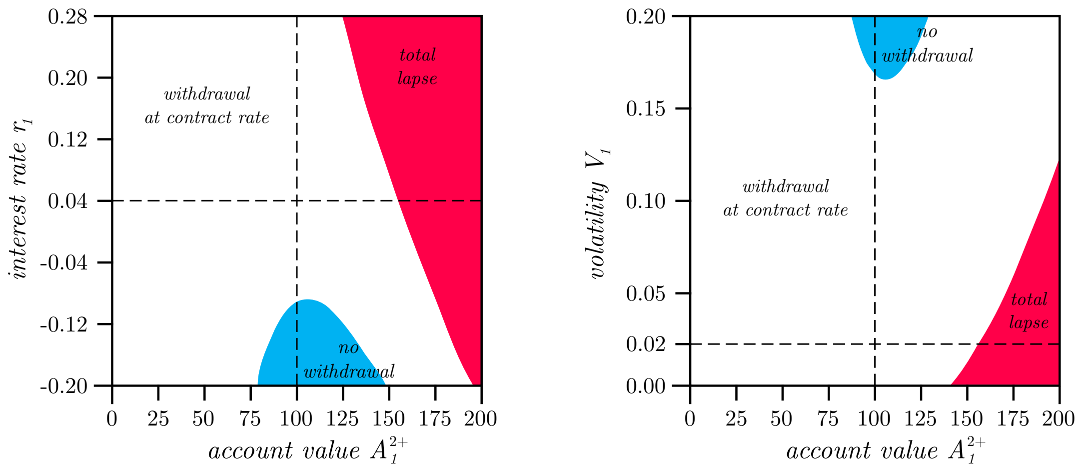

Figure 3.5.1: Optimal strategy at the first event time ($t = 1$) for the BS HW model and the Heston model, assuming $B_1^{2+} = 100$. Model parameters are available in Tables 3.4 and 3.6. Product parameters are available in Table 3.12, and $\alpha_g = 135$ bp for both cases.

3.6 Conclusions

In this Chapter we have developed four methods to price GLWB contracts under different conditions. Regarding the stochastic model, both stochastic interest rate and stochastic volatility effects have been considered. Regarding the policy holder's behavior, both Static and Dynamic strategy have been considered.

Since GLWB variable annuities are such a long maturity products, the effects of stochastic interest rate and stochastic volatility cannot be overlook. In particular, the impact of stochastic interest rate seems to be more relevant. Also Forsyth and Vetzal in [21] used a regime switching model having both stochastic interest rate and volatility, but our approach, based on SDE, is more realistic, and suitable for hedging.

All four methods gave compatible results both for pricing and delta calculation. The fair hedging fee (i.e. the cost of maintaining the replicating portfolio) is determined using a sequence of parameters refinements. The PDE methods proved to be not very expensive, while MC methods proved to be more expensive. The Hybrid PDE seemed to be the more performing than the others for its convergence speed and stability of results. Also ADI PDE behaved very well but the implementation was harder than Hybrid PDE one. In the BS HW model case, Standard MC, thanks to its exact simulation, outperformed the hybrid method while, in the Heston model case, the MC methods proved to be roughly equivalent, even if the Hybrid MC was easier to be implemented.

As we said before, PDE methods proved to be much more efficient than MC methods, especially in Dynamic case where it's much more simple to implement the optimal withdrawal choice. Similarity reduction reduces the dimension of the problem to two and therefore PDE methods perform well. Anyway, we have to remark that MC methods offer a confidence interval for results, they are useful in risk measures calculation (for example VAR or ES), and they are preferred by insurance companies because of their attachment to the concept of scenario.

A future development that could be treated is to combine stochastic interest rate and stochastic volatility: the combined model could be an element of greater realism.

We conclude by pointing out that our methods are quite flexible in that they can accommodate a wide variety of policy holder withdrawal strategies such as ones derived from utility-based models.

Chapter 4

Pricing and Hedging GMWB in the Heston and in the Black-Scholes with Stochastic Interest Rate Models

4.1 Introduction

This Chapter presents the results about the research paper [24]. We consider a Guaranteed Minimum Withdrawal Benefit (GMWB) annuity. We restrict our attention to a simplified form of a GMWB which is initiated by making a lump sum payment to an insurance company. This lump sum is then invested in risky assets, usually a mutual fund. The benefit base, or guarantee account balance, is initially set to the amount of the lump sum payment. The holder of the policy (hereinafter, we will abbreviate it with PH) is entitled to withdraw a fixed sum, even if the actual investment in the risky asset declines to zero. The withdrawal period may start immediately or later: in this case the benefit base and the account value may be reset to the maximum between their value and a fixed value. Finally, the PH may withdraw more than the contractually specified amount, including complete surrender of the contract, upon payment of a penalty. Complete surrender here means that the PH withdraws the entire amount remaining in the investment account, and the contract terminates. In most cases, this penalty for full or partial surrender declines to zero after five to seven years. During contract execution, a death benefit may come with the PH's death: in this case, his (her) heirs receive the remaining amount in the risky asset account.

The hedging costs for this guarantee are offset by deducting a proportional fee from the risky asset account. From an insurance point of view, these products are treated as financial ones: the products are hedged as if they were pure financial products, and the mortality risk is hedged using the law of large numbers. Therefore, it's very important for insurance companies to be able to price quickly these products. Moreover these products have long maturities that could last almost 25 years. The Black-Scholes model, with constant interest rate and volatility seems to be unsuitable for those products: that's why we present our pricing methods in two frameworks, modeling stochastic volatility (Heston model [26]) and stochastic interest rate (Hull-White model [28]) .

There have been several recent articles on pricing GMWBs. In particular, we would remember Chen and Forsyth [13] and Chen, Vetzal and Forsyth [14]. In the first paper, the authors used an impulse stochastic control formulation for pricing variable annuities GMWB, assuming the PH to be allowed to withdraw funds continuously, or only at anniversaries. In the second one, the authors analyzed the impact of several product and model parameters using the same PDE approach. The use of PDEs proved to be very fast and accurate, and we used it as a reference for our work.

Another research work about GMWB is Yang and Dai's one [46]: they used a flexible tree for evaluating GMWB contracts with various provisions. Yang-Dai's product is slightly different from Chen-Forsyth's one: that's why we treat the two apart.

We have made reference also to Bacinello et al. [5]: variable annuities (including GMWBs) are priced using a Monte Carlo approach. The PH's behavior is assumed to be semi-Static, i.e. the holder withdraws at the contract rate or surrenders the contract.

In this Chapter, we price two types of GMWBs guarantees and we find the no-arbitrage fee in the Heston model and the Black-Scholes with stochastic interest rate model (*BS HW model*). First, we treat a Static withdrawal strategy: the PH withdraws at the contract rate. Then, taking the point of view of the worst case for the hedger, we price the guarantees assuming that the PH follows a Dynamic withdrawal strategy. We also used these methods to calculate the Greeks for hedging and risk management. For this purpose we present four numerical methods: a hybrid tree-finite difference method and a Hybrid Monte Carlo method (both introduced by Briani et al. [10]) an ADI finite difference scheme (Haentjens and Hout [25]), and a Standard Monte Carlo method with Longstaff-Schwartz least squares regression (Longstaff and Schwartz [33]).

We use the term *no-arbitrage fee* in the sense that this is the fee which is required to maintain a replicating portfolio. A description of the replicating portfolio for these types of guarantees is given in Chen et al. [13] and Belanger et al. [6].

The main results of this Chapter are the following ones:

- We formulate the determination of the no-arbitrage fee (i.e. the cost of maintaining a replicating hedging portfolio) in the Heston model and in the BS HW model using different pricing methods;
- We present the effects of stochastic volatility and stochastic interest rate on pricing and Greeks calculation, and the sensitivity of the GMWB fee to various modeling parameters;
- We use different numerical methods to price the GMWB contracts;
- We present numerical examples which show the convergence of these methods.

The Chapter is organized as follows: in Section 2, we describe the main features of the contracts such as event times, withdrawals and penalties. In Section 3, we provide a brief review of the stochastic models used afterward. In Section 4, we present the numerical methods, and how to implement them to solve the GMWB contract pricing problem. In Section 5 we perform tests in order to show their behavior and we study the sensitivity of the no-arbitrage fee to economic and contractual assumptions. Finally, in Section 6, we present the conclusions.

4.2 The GMWB Contracts

In the following, we will refer to the contracts described in the paper of Chen and Forsyth [13] and in the paper of Yang and Dai [46]. We are calling GMWB-CF the contract described in [13] and GMWB-YD to the contract described in [46]. Now, we make a brief summary of the main features of the two contracts.

4.2.1 Mortality

Similar to the work of Chen and Forsyth [13], Milevsky and Salisbury [34] and Dai et al. [19], we will ignore mortality effects in the following. We plan to study the effects of mortality in a future work.

4.2.2 Contract State Parameters

At time $t = 0$ the policy holder pays with lump sum the premium P to the insurance company. The premium P is invested in a fund whose price is denoted by the variable S_t.

For both the two contracts, we suppose that there is a set of discrete times $\{t_i, i = 1, \ldots N\}$, which we term *event times*; at these times withdrawals may occur. We suppose $\Delta t_i = t_{i+1} - t_i$ to be constant, and denoted by Δt. We also consider $t_0 = t_1 - \Delta t$, and to be consistent with Yang Dai's notation we will write T_1 instead of t_0 ($T_1 = t_0$). Then, we write $T_1 < T_1 + \Delta t = t_1 < t_2 < \cdots < t_N = T_2$; the time lag $[T_1, T_2]$ is called *payout phase*. We remark that no withdrawals takes place in T_1.

GMWB-CF

The state parameters of the contract are:

- Account value: A_t, $A_0 = P$.
- Base benefit: B_t, $B_0 = P$.

Both these two variables are initially set equal to the premium. We define $T_1 = 0$ the time of the contract beginning, and $T_2 = t_N$ the time of the last possible withdrawal. Usually, the first withdrawal takes place in $t_1 = 1\, y$ or $t_1 = 0.5\, y$.

GMWB-YD

The state parameters of the contract are:

- Account value: A_t, $A_0 = P$.
- Guaranteed minimum withdrawal: G.

The variable A_t is initially set equal to the premium, while G is not defined until the beginning of the withdrawal period at time T_1. For this type of contract we don't need to define the Benefit Base variable because its value is deterministic until the PH decides to lapse.

For this product, there exist two time parameters, T_1 and T_2 that express the begin and the end of the payout phase. Yang and Dai used integers values for T_1 and T_2 and $\Delta t = 1\,y$ in their numerical tests. No withdrawals happen during the deferred time, i.e. for $t \in [0, T_1]$: in that period the account value evolves as explained in the next subsection (see Formula (4.2.2)). At time T_1 also the account value is reset to

$$A_{T_1^{(+)}} = \max\left[C(T_1), A_{T_1^{(-)}}\right],$$

and the value of G is fixed as

$$G = \frac{A_{T_1^{(+)}}}{m(T_2 - T_1)}, \tag{4.2.1}$$

where m denotes the number of withdrawals per year (usually $m = 1$), and $C(T_1)$ is a contract specified value that can be interpreted as the lower bound of the total guaranteed withdrawal. That value is specified as the return on the initial investment with a roll-up interest rate guaranteed interest rate i, as follows:

$$C(T_1) = P(1+i)^{T_1}.$$

If $T_1 = 0$, the reset is trivial: $A_{T_1^{(+)}} = A_0 = P$.

4.2.3 Evolution of the Contracts in the Deferred Time and between Event Times.

We call *deferred time* the time between 0 and the beginning of the payoff phase T_1: $0 \leq t < T_1$. This time set is empty unless for deferred GMWB-YD products; the other products have $T_1 = 0$ so there is no deferred time. We first consider the evolution of the value of the guarantee excluding event times t_i. Let $t \in [0, T_1[\subseteq [0, T_2]$ or $t \in]t_i, t_{i+1}[\subseteq [0, T_2]$. As we said before, S_t denotes the underlying fund driving the account value. The dynamics of S_t will be described in the next Section. The account value A_t follows the same dynamics of S_t with the exception of the fact that some fees may be subtracted continuously:

$$dA_t = \frac{A_t}{S_t} dS_t - \alpha_{tot} A_t dt. \tag{4.2.2}$$

We suppose that the total annual fees are charged to the PH and withdrawn continuously from the investment account A_t. These fees include the mutual fund management fees α_m and the fee charged to fund the guarantee (also known as the rider) α_g, so that

$$\alpha_{tot} = \alpha_m + \alpha_g.$$

The only portion used by the insurance company to hedge the contract is that coming from α_g: the other part of the fees has to be considered as an outgoing money flow as PH's withdrawals are.

4.2.4 Event Times and Final Payoff

Let G be the withdrawal guaranteed amount: for a CF product type, this parameter is a contract input, while for a YD type this value is determined at time T_1 according to formula 4.2.1.

We denote W_i the withdrawal of the PH at time t_i. As in [13], we observe that W_i is a control variable.

GMWB-CF

Usually the first event time takes place at time $t_1 = \Delta t = 1\,y$ or $t_1 = \Delta t = 0.5\,y$; moreover $t_i = i \cdot \Delta t$.

We denote with $\left(A_{t_i^{(-)}}, B_{t_i^{(-)}}, t_i\right)$ the state variables just before an event time that occurs at time t_i and with $\left(A_{t_i^{(+)}}, B_{t_i^{(+)}}, t_i\right)$ the state variables just after it.

We distinguish two pricing frameworks: at each event time the PH can withdraw according to the contract rate G (*Static approach*) or to a different rate (*Dynamic approach*). If $W_i \leq G$, then there is no penalty imposed; if $W_i > G$ there is a proportional penalty charge $\kappa\,(W_i - G)$. Anyway, the value of W_i chosen by the PH cannot exceed the guaranteed withdrawal amount $B_{t_i^{(-)}}$: it must be $W_i \in \left[0, B_{t_i^{(-)}}\right]$.

As we said before, the PH may not receive all the money he (she) withdraws from the account value. Let $f_i\,(W) : \left[0, B_{t_i^{(-)}}\right] \to \mathbb{R}$ be a function of W_i denoting the rate of cash flow received by the PH due to the withdrawal at time t_i. Then,

$$f_i\,(W_i) = \begin{cases} W_i & \text{if } W_i \leq G \\ W_i - \kappa\,(W_i - G) & \text{if } W_i > G. \end{cases}$$

The new state variables are

$$\left(A_{t_i^{(+)}}, B_{t_i^{(+)}}, t_i\right) = \left(\max\left(A_{t_i^{(-)}} - W_i, 0\right), B_{t_i^{(-)}} - W_i, t_i\right) \tag{4.2.3}$$

At time $t = T_2$ the last event time takes place: the PH withdraws as in the previous event times; then he (she) receives the final payoff which is worth

$$FP = \max\left(A_{T_2}, (1-\kappa)\,B_{T_2}\right).$$

This final payoff is applied also in the static case.

It is possible to prove that the optimal withdrawal at time T_2 is $W_N = \min\left(G, B_{T_2^{(-)}}\right)$; in this case, the value of the contract before the withdrawal is

$$\mathcal{V}\left(A_{T_2^{(-)}}, B_{T_2^{(-)}}, T_2\right) = \max\left(A_{T_2^{(-)}}, (1-\kappa)\,B_{T_2^{(-)}} + \kappa \min\left(G, B_{T_2^{(-)}}\right)\right).$$

Therefore, this remark simplifies the research of the optimal withdrawal in the Dynamic framework.

We notice that, if $A_{t_i^{(-)}} > B_{t_i^{(-)}}$ the contract can not be fully terminated in t_i: if the PH withdraws at the maximal rate, then $W_i = B_{t_i^{(-)}}$ and $\left(A_{t_i^{(+)}}, B_{t_i^{(+)}}, t_i\right) = \left(A_{t_i^{(-)}} - B_{t_i^{(-)}}, 0, t_i\right)$. In this case the PH won't be able to make any withdrawal in following event times because of $B = 0$, but he will receive the final payoff $FP = A_{T_2}$ at time T_2.

GMWB-YD

This kind of products can be deferred or not. If we set $\Delta t = {}^{(T_2 - T_1)}/_N$, then $t_i = T_1 + \Delta t \cdot i$ for $i = 1, \dots, N$. Usually $\Delta t = 1\,y$.

We denote with $\left(A_{t_i^{(-)}}, G^{(-)}, t_i\right)$ the state variables just before an event time that occurs at time t_i and with $\left(A_{t_i^{(+)}}, G^{(+)}, t_i\right)$ the state variables just after it.

According to [46], we distinguish two pricing frameworks: at each event time the PH can withdraw according to the contract rate G (*Static approach*) or fully surrender (*Dynamic approach*). In the first case, he (she) receives G at all event times after T_1 ($T_2 - T_1$ payments) and the state change is given by

$$\left(A_{t_i^{(+)}}, G^{(+)}, t_i\right) = \left(\max\left(0, A_{t_i^{(+)}} - G^{(-)}\right), G^{(-)}, t_i\right). \tag{4.2.4}$$

At time $t = T_2$, the PH receives G plus the final payoff:

$$FP = A_{T_2^{(+)}}.$$

In the second case, the PH receives G until the surrender event, and the equation (4.2.4) still holds. Let's suppose that the PH decides to surrender at time the event time t_{i^*}; then

$$\left(A_{t_{i^*}^{(+)}}, G, t_{i^*}\right) = (0, 0, t_{i^*}).$$

The final payoff is paid out at time t_{i^*}, and the contract becomes valueless:

$$FP = G + (1 - \kappa) \max\left(0, A_{t_{i^*}^{(-)}} - G\right).$$

4.2.5 Similarity Reduction

An important property of GMWB-YD contract is the fact that this contract behaves good under scaling transformations as also GLWB variable annuities do. If $\mathcal{V}(A, G, t)$ denotes the value of a contract, it is possible to prove that for any scalar $\eta > 0$

$$\eta \mathcal{V}(A, G, t) = \mathcal{V}(\eta A, \eta G, t). \tag{4.2.5}$$

Then, we just have to treat the case $G = \hat{G}$ for a fixed value $\hat{G}$ (for example $\hat{G} = P/(T_2 - T_1)$), and then, choosing $\eta = {}^{\hat{G}}/_G$, we can obtain

$$\mathcal{V}(A, G, t) = \frac{G}{\hat{G}} \mathcal{V}\left(\frac{\hat{G}}{G} A, \hat{G}, t\right),$$

which means that we can solve the pricing problem only for a single representative value of G. This effectively reduces the problem dimension.

The previous property can be applied at time T_1 when A and G are reset. Some simple calculations show that

$$\mathcal{V}\left(A_{T_1^{(+)}}, G_{T_1}^{+}, T_1\right) = \frac{A_{T_1^{(+)}}}{P}\mathcal{V}\left(P, \hat{G}, T_1\right).$$

The similarity reduction (4.2.5) was also exploited from Shah et Bertsimas in [43]. We would remark that Yang and Dai didn't use this technique for their product: therefore, their resolution of the problem of pricing is more complicated and computationally expensive.

According to GMWB-CF contracts, the similarity reduction can't be applied directly. In fact, we can prove that

$$\eta\mathcal{V}\left(A, B, G, t\right) = \mathcal{V}\left(\eta A, \eta B, \eta G, t\right) \tag{4.2.6}$$

but in this case we have to scale also the guaranteed withdrawal amount G and therefore it is not useful to reduce problem's dimension.

4.3 The stochastic models of the fund S

To understand the different impacts of stochastic volatility and stochastic interest rate over such a long maturity contract, we price the GMWB VA according to two models: the Heston model, which provides stochastic volatility, and the Black-Scholes Hull-White model, which provide stochastic interest rate. As we said before, the process S represents the underlying fund driving the account value A_t of the product.

The dynamics of the Heston model and of the Black-Scholes Hull-White Model are the same as those fixed in Section 3.3.1 and 3.3.2.

4.4 Numerical methods of pricing

In this Section we describe the four pricing methods: a Hybrid Monte Carlo method, a Standard Monte Carlo method, a Hybrid PDE method, and an ADI PDE method. These methods have already been described in Chapter 3 but in this case problem dimension may change. In fact, GLWB pricing problem has dimension equal to 2. GMWB-CF with Dynamic PH's behavior has dimension equal to 3, and all other cases 2.

We remember that our aim is to find the fair value for α_g: it's the charge that makes the initial value of the policy equal to the initial premium. To achieve this target, we price the policy (with one of the following procedures) and then we use the secant method to approach the correct value for α_g. Therefore, the main goal is to be able to find the initial value for a given value of α_g: $\mathcal{V}\left(A_0, B_0, 0\right)\left(\alpha_g\right)$.

We remark that we want to calculate the value of the policy from the point of view of the insurance company: the management fees are treated as a outgoing cash flows, and if we assume that the policy holder follows a withdrawal strategy, we consider the worst one for the insurance company.

4.4.1 The Hybrid Monte Carlo method

The value of a GLWB policy can be calculated through a Monte Carlo set of simulations. This procedure is based on two steps: generation of a scenario (a sampling of the underlying values along the life of the product), and projection of the product into the scenario. According to the way we obtain the scenarios, we distinguish two Monte Carlo models: Hybrid MC (HMC) and Standard MC (SMC).

The Hybrid MC method has been explained in Section 1.4.

4.4.1.1 Scenario generation

The generations is done according to 1.4.1.1 and 1.4.1.2

4.4.1.2 Projection

Once we have generated the scenarios set $\mathcal{S} = \{s_k, k = 1, \ldots, n_s\}$, we project the policy into all of it's scenarios: this means we calculate the initial value V_s of the contract as the sum of discounted cash flows determined according to each scenario $s \in \mathcal{S}$. Then, the initial value of the contract $\mathcal{V}$ can be approximates as the average of the initial values among all scenarios:

$$\mathcal{V} \approx \sum_{k=1}^{n_s} \frac{V_{s_k}}{n_s}.$$

This calculation depends on whether we take an optimized strategy or not.

Constant Withdrawal In this case the strategy of the PH is fixed. For a GMWB-CF product, the value of the base benefit B_{t_i} is certain: $B_{t_i^{(-)}} = P - G(i-1)$ and $B_{t_i^{(+)}} = B_{t_i^{(-)}} - G$. We can just write $V_s = V_s(A, t)$ to denote the value of the GMWB having account value equal to A at time t. This fact sets the problem dimension to 2. In this case, GMWB-CF and GMWB-YD collapse in the same product.

For each scenarios s, first we calculate the values $A_{t_i^{(+)}}$ for all t_i:

$$\begin{cases} A_0 = P \\ A_{t_i^{(-)}} = A_{t_{i-1}^{(+)}} \frac{S_{t_i}}{S_{t_{i-1}}} e^{-\alpha_{tot}\Delta t} \\ A_{t_i^{(+)}} = \max\left(0, A_{t_i^{(-)}} - G\right). \end{cases}$$

Then we set

$$V_s\left(A_{T_2^{(+)}}, T_2\right) = A_{T_2^{(+)}};$$

for all $T_1 < t_i < T_2$ we have

$$V_s\left(A_{t_i^{(+)}}, t_i\right) = e^{-\int_{t_i}^{t_{i+1}} r_s ds}\left[V_s\left(A_{t_{i+1}^{(+)}}, t_{i+1}\right) + G\right],$$

and finally

$$V_s(A_{T_1}, T_1) = e^{-\int_{T_1}^{t_1} r_s ds}\left[V_s\left(A_{t_1^{(+)}}, t_1\right) + G\right].$$

If we're price a deferred product, (i.e. $T_1 > 0$), then we set

$$G = \frac{P}{m(T_2 - T_1)}$$

and we use similarity reduction to obtain

$$V_s(A_0, 0) = e^{-\int_0^{T_1} r_s ds} V_s(P, T_1) \cdot \frac{\max\left(P, S_{T_1} e^{-\alpha_{tot} T_1}\right)}{P}.$$

Optimal Withdrawal The Optimal Withdrawal is a case of Dynamic Withdrawal and it applies only to GMWB-CF product. In their articles, Chen and Forsyth suppose the PH to be entitled to do optimal withdrawals, i.e. chose at each event time how much withdraw. In this case we suppose that at each event time t_i the PH can withdraw a fraction of the regular amount. To price in this case, we suppose that the PH chooses the value of W_i that causes the worst hedging case for the insurance company. In this case, we denote $\mathcal{V}(A, B, t)$ the expected value at time t of a generic policy whose state parameters are A, B :

$$\mathcal{V}(A, B, t) = \mathbb{E}[V_s(A, B, t)].$$

So, we suppose that the PH chooses W_i such that

$$W_i = \underset{w_i \in \left[0, B_{t_i^{(-)}}\right]}{\operatorname{argmax}} \mathcal{V}\left(\max\left(A_{t_i^{(-)}} - w_i, 0\right), B_{t_i^{(-)}} - w_i, t_i\right) + f_i(w_i).$$

This expected value can be calculated with a Longstaff-Schwartz approach:

1. Simulate N random scenarios and price the policies into these scenarios.
2. For $i = N$ to $i = 0$ (from $t_N = T_2$ to $t_0 = T_1 = 0$):
   (a) Approximate the function $\mathcal{V}(A, B, t_i)$ using the least squares projection into a space of functions (usually polynomials).
   (b) For each scenario s find the optimal withdrawal W_i (if $t_i > 0$).
   (c) Recalculate the upcoming state variables from $\tau = t_i$ to $\tau = T_2$ assuming that the PH chooses the best value for W_τ.
3. Calculate the average of the initial value $V_s(P, P, 0)$ over all the scenarios s to obtain an approximation of $\mathcal{V}(P, P, 0)$.

The search for the optimal withdrawal for this type of product is a stiff purpose. The approximation of the function $\mathcal{V}(A, B, t)$ with polynomials is hard: this is due to the fact that this function is very curved when the account value A_t is close to B_t, and is very straight otherwise.

We developed the projection algorithm in two different ways, to improve the computational time or the convergence to the right value. We call the fast algorithm "Full Regression" and the accurate one "Regression by Lines".

Full Regression In this case, the regression at each event time t_i is done using two polynomials with 3 variates: $Q_{t_i}^{up}(A,B,u)$ and $Q_{t_i}^{dw}(A,B,u)$ where u is r in the BS HW model and v in the Heston model. Here the most important remarks

- Create a grid of constant points $\mathcal{G} = \mathcal{A} \times \mathcal{B} = \{(a_k, b_h), 0 \le k \le K, 0 \le h \le H\}$ to be used as initial values to diffuse the couple (A,B) using random scenarios. This grid lets us be sure that at each event time, the set of initial values is well distributed and useful for polynomial regression. In our tests we used $\mathcal{B}$ as a set of Chebychev nodes from 0 to P, and $\mathcal{A}$ as a set of uniform nodes from 0 to $3P$. See Figure 4.4.1.
- Separate the space in two regions $\mathcal{U} = \{(a,b)\,|a \ge b\}$ and $\mathcal{D} = \{(a,b)\,|a < b\}$ and perform regression using $Q_{t_i}^{up}(A_{t_i}, B_{tt_i}, u_{t_i})$ for the first set, and $Q_{t_i}^{dw}(A_{t_i}, B_{t_i}, u_{t_i})$ for the second.
- Use shift and scaling technique to improve regression.
- As remarked before, the optimal withdrawal at last event time $t_i = T_2$, is always equal to $\min\left(G, B_{T_2^{(-)}}\right)$.
- To find the best value for the withdrawal amount W_i, numerical tests proved that, if G divides exactly P, then it's enough to search among the multiples of G.

Here a pseudo code:

```
Full_regression(){
        int ETs= T2*WD_rate;
        Scenario_generation_step();
        Forward_initial_step();
        for(int ti= ETs-1;ti>0;ti--){
                Backward_step_GMWB(ti);
                Least_Squares_step_GMWB(ti);
                Forward_Dynamic_step_GMWB(ti);
        }
        Backward_step_GMWB(0);
}
```

The functions that we used are the following ones:

- Scenario_generation_step(). Generate the scenarios: S and v or r .
- Forward_initial_step(). For all the scenarios s, chose a node (a,b) of the grid $\mathcal{G}$ (covering all the grid as s changes), and set $\left(A_{t_i}^s, B_{t_i}^s\right) = (a,b)$ for all t_i.
- Backward_step_GMWB(ti). For all the scenarios, calculate the value of the policy at the event time t_i as the sum of discounted future cash flows V_{t_i}.
- Least_Squares_step_GMWB(ti). Perform polynomial regression. Calculate $Q_{t_i}^{up}(A,B,u)$ using the value $(A_{t_i}, B_{t_i}, u_{t_i}, V_{t_i})$ such that $A_{t_i} \ge B_{t_i}$. Calculate $Q_{t_i}^{dw}(A,B,u)$ using the value $(A_{t_i}, B_{t_i}, u_{t_i}, V_{t_i})$ such that $A_{t_i} \le B_{t_i}$.

- Forward_Dynamic_step_GMWB(ti). Keeping fixed the value of A_{t_i} and B_{t_i} as stated by the Forward_initial_step function, calculate the state parameters of the policy at all the event times t_j after t_i, using at each time the best withdrawal. As we are proceeding backward and $t_j \geq t_i$, we can find the best withdrawal using the polynomial $Q^{up}_{t_j}(A, B, u)$ and $Q^{dw}_{t_j}(A, B, u)$ calculated at the previous steps.

Regression by Lines In this case, the regression at each event time t_i is done using 3 polynomials with 2 variates for each value of base benefit B and event time t_i: $Q^{up}_{t_i,B}(A, u)$, $Q^{md}_{t_i,B}(A, u)$ and $Q^{dw}_{t_i,B}(A, u)$ where u is r in the BS HW model and v in the Heston model. These polynomials are supposed to have all the same degree d. Here the most important remarks

- Create a grid of constant points $\mathcal{G} = \mathcal{A} \times \mathcal{B} = \{(a_k, b_l), 0 \leq k \leq K, 0 \leq l \leq L\}$ to be used as initial values to diffuse the couple (A, B) using random scenarios. This grid lets us be sure that at each event time, the set of initial values is well distributed and useful for polynomial regression. In our tests we used $\mathcal{B}$ as set of uniform nodes from 0 to P with $L = P/G$: $\mathcal{B} = \{0, G, 2G, \ldots, P\}$. The set $\mathcal{A}$ is more complicated. It contains points from $A_{min} = 0$ to $A_{max} = 3P$; we also tried other values for A_{max}, and $3P$ gave the best results. For each level of $B \in \mathcal{B}$, we divide the interval $[0, 3P]$ in 3 subsets: $DW_B = \left[0, \frac{1}{2}B\right]$, $MD_B = \left[\frac{1}{2}B, \frac{3}{2}B\right]$ and $UP_B = \left[\frac{3}{2}B, 3P\right]$. In each of these subsets we define $d + 1$ Chebychev nodes. These nodes defines the grid. See Figure 4.4.1.
- For each level B, the polynomials $Q^{up}_{t_i,B}(A, u)$, $Q^{md}_{t_i,B}(A, u)$ and $Q^{dw}_{t_i,B}(A, u)$ are obtained by regression, diffusing the state parameters of the policy from the nodes in the sets DW_B, MD_B and UP_B.
- As remarked before, the optimal withdrawal at last event time $t_i = T_2$, is always equal to $\min\left(G, B^-_{T_2}\right)$.
- To find the best value for the withdrawal amount W_{t_i}, numerical tests proved that, if G divides exactly P, then it's enough to search among the multiples of G. This means that when we search the best withdrawals, the possible value of B are those of $\mathcal{B}$.

Here a pseudo code:

```
Regression_by_lines(){
        int ETs= T2*WD_rate;        int H=P/G;
        Scenario_generation_step();
        for(int ti= ETs-1;ti>0;ti--){
                for(int l=0;l<L+1;l++){
                        for(sector_l= DW_l, MD_l, UP_l){
                                if(sector_l is not empty){
                                  Forward_Dynamic_step(ti,l,sector_l);
                                  Backward_step_GMWB(ti);
                                  Least_Squares_step(ti,l,sector_l);
        }}}}
        Last_Forward_Dynamic_step();
        Backward_step(0);
}
```

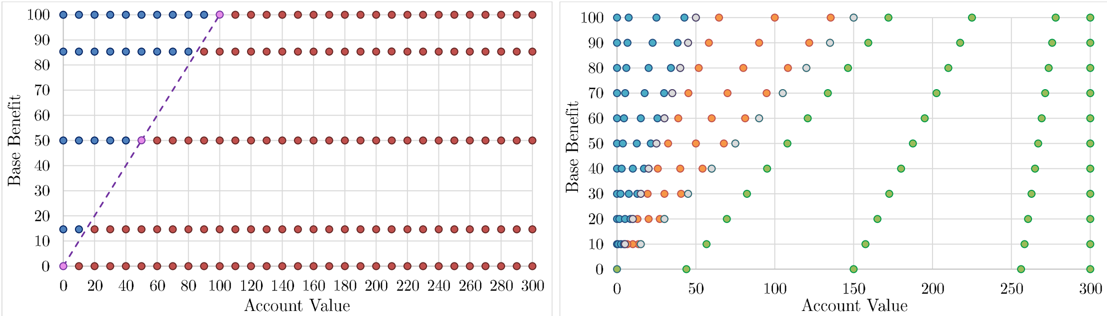

Figure 4.4.1: The grids used in the Full Regression method and in the Regression by Lines for a GMWB with $T_2 = 10$ and annual withdrawals. In the first picture, violet points are used in both the two regions. In the second picture, for each B level, the gray points border the different sectors and they are shared. The degree of the polynomials for this last case is 4.

The functions that we used are the following ones:

- Scenario_generation_step(). Generate the scenarios: S and v or r .
- Forward_Dynamic_step(ti,l,DW). For all the scenarios, setting $B_{t_i} = b_l = l \cdot P/G$, and choosing A_{t_i} in the node set $DW_{B_{t_i}}$, calculate the state parameters of the policy at all the event times t_j after t_i, using at each time t_j the best withdrawal. This functions does the same for the the sectors $MD_{B_{t_i}}$ and $UP_{B_{t_i}}$.
- Backward_step_GMWB(ti). For all the scenarios, calculate the value of the policy at event time t_i as the sum of discounted future cash flows V_{t_i}.
- Least_Squares_step(ti,l,DW). Perform polynomial regression. Calculate $Q^{dw}_{t_i,B}(A,u)$ using the value $(A_{t_i}, B_{t_i}, u_{t_i}, V_{t_i})$ diffused. This functions does the same for the the sectors $MD_{B_{t_i}}$ and $UP_{B_{t_i}}$.
- Last_Forward_Dynamic_step(). For all the scenarios, compute the state parameters of the policy, starting from $t = 0$ and $A_0 = B_0 = P$.

Optimal surrender This case concerns GMWB-YD products. In their articles, Yang and Dai suppose the PH to be entitled to surrender when optimal.

In this case we suppose that at each event time $t_i \in \{t_1, \ldots, t_N\}$ the PH can withdraw the contract amount, or fully surrender. As we did before, similarity reduction let us fix the value of G. We denote $\mathcal{V}(A,t)$ the expected value at time t of a generic policy whose state parameter is A (similarity reduction let us use only A as variable) :

$$\mathcal{V}(A,t) = \mathbb{E}\left[V_s(A,t)\right].$$

So, we suppose that the PH surrenders at time t_{i^*} if

$$(1-\kappa)\max\left(A_{t^-_{i^*}} - G, 0\right) \geq \mathcal{V}\left(\max\left(A_{t^-_{i^*}} - G, 0\right), t\right).$$

The expected value $\mathcal{V}$ can be calculated with a standard Longstaff-Schwartz approach:

1. Simulate N random scenarios and price the policy into these scenarios assuming that the PH follows a static approach.

2. For $t_i = t_N$ to $t_i = t_1$:

   (a) Approximate the function $\mathcal{V}(A, t_i)$ using the least squares projection into a space of functions (usually polynomials).

   (b) For each scenario evaluate if t_i is the good stopping time.

3. Use at time T_1 similarity reduction to include account value's reset.

4. Calculate the average of the initial value $V_s(P, 0)$ for all the scenarios to obtain an approximation of $\mathcal{V}(P, 0)$.

4.4.2 Standard Monte Carlo method

The Monte Carlo method is very similar to the Hybrid Monte Carlo one. The only different thing, is the way we produce the random scenarios. The projection phase is the same as Hybrid Monte Carlo one.

4.4.2.1 Scenario generation

We distinguish two cases for the two models.

The Heston model The generation of the scenarios (underlying and volatility) in this case has been done using a third order scheme described in Alfonsi [2].

The Black-Scholes Hull-White model The generation of the scenarios (underlying and interest rate) in this case has been done using an exact scheme described in Ostrovski [39], with a few changes in order to incorporate the correlation between underlying and interest rate.

4.4.3 PDE Hybrid Method

The Hybrid PDE approach is different from the previous ones. In fact it's a PDE pricing method and it's based on Briani et al. [10], [11] both for Heston and Hull-White case. Using a tree to diffuse volatility or interest rate, we freeze these values between two tree-levels and we solve a Black Scholes PDE for each node of the tree, using as initial data a weighted mix of the data of the upcoming nodes.

We can resume the pricing methods in three features: model, algorithm structure and pricing.

We start describing the model between the event times.

4.4.3.1 The Heston Model

Starting from the model for the found S_t in (3.3.1), we call $\bar{\rho} = \sqrt{1-\rho^2}$ and we write $Z_t^A = \rho Z_t^v + \bar{\rho}\bar{Z}_t^A$, where $\bar{Z}^A$ is a Brownian motion uncorrelated with Z^v. Then,

$$\begin{cases} dA_t = (r - \alpha_{tot})\, A_t dt + \sqrt{v_t} A_t \left(\rho dZ_t^v + \bar{\rho} d\bar{Z}_t^A\right) & v_0 = \bar{v}_0, \\ dv_t = k(\theta - v_t)\, dt + \omega\sqrt{v_t} dZ_t^V & A_0 = \bar{A}_0, \end{cases} \quad d\left\langle \bar{Z}_t^A, Z_t^v \right\rangle = 0,$$

covering the behavior of A_t between two event times, we define the process

$$Y_t^E = \ln(A_t) - \frac{\rho}{\omega} v_t,\ Y_0^E = \ln(A_0) - \frac{\rho}{\omega} v_0$$

Then,

$$A_t = \exp\left(Y_t^E + \frac{\rho}{\omega} v_t\right) \tag{4.4.1}$$

and

$$dY_t^E = \left(r - \alpha_{tot} - \frac{v_t}{2} - \frac{\rho}{\omega} k(\theta - v_t)\right) dt + \bar{\rho}\sqrt{v_t} d\bar{Z}_t^A.$$

This process Y^E is important because it's a process uncorrelated with the volatility process v, and we introduce it as in [10]. We are going to use it to define a PDE to be solved along the tree.

We define $\hat{\mathcal{V}}^{He}\left(t, Y_t^E\right) = \mathcal{V}(t, A_t)$.

If, in a small time lag $[\tau, \tau + \Delta\tau]$, we approximate the process Y_t^E by the process $\bar{Y}_t^E$ whose dynamics is given by

$$d\bar{Y}_t^E = \left(r - \alpha_{tot} - \frac{v_\tau}{2} - \frac{\rho}{\omega} k(\theta - v_\tau)\right) dt + \bar{\rho}\sqrt{v_\tau} d\bar{Z}_t^A.$$

Then, $\hat{\mathcal{V}}^{He}\left(t, \bar{Y}_t^E\right)$ verifies the following PDE

$$\frac{\partial \hat{\mathcal{V}}^{He}}{\partial t} + \left(r - \alpha_{tot} - \frac{v_\tau}{2} - \frac{\rho}{\omega} k(\theta - v_\tau)\right) \frac{\partial \hat{\mathcal{V}}^{He}}{\partial \bar{Y}_t^E} + \frac{\bar{\rho}^2 v_\tau}{2} \frac{\partial^2 \hat{\mathcal{V}}^{He}}{\partial^2 \bar{Y}_t^E} - r\hat{\mathcal{V}}^{He} = 0. \tag{4.4.2}$$

4.4.3.2 The Black-Scholes Hull-White Model

Starting from the model for the found S_t in (3.3.2), we call $\bar{\rho} = \sqrt{1-\rho^2}$ and we write $Z_t^A = \rho Z_t^r + \bar{\rho}\bar{Z}_t^A$, where $\bar{Z}^A$ is a Brownian motion uncorrelated with Z^r. Then,

$$\begin{cases} dA_t = A_t (r - \alpha_{tot})\, dt + \sigma A_t \left(\rho dZ_t^r + \bar{\rho} d\bar{Z}_t^A\right) & A_0 = \bar{A}_0, \\ dX_t = -kX_t dt + dZ_t^r & X_0 = 0, \\ r_t = \omega X_t + \beta(t), & \end{cases} \quad d\left\langle \bar{Z}_t^A, Z_t^r \right\rangle = 0.$$

We define the process

$$Y_t^U = \ln(A_t) - \rho\sigma X_t,\ Y_0^U = \ln(A_0)$$

Then,

$$A_t = \exp(Y_t + \rho\sigma X_t) \tag{4.4.3}$$

and
$$dY_t^U = \left(r_t - \alpha_{tot} - \frac{\sigma^2}{2} + \sigma\rho k X_t \right) dt + \sigma\bar{\rho} d\bar{Z}_t^A.$$

This process Y^U is important because it's a process uncorrelated with the mean-reverting process X, and we introduce it as in [10]. We are going to use it to define a PDE to be solved along the tree.

We define $\hat{\mathcal{V}}^{HW}\left(t, Y_t^U\right) = \mathcal{V}\left(t, A_t\right)$. If, in a small time lag $[\tau, \tau + \Delta\tau]$, we approximate the process Y_t^U by the process $\bar{Y}_t^U$ whose dynamics is given by

$$d\bar{Y}_t^U = \left(r_\tau - \alpha_{tot} - \frac{\sigma^2}{2} + \sigma\rho k X_\tau \right) dt + \sigma\bar{\rho} dZ_t,$$

and the interest rate process by r_τ, then, $\hat{\mathcal{V}}^{HW}\left(t, \bar{Y}_t^U\right)$ verifies the following PDE

$$\frac{\partial\hat{\mathcal{V}}^{HW}}{\partial t} + \left(r_\tau - \alpha_{tot} - \frac{\sigma^2}{2} + \sigma\rho k X_\tau \right) \frac{\partial\hat{\mathcal{V}}^{HW}}{\partial\bar{Y}_t^U} + \frac{\bar{\rho}^2\sigma^2}{2} \frac{\partial^2\hat{\mathcal{V}}^{HW}}{\partial^2\bar{Y}_t^U} - r_\tau\hat{\mathcal{V}}^{HW} = 0. \tag{4.4.4}$$

4.4.3.3 Algorithm structure

The structures for this algorithm consist in a tree and a PDE solver. As described in Briani et al. [10], [11], we use a tree to diffuse the volatility (or the interest rate) along the life of the product, and we solve backward a 1D PDE freezing at each node of the tree the volatility (or the interest rate). The tree is built according to Section 1.2.4 (quadrinomial tree, matching the first three moments of the process), and the PDE is solved with a finite difference approach. We have to solve the PDE between event times, and at each event time we apply the changes to the states to reproduce the effects of the events.

We remark that we solve the PDEs doing a single time step that requires only a linear complexity because we have to solve a linear system with tridiagonal matrix. The computational cost is low as observed in [10] and [11]. We observe that X_t and V_t processes are mean reverting. Thanks to the way the trees are built, there are many nodes in the trees that cannot be visited by the approximating Markov chain. Therefore their probability $p_{n,j}$ to be visited is worth 0 and they have no impact on the values at the root of the tree. There is no reason to do any operation for those nodes. So, to save time, we do the standard step (mix up the vectors according to the transition probabilities and solve backward a PDE) only for those nodes having $p_{n,j} > 0$. This curtailing technique reduces the computational time, and the convergence of the method is preserved. A similar approach is used in [3].

4.4.3.4 Pricing

We distinguish 3 cases.

Static case

This case is common to both GMWB-CF and GMWB-YD products. Problem's dimension is 2: about GMWB-CF, at each event time the value of the the base benefit B_{t_i} is equal to $P - G \cdot i$

and thus it's not a problem's variable; about GMWB-YD similarity reduction reduce problem's dimension to 2.

For each node of the tree we have to solve one PDE using the mixture of the the data of the upcoming nodes: the mixture is done according to transition probabilities. The PDE to be solved are those in (4.4.2) and (4.4.4) where $[\tau, \tau + \Delta\tau]$ denotes the time lag between two tree's node.

The variables $\bar{r}$, $\bar{X}$ and $\bar{v}$ will denote the frozen values of r_t, X_t and v_t using the data of the actual node. We used a finite differences approach using equally spaced nodes for Y_t processes. To reduce the run time, we do this only for most relevant nodes: this cutting technique dramatically reduced calculation times without compromising the quality of results. Then, using the inverse transformations (4.4.1) and (4.4.3), we apply the event times actions in (4.2.3) or equivalently (4.2.4).

Optimal Withdrawal

This case is about GMWB-CF products. This is the hardest to be treated because the problem's dimension is 3. We solve the same PDE as in Static case, but this time we have to solve them for different values of B_t and chose the best withdrawal W_t at each event time. Numerical test showed that it's enough to search the best withdrawal between multiples of G equal or smaller than the base benefit. Then we decided to solve the problem for all B_t values of which are multiples of G and are smaller than the initial premium P: $B = 0$, $B = G$, $B_t = 2G$, $\ldots B = nG = P$. Then, we solve n 2-dimensional problems rather then one 3-dimensional problem. This approach is similar to "Regression by Lines" defined for MC methods.

Best withdrawal's search is performed searching among permissible withdrawals which are multiples of G: $W = 0$, $W = G$,$\ldots W = mG = B$. The estimate of $\mathcal{V}(A, B)$ for those values of A that aren't on the grid, is done using splines.

In Figure 4.4.2 we can see a scheme that represents what happens for a product with $G = 20$: for example, a 5 years maturity GM with annual withdrawal rate and $P = 100$ ($G = 20$). Nodes are exponentially distributed (uniformly for Y process) and for each B value, we add a node that represents $A = 0$ (blue nodes). For each node, first we mix the data vectors of the upcoming nodes according to transition probabilities. Then we solve a PDE backward starting from the mixture of the data. Then we apply withdrawal step: for each node we consider admissible withdrawals of the type $W = kG$ and we chose the value that maximize PH's benefit: cash flow plus policy's value. This research is shown in the Figure (see yellow nodes that corresponds to possible withdrawals).

Optimal surrender

This case is about GMWB-YD products. It's much more simpler than optimal withdrawal. In fact, the PH can only chose between withdrawal at the contract rate and fully surrender.

Withdrawal step at event time t_i consists into replacing $\mathcal{V}(A, t_i)$ by

$$\max\left[G + \mathcal{V}\left(\max\left(A - G, 0\right), t_i\right); G + (1 - \kappa)(A - G)\right].$$

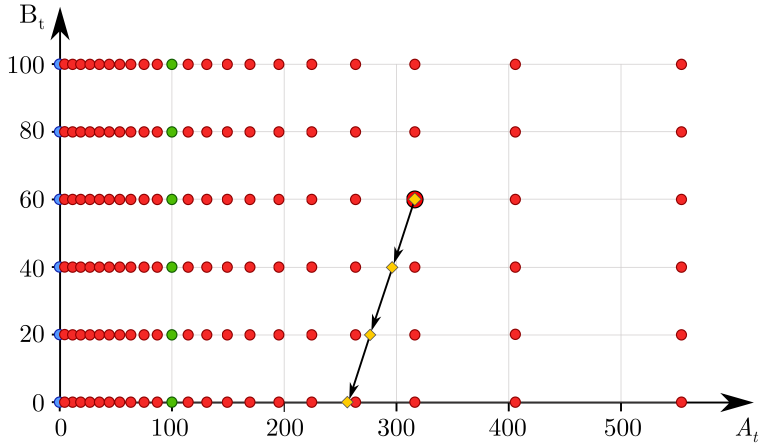

Figure 4.4.2: The scheme for PDE optimal withdrawal for a given node.

4.4.4 PDE ADI Method

We propose a PDE pricing method with alternative direction implicit scheme which has been already successfully used for European financial product (see [25]) and for an insurance GLWB product (see [23]). This method permits to treat the Heston model and the Black-Scholes Hull-White model. This method is fast and accurate. Moreover it is easy to take into account the similarity reduction and the optimal behavior. For this method, we followed the same principles of HPDE method about taking in account the event times.

The PDEs to be solved are

$$\mathcal{V}_t^{He} + \frac{VA^2}{2}\mathcal{V}_{AA}^{He} + \frac{\omega^2 V}{2}\mathcal{V}_{VV}^{He} + (r - \alpha_{tot})\, A\mathcal{V}_A^{He} + \rho\omega AV\mathcal{V}_{AV}^{He} + k\,(\theta - V)\,\mathcal{V}_V^{He} - r\mathcal{V}^{He} = 0 \qquad \text{(He 2b)}$$

$$\mathcal{V}_t^{HW} + \frac{\sigma^2 A^2}{2}\mathcal{V}_{AA}^{HW} + \frac{\omega^2}{2}\mathcal{V}_{rr}^{HW} + (r - \alpha_{tot})\, A\mathcal{V}_A^{HW} + \rho\omega A\sigma\mathcal{V}_{Ar}^{HW} + k\,(\theta_t - r)\,\mathcal{V}_r^{HW} - r\mathcal{V}^{HW} = 0 \qquad \text{(HW 2b)}$$

There are multiple numerical parameters which have to be carefully chosen. We have to choose the grids for the benefit base, the account value, the rate in the Hull-White model and the volatility in the Heston model. We have chosen to use the meshes described in [25] with the parameters

$$A_{left} = 0.8S_0 \quad A_{right} = 1.2S_0 \quad A_{max} = 1000 \cdot T2 \cdot S_0 \quad \text{and } d_1 = S_0/20,$$

for the mesh of variable A,

$$R_{\max} = 0.8, \quad c = R_0 \quad \text{and } d_2 = R_{\max}/400$$

for the mesh of variable r in the Black-Scholes Hull-White model, and

$$V_{\max} = MIN(MAX(100V_0, 1), 5) \quad \text{and } d_3 = V_{\max}/500.$$

for the mesh of variable V in the Heston model. Some grid are uniform, or based on hyperbolic grid. Moreover the boundary conditions are completely unknown, and an asymptotic study would be necessary to chose them. We have chosen homogeneous Neumann boundary conditions, and we have chosen very large grids to avoid that this choice impacts the results. We have only used the Douglas scheme, but other schemes are possible to have better order of convergence in time. Thus many possibilities are possible to improve the ADI scheme, but the easier is already enough to obtain good results.

4.5 Numerical results

In this Section we compare the numerical methods used in Section 4.4: Hybrid Monte Carlo (*HMC*), Standard Monte Carlo (*SMC*), Hybrid PDE (*HPDE*), and ADI PDE (*APDE*). In particular we compare pricing and Greeks computation in *Static Case* and *Dynamic Case* for both the two product types.

We chose the parameters of the methods according to 4 configurations (A, B, C, D), with an increasing number of steps and so that the calculation time for the various methods in each configuration were close. The 4 configurations are in Table 4.1 and in Table 4.2 with the notation (time steps per year × space steps × vol steps) for the ADI PDE method, (time steps per year × space steps) for the Hybrid PDE method approaches and (time steps per year × number of simulations) for the MC ones. In Monte Carlo for Dynamic case, we also add the degree of the approximating polynomial. These values have been chosen to achieve approximately these run times: (A) 30 s, (B) 120 s, (C) 480 s, (D) 1920 s. To reduce the run time we do the secant iterations using an increasing number of time steps for all the methods: the values in Table 4.1 are those used for the last 3 iterations.

We use the Standard MC both as a pricing method, both as a benchmark (BM). About the benchmark, in the Static case we used 10^8 independent runs. In the Dynamic case we used 10^6 independent runs; in each sub runs the expected value has been approximated by a 4 order polynomial.

The search for the fair α_g value has been driven by the secant method. The initial values for this method were $\alpha_g = 0$ bp and $\alpha_g = 200$ bp.

To achieve Delta calculation in Monte Carlo methods we used a 1‰ shock in Static case and 1% in Dynamic case.

4.5.1 Static Withdrawal for GMWB-CF

In the Static Withdrawal case we suppose the PH to withdrawal exactly at the guaranteed rate.

The Static Tests 1 and 2 are inspired by [13]: in their article, Chen and Forsyth price a GMWB contract according to an optimal withdrawal framework, under the Black Scholes model. First we priced their product for different maturities and withdrawal rates, assuming Static withdrawals in Black and Scholes model to get a reference price in this model; we got the α value using both a standard Monte Carlo method and a standard PDE method. As we easily got the correct values for the simple Black-Scholes model, then we add stochastic interest

	BS HW Static				Heston Static			
	HMC	SMC	HPDE	APDE	HMC	SMC	HPDE	APDE
A	$4\times 9.2\cdot 10^5$	$1\times 1.7\cdot 10^6$	260×250	$25\times 250\times 505$	$4\times 5.8\cdot 10^5$	$4\times 5.2\cdot 10^5$	270×250	$25\times 250\times 505$
B	$8\times 1.8\cdot 10^6$	$1\times 5.7\cdot 10^6$	420×500	$40\times 400\times 85$	$8\times 1.2\cdot 10^6$	$8\times 1.2\cdot 10^6$	520×500	$40\times 400\times 80$
C	$12\times 6.3\cdot 10^6$	$1\times 2.9\cdot 10^7$	780×1000	$60\times 620\times 125$	$12\times 3.9\cdot 10^6$	$12\times 3.4\cdot 10^6$	850×1000	$60\times 620\times 120$
D	$16\times 1.9\cdot 10^7$	$1\times 1.2\cdot 10^8$	1200×2000	$100\times 10^3\times 215$	$16\times 1.2\cdot 10^7$	$16\times 1.1\cdot 10^7$	1400×2000	$100\times 10^3\times 200$
	BS HW Dynamic				**Heston Dynamic**			
	HMC	SMC	HPDE	APDE	HMC	SMC	HPDE	APDE
A	$4\times 6.0\cdot 10^4\times 1$	$1\times 6.5\cdot 10^4\times 1$	70×250	$8\times 95\times 30$	$4\times 5.1\cdot 10^4\times 1$	$4\times 5.5\cdot 10^4\times 1$	88×250	$10\times 125\times 25$
B	$8\times 8.7\cdot 10^4\times 2$	$1\times 9.5\cdot 10^4\times 2$	160×500	$14\times 150\times 48$	$8\times 1.2\cdot 10^4\times 2$	$8\times 1.3\cdot 10^4\times 2$	160×500	$15\times 200\times 40$
C	$12\times 1.8\cdot 10^5\times 3$	$1\times 1.9\cdot 10^5\times 3$	270×1000	$22\times 250\times 75$	$12\times 2.3\cdot 10^5\times 3$	$12\times 2.5\cdot 10^5\times 3$	266×1000	$25\times 320\times 60$
D	$16\times 3.5\cdot 10^5\times 4$	$1\times 3.5\cdot 10^5\times 4$	360×2000	$35\times 400\times 120$	$16\times 4.2\cdot 10^5\times 4$	$16\times 5.0\cdot 10^5\times 4$	350×2000	$40\times 500\times 90$

Table 4.1: Configuration parameters for the BS HW model and for the Heston model, Static and Dynamic for the GMWB-CF product with $T_2 = 10$ and $WF = 1$.

	BS HW Static				Heston Static			
	HMC	SMC	HPDE	APDE	HMC	SMC	HPDE	APDE
A	$4\times 3.2\cdot 10^5$	$1\times 6.0\cdot 10^5$	130×250	$10\times 245\times 50$	$4\times 2.3\cdot 10^5$	$4\times 2.0\cdot 10^5$	120×250	$10\times 250\times 50$
B	$8\times 6.4\cdot 10^5$	$1\times 2.3\cdot 10^6$	215×500	$15\times 375\times 80$	$8\times 4.6\cdot 10^5$	$8\times 3.8\cdot 10^5$	220×500	$15\times 380\times 80$
C	$12\times 2.2\cdot 10^6$	$1\times 1.2\cdot 10^7$	415×1000	$35\times 520\times 110$	$12\times 1.6\cdot 10^6$	$12\times 1.3\cdot 10^6$	425×1000	$36\times 530\times 110$
D	$16\times 6.8\cdot 10^6$	$1\times 4.5\cdot 10^7$	480×2000	$55\times 880\times 180$	$16\times 4.8\cdot 10^6$	$16\times 4.0\cdot 10^6$	480×2000	$55\times 890\times 180$
	BS HW Dynamic				**Heston Dynamic**			
	HMC	SMC	HPDE	APDE	HMC	SMC	HPDE	APDE
A	$4\times 6.8\cdot 10^4\times 1$	$1\times 8.1\cdot 10^4\times 1$	130×250	$10\times 245\times 50$	$4\times 5.5\cdot 10^4\times 2$	$4\times 5.8\cdot 10^4\times 2$	120×250	$10\times 250\times 50$
B	$8\times 2.5\cdot 10^5\times 2$	$1\times 3.4\cdot 10^5\times 2$	215×500	$15\times 375\times 80$	$8\times 2.2\cdot 10^5\times 3$	$8\times 2.0\cdot 10^5\times 3$	220×500	$15\times 380\times 80$
C	$12\times 6.9\cdot 10^5\times 3$	$1\times 9.7\cdot 10^5\times 3$	415×1000	$35\times 520\times 110$	$12\times 5.9\cdot 10^5\times 4$	$12\times 5.6\cdot 10^5\times 4$	425×1000	$36\times 530\times 110$
D	$16\times 1.8\cdot 10^6\times 4$	$1\times 1.8\cdot 10^6\times 4$	480×2000	$55\times 880\times 180$	$16\times 1.5\cdot 10^6\times 5$	$16\times 1.5\cdot 10^6\times 5$	480×2000	$55\times 890\times 180$

Table 4.2: Configuration parameters for the BS HW model and for the Heston model, Static and Dynamic for the GMWB-YD product with $(T_1, T_2) = (10, 25)$

rate and stochastic volatility. Model parameters are available in Table 4.3, and the values of α_g that we got in the Black-Scholes case are given in Table 4.4.

4.5.1.1 Test 1: Static GMWB-CF in the Black-Scholes Hull-White Model

In this test we want to price a GMWB-CF product according to BS HW model. We use the same corresponding parameters as in the Black Scholes model. Model parameters are shown in Table 4.5. Results are available in Table 4.6.

All the four methods behaved well and in the configuration D they gave results consistent with the benchmark. HPDE proved to be the best: all of its configurations gave results consistent with the benchmark. Then APDE and SMC, and HMC gave good results too. SMC performed a little better than HMC: the first method simulates the underlying value and the interest rate exactly and so it is enough to simulate the values at each event time. HMC matches the first three moments of the BS HW r process, but doesn't reproduce exactly its law: therefore it is right to increase the number of steps per year. So, for a given run time, we can simulate less scenarios in HMC than SMC: effectively, the confidence interval of HMC is larger than SMC one. Moreover, SMC over performed the benchmark when using configuration D. The two PDE methods returned stable results, and they often converged in a monotone way.

With regard to the numerical results, we observe that the α_g values decrease with increasing maturity, just as in the Black-Scholes case, and increase a little, with increasing withdrawal frequency.

4.5.1.2 Test 2: Static GMWB-CF in the Heston Model

In this test we want to price a GMWB-CF product according to the Heston model. Model parameters are shown in the Table 4.7. Results are shown in Table 4.8.

In this Test, MC methods had more difficulties; all the values of PDE methods were close to the benchmark, while some values from MC methods were less accurate, but compatibles with the benchmark (the value of BM is inside the MC confidence interval). If we compare the two MC approaches, in this case they are equivalent: HMC proved to be faster than SMC when using few time steps (we could exploit +11% simulations in configuration A), while SMC proved to be slightly faster in high time steps simulations, because of more time needed to build the volatility tree (−8% simulations in configuration D). HPDE shows to be very stable (case $T_2 = 10$, $WF = 2$, α_g didn't change through configurations B-D), APDE behaved well to (often monotone convergence).

With regard to the numerical results, we observe that the α_g values decrease with increasing maturity, just as in the Black-Scholes case, and increase a little, with increasing withdrawal frequency.

4.5.1.3 Test 3: Hedging for Static GMWB-CF

To reduce financial risks, insurance companies have to hedge the sold VA: to accomplish this target they must calculate the Greeks of products.

In this test we want to show how the different methods can be used to calculate the main Greeks. This can be done through finite differences for small shocks on the variables. In general, the PDE methods are ahead w.r.t. MC methods: the price is computed through finite differences and so the price under shock is already computed. For MC methods this is quite harder because the pricing has to be repeated changing the inputs.

To start, we calculate the underlying greek Delta, for the products of Test 1 and Test 2. As in this case we don't want to compute the fair fee α_g, we fix it arbitrarily: see Table 4.9 and Table 4.11. The values chosen are such as to cover the costs of the insurer, and may be plausible on a real case. Results are available in Table 4.10 (all values in Table must be multiplied by 10^{-4}).

In this Test, we got very accurate results with all method. Anyway, HPDE and APDE proved to be the best: they both gave stable and accurate results; in this Test, the two PDE methods were equivalent. We remark that despite fair fee changes a lot when changing the maturity of the policy, the value of Delta changes much less. Delta calculation proved to be slightly harder in the Heston model case than in the BS HW model case: see confidence intervals.

Expiry Time T	5, 10, 20 Years	GMW G	$100.0/(T \cdot WF)$
Withdrawal Frequency WF	1 or 2 per Year	Initial Premium	100.0
Initial account value A_0	100.0	S_0	100.0
Initial base b. value B_0	100.0	r	0.05
Withdrawal penalty κ	0.10	σ	0.20
Management fees α_m	0		

Table 4.3: Parameters used by Chen and Forsyth in [13].

T_2	$WF = 1$		$WF = 2$	
	PDE	MC	PDE	MC
5	235.24	235.11 ± 0.42	243.96	243.80 ± 0.42
10	92.41	92.28 ± 0.30	94.62	94.84 ± 0.30
20	27.64	27.79 ± 0.24	28.09	28.39 ± 0.24

Table 4.4: Fair bp values of α_g in Black Scholes model, for Static GMWB-CF with the same parameters as in [13].

S_0	r	*curve*	k	ω	ρ	σ
100	0.05	*flat*	1.0	0.2	−0.5	0.20

Table 4.5: The model parameters about Test 1

T_2		$WF = 1$					$WF = 2$				
		HMC	SMC	HPDE	APDE	BM	HMC	SMC	HPDE	APDE	BM
5	A	191.05 ± 0.83	190.81 ± 0.66	191.03	191.16	191.34	196.87 ± 0.88	196.79 ± 0.94	196.39	196.50	196.77
	B	191.79 ± 0.59	191.25 ± 0.33	191.18	191.58		196.88 ± 0.62	197.02 ± 0.47	196.55	196.60	
	C	191.34 ± 0.32	191.26 ± 0.15	191.25	191.47		196.72 ± 0.34	196.64 ± 0.21	196.62	196.88	
	D	191.20 ± 0.18	191.25 ± 0.07	191.27	191.38	±0.11	196.55 ± 0.19	196.67 ± 0.10	196.65	196.68	±0.11
10	A	79.71 ± 0.84	79.26 ± 0.66	79.41	79.33	79.44	81.38 ± 0.88	81.13 ± 0.94	80.98	81.32	80.97
	B	79.80 ± 0.60	79.43 ± 0.33	79.39	79.41		80.95 ± 0.63	81.42 ± 0.47	80.98	80.64	
	C	79.61 ± 0.32	79.56 ± 0.15	79.39	79.41		81.12 ± 0.34	81.12 ± 0.21	80.98	81.01	
	D	79.35 ± 0.18	79.44 ± 0.07	79.38	79.40	±0.08	80.90 ± 0.19	80.98 ± 0.10	80.99	81.01	±0.08
20	A	26.33 ± 0.98	25.04 ± 0.77	24.90	24.90	24.81	25.04 ± 1.07	25.72 ± 1.06	25.27	25.20	25.16
	B	25.92 ± 0.69	25.23 ± 0.39	24.86	24.67		25.91 ± 0.75	25.59 ± 0.54	25.23	25.16	
	C	25.16 ± 0.37	24.91 ± 0.17	24.84	24.81		25.16 ± 0.41	25.35 ± 0.24	25.21	25.18	
	D	24.99 ± 0.21	24.81 ± 0.09	24.84	24.82	±0.07	25.40 ± 0.23	25.13 ± 0.12	25.20	25.18	±0.07

	HMC	SMC	HPDE	APDE
A	31 s	30 s	30 s	30 s
B	121 s	121 s	120 s	118 s
C	482 s	484 s	464 s	481 s
D	1920 s	1899 s	1893 s	1909 s

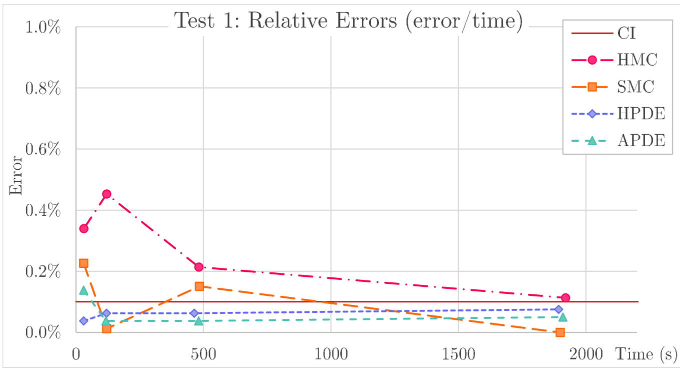

Table 4.6: Test 1. In the first Table, the fair fee α_g in bp for the Black-Scholes Hull-White model, with annual or six-monthly withdrawal. In the second Table the run times for the case $WF = 1$, $T_2 = 10$. Finally, the plot of relative error (w.r.t. BM value) for the four methods in the same case of run-times Table. The parameters used for this test are available in Table 4.3 and in Table 4.5.

S_0	v_0	θ	k	ω	ρ	r
100	0.20^2	0.20^2	1.0	0.2	-0.5	0.05

Table 4.7: The model parameters about Test 2.

T_2		$WF = 1$					$WF = 2$				
		HMC	SMC	HPDE	APDE	BM	HMC	SMC	HPDE	APDE	BM
5	A	232.30 ± 0.97	232.83 ± 1.03	231.20	231.62	231.38	239.82 ± 1.00	240.88 ± 1.05	239.34	239.51	239.54
	B	231.72 ± 0.68	231.73 ± 0.75	231.37	231.55		240.41 ± 0.71	239.76 ± 0.77	239.52	239.38	
	C	231.39 ± 0.37	231.56 ± 0.41	231.43	231.48	±0.10	239.84 ± 0.38	239.76 ± 0.41	239.59	239.72	±0.11
	D	231.42 ± 0.21	231.56 ± 0.23	231.45	231.47		239.48 ± 0.20	239.71 ± 0.23	239.61	239.64	
10	A	97.13 ± 1.10	97.55 ± 1.07	95.86	95.91	95.81	98.29 ± 1.05	99.50 ± 1.10	97.99	98.25	97.98
	B	96.23 ± 0.73	97.07 ± 0.78	95.86	95.88		98.58 ± 0.74	98.62 ± 0.78	98.01	98.78	
	C	95.65 ± 0.39	95.81 ± 0.42	95.87	95.89	±0.08	98.12 ± 0.40	97.78 ± 0.43	98.01	98.01	±0.09
	D	95.88 ± 0.23	95.84 ± 0.24	95.87	95.86		97.93 ± 0.23	97.95 ± 0.24	98.01	98.00	
20	A	31.84 ± 1.17	31.84 ± 1.23	30.71	30.68	30.57	31.39 ± 1.20	32.06 ± 1.27	31.18	31.26	31.05
	B	31.42 ± 0.84	31.78 ± 0.90	30.64	30.60		31.69 ± 0.85	31.12 ± 0.90	31.11	31.13	
	C	30.53 ± 0.45	30.99 ± 0.47	30.63	30.63	±0.06	31.02 ± 0.45	31.04 ± 0.50	31.10	31.09	±0.07
	D	30.73 ± 0.26	30.61 ± 0.27	30.63	30.63		31.47 ± 0.26	31.10 ± 0.27	31.09	31.08	

	HMC	SMC	HPDE	APDE
A	30 s	30 s	30 s	31 s
B	122 s	118 s	121 s	120 s
C	486 s	477 s	483 s	479 s
D	1951 s	1924 s	1956 s	1939 s

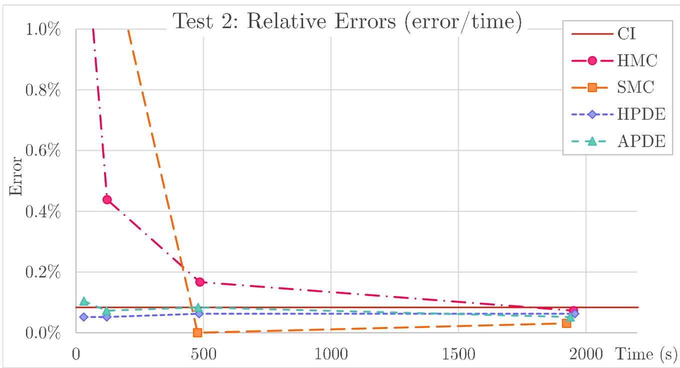

Table 4.8: Test 2. In the first Table, the fair fee α_g in bp for the Heston model, with annual or six-monthly withdrawal. In the second Table the run times for the case $WF = 1$, $T_2 = 10$. Finally, the plot of relative error (w.r.t. BM value) for the four methods in the same case of run-times Table. The parameters used for this test are available in Table 4.3 and in Table 4.7.

T_2	$WF = 1$	$WF = 2$
5	200	200
10	100	100
20	50	50

Table 4.9: The α_g values used for Delta calculation in the Static BS HW case (bp).

T_2		$WF = 1$					$WF = 2$				
		HMC	SMC	HPDE	APDE	BM	HMC	SMC	HPDE	APDE	BM
5	A	6212 ± 4	6214 ± 3	6213	6212	6213	6178 ± 4	6180 ± 4	6181	6180	6180
	B	6213 ± 3	6213 ± 1	6213	6213		6180 ± 3	6180 ± 2	6180	6180	
	C	6211 ± 1	6213 ± 1	6213	6213	± 1	6179 ± 1	6180 ± 1	6180	6180	± 1
	D	6213 ± 0	6213 ± 1	6213	6213		6179 ± 1	6180 ± 1	6180	6180	
10	A	7153 ± 7	7154 ± 6	7155	7153	7154	7138 ± 7	7129 ± 8	7133	7127	7132
	B	7155 ± 5	7152 ± 3	7154	7154		7134 ± 5	7132 ± 4	7132	7131	
	C	7152 ± 3	7153 ± 1	7154	7154	± 1	7132 ± 3	7131 ± 2	7132	7131	± 1
	D	7157 ± 2	7154 ± 1	7154	7154		7133 ± 2	7131 ± 1	7132	7131	
20	A	8018 ± 16	8010 ± 13	8017	8008	8016	8010 ± 20	8005 ± 20	8005	7995	8004
	B	8023 ± 11	8016 ± 7	8017	8014		8014 ± 14	8005 ± 10	8005	8002	
	C	8025 ± 6	8013 ± 3	8016	8015	± 1	8013 ± 7	8002 ± 4	8004	8003	± 1
	D	8020 ± 3	8015 ± 1	8016	8015		8007 ± 4	8001 ± 2	8004	8003	

Table 4.10: Test 3. Delta calculation for the Static BS HW case. All results must be multiplied by 10^{-4}. The parameters used for this test are available in Table 4.3, 4.7 and in Table 4.9 .

T_2	$WF = 1$	$WF = 2$
5	250	250
10	100	100
20	50	50

Table 4.11: The α_g values used for Delta calculation in the Static Heston case (bp).

T_2		$WF = 1$					$WF = 2$				
		HMC	SMC	HPDE	APDE	BM	HMC	SMC	HPDE	APDE	BM
5	A	6132 ± 4	6141 ± 5	6132	6131	6131	6101 ± 5	6107 ± 5	6099	6098	6098
	B	6134 ± 3	6136 ± 3	6131	6131		6101 ± 3	6104 ± 3	6098	6098	
	C	6131 ± 2	6131 ± 2	6131	6131	±1	6099 ± 2	6097 ± 2	6098	6098	±1
	D	6131 ± 1	6131 ± 1	6131	6131		6098 ± 1	6098 ± 1	6098	6098	
10	A	7287 ± 8	7297 ± 9	7286	7284	7285	7277 ± 9	7273 ± 9	7263	7261	7262
	B	7289 ± 6	7287 ± 6	7285	7284		7266 ± 6	7269 ± 6	7262	7263	
	C	7287 ± 3	7287 ± 3	7284	7284	±1	7264 ± 3	7262 ± 3	7262	7262	±1
	D	7285 ± 2	7287 ± 2	7284	7284		7263 ± 2	7264 ± 2	7262	7262	
20	A	8051 ± 19	8084 ± 19	8059	8058	8056	8048 ± 19	8053 ± 19	8048	8045	8047
	B	8067 ± 13	8074 ± 14	8058	8056		8055 ± 13	8072 ± 14	8047	8045	
	C	8060 ± 7	8068 ± 7	8057	8056	±1	8050 ± 7	8047 ± 8	8046	8045	±1
	D	8060 ± 4	8063 ± 4	8057	8056		8051 ± 4	8048 ± 4	8046	8045	

Table 4.12: Test 3. Delta calculation for the Static Heston case. All results must be multiplied by 10^{-4}. The parameters used for this test are available in Table 4.3, 4.7 and in Table 4.11 .

4.5.2 Static Withdrawal for GMWB-CF

In the Static Withdrawal case we suppose the PH to withdrawal exactly at the guaranteed rate.

The Static Tests 1 and 2 are inspired by [13]: in their article, Chen and Forsyth price a GMWB contract according to an optimal withdrawal framework, under the Black Scholes model. First we priced their product for different maturities and withdrawal rates, assuming Static withdrawals in Black and Scholes model to get a reference price in this model; we got the α value using both a standard Monte Carlo method and a standard PDE method. As we easily got the correct values for the simple Black-Scholes model, then we add stochastic interest rate and stochastic volatility. Model parameters are available in Table 4.13, and the values of α_g that we got in the Black-Scholes case are given in Table 4.14.

4.5.2.1 Test 1: Static GMWB-CF in the Black-Scholes Hull-White Model

In this test we want to price a GMWB-CF product according to BS HW model. We use the same corresponding parameters as in the Black Scholes model. Model parameters are shown in Table 4.15. Results are available in Table 4.16.

All the four methods behaved well and in the configuration D they gave results consistent with the benchmark. HPDE proved to be the best: all of its configurations gave results consistent with the benchmark. Then APDE and SMC, and HMC gave good results too. SMC performed a little better than HMC: the first method simulates the underlying value and the interest rate exactly and so it is enough to simulate the values at each event time. HMC matches the first three moments of the BS HW r process, but doesn't reproduce exactly its law: therefore it is right to increase the number of steps per year. So, for a given run time, we can simulate less scenarios in HMC than SMC: effectively, the confidence interval of HMC is larger than SMC one. Moreover, SMC over performed the benchmark when using configuration D. The two PDE methods returned stable results, and they often converged in a monotone way.

With regard to the numerical results, we observe that the α_g values decrease with increasing maturity, just as in the Black-Scholes case, and increase a little, with increasing withdrawal frequency.

4.5.2.2 Test 2: Static GMWB-CF in the Heston Model

In this test we want to price a GMWB-CF product according to the Heston model. Model parameters are shown in the Table 4.17. Results are shown in Table 4.18.

In this Test, MC methods had more difficulties; all the values of PDE methods were close to the benchmark, while some values from MC methods were less accurate, but compatibles with the benchmark (the value of BM is inside the MC confidence interval). If we compare the two MC approaches, in this case they are equivalent: HMC proved to be faster than SMC when using few time steps (we could exploit +11% simulations in configuration A), while SMC proved to be slightly faster in high time steps simulations, because of more time needed to build the volatility tree (−8% simulations in configuration D). HPDE shows to be very stable (case $T_2 = 10$, $WF = 2$, α_g didn't change through configurations B-D), APDE behaved well to (often monotone convergence).

With regard to the numerical results, we observe that the α_g values decrease with increasing maturity, just as in the Black-Scholes case, and increase a little, with increasing withdrawal frequency.

4.5.2.3 Test 3: Hedging for Static GMWB-CF

To reduce financial risks, insurance companies have to hedge the sold VA: to accomplish this target they must calculate the Greeks of products.

In this test we want to show how the different methods can be used to calculate the main Greeks. This can be done through finite differences for small shocks on the variables. In general, the PDE methods are ahead w.r.t. MC methods: the price is computed through finite differences and so the price under shock is already computed. For MC methods this is quite harder because the pricing has to be repeated changing the inputs.

To start, we calculate the underlying greek Delta, for the products of Test 1 and Test 2. As in this case we don't want to compute the fair fee α_g, we fix it arbitrarily: see Table 4.19 and Table 4.21. The values chosen are such as to cover the costs of the insurer, and may be plausible on a real case. Results are available in Table 4.20 (all values in Table must be multiplied by 10^{-4}).

In this Test, we got very accurate results with all method. Anyway, HPDE and APDE proved to be the best: they both gave stable and accurate results; in this Test, the two PDE methods were equivalent. We remark that despite fair fee changes a lot when changing the maturity of the policy, the value of Delta changes much less. Delta calculation proved to be slightly harder in the Heston model case than in the BS HW model case: see confidence intervals.

Expiry Time T	5, 10, 20 Years	GMW G	$100.0/(T \cdot WF)$
Withdrawal Frequency WF	1 or 2 per Year	Initial Premium	100.0
Initial account value A_0	100.0	S_0	100.0
Initial base b. value B_0	100.0	r	0.05
Withdrawal penalty κ	0.10	σ	0.20
Management fees α_m	0		

Table 4.13: Parameters used by Chen and Forsyth in [13].

T_2	$WF = 1$		$WF = 2$	
	PDE	MC	PDE	MC
5	235.24	235.11 ± 0.42	243.96	243.80 ± 0.42
10	92.41	92.28 ± 0.30	94.62	94.84 ± 0.30
20	27.64	27.79 ± 0.24	28.09	28.39 ± 0.24

Table 4.14: Fair bp values of α_g in Black Scholes model, for Static GMWB-CF with the same parameters as in [13].

S_0	r	$curve$	k	ω	ρ	σ
100	0.05	$flat$	1.0	0.2	−0.5	0.20

Table 4.15: The model parameters about Test 1

T_2		$WF=1$					$WF=2$				
		HMC	SMC	HPDE	APDE	BM	HMC	SMC	HPDE	APDE	BM
5	A	191.05 ± 0.83	190.81 ± 0.66	191.03	191.16	191.34	196.87 ± 0.88	196.79 ± 0.94	196.39	196.50	196.77
	B	191.79 ± 0.59	191.25 ± 0.33	191.18	191.58		196.88 ± 0.62	197.02 ± 0.47	196.55	196.60	
	C	191.34 ± 0.32	191.26 ± 0.15	191.25	191.47	±0.11	196.72 ± 0.34	196.64 ± 0.21	196.62	196.88	±0.11
	D	191.20 ± 0.18	191.25 ± 0.07	191.27	191.38		196.55 ± 0.19	196.67 ± 0.10	196.65	196.68	
10	A	79.71 ± 0.84	79.26 ± 0.66	79.41	79.33	79.44	81.38 ± 0.88	81.13 ± 0.94	80.98	81.32	80.97
	B	79.80 ± 0.60	79.43 ± 0.33	79.39	79.41		80.95 ± 0.63	81.42 ± 0.47	80.98	80.64	
	C	79.61 ± 0.32	79.56 ± 0.15	79.39	79.41	±0.08	81.12 ± 0.34	81.12 ± 0.21	80.98	81.01	±0.08
	D	79.35 ± 0.18	79.44 ± 0.07	79.38	79.40		80.90 ± 0.19	80.98 ± 0.10	80.99	81.01	
20	A	26.33 ± 0.98	25.04 ± 0.77	24.90	24.90	24.81	25.04 ± 1.07	25.72 ± 1.06	25.27	25.20	25.16
	B	25.92 ± 0.69	25.23 ± 0.39	24.86	24.67		25.91 ± 0.75	25.59 ± 0.54	25.23	25.16	
	C	25.16 ± 0.37	24.91 ± 0.17	24.84	24.81	±0.07	25.16 ± 0.41	25.35 ± 0.24	25.21	25.18	±0.07
	D	24.99 ± 0.21	24.81 ± 0.09	24.84	24.82		25.40 ± 0.23	25.13 ± 0.12	25.20	25.18	

	HMC	SMC	HPDE	APDE
A	31 s	30 s	30 s	30 s
B	121 s	121 s	120 s	118 s
C	482 s	484 s	464 s	481 s
D	1920 s	1899 s	1893 s	1909 s

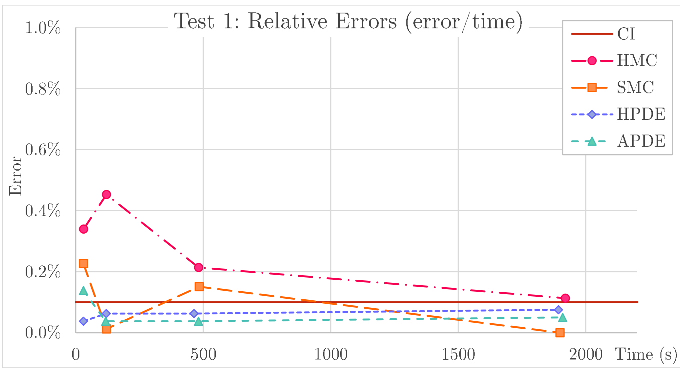

Table 4.16: Test 1. In the first Table, the fair fee α_g in bp for the Black-Scholes Hull-White model, with annual or six-monthly withdrawal. In the second Table the run times for the case $WF=1$, $T_2=10$. Finally, the plot of relative error (w.r.t. BM value) for the four methods in the same case of run-times Table. The parameters used for this test are available in Table 4.13 and in Table 4.15.

S_0	v_0	θ	k	ω	ρ	r
100	0.20^2	0.20^2	1.0	0.2	-0.5	0.05

Table 4.17: The model parameters about Test 2.

T_2		$WF = 1$					$WF = 2$				
		HMC	SMC	HPDE	APDE	BM	HMC	SMC	HPDE	APDE	BM
5	A	232.30 ± 0.97	232.83 ± 1.03	231.20	231.62	231.38	239.82 ± 1.00	240.88 ± 1.05	239.34	239.51	239.54
	B	231.72 ± 0.68	231.73 ± 0.75	231.37	231.55		240.41 ± 0.71	239.76 ± 0.77	239.52	239.38	
	C	231.39 ± 0.37	231.56 ± 0.41	231.43	231.48	± 0.10	239.84 ± 0.38	239.76 ± 0.41	239.59	239.72	± 0.11
	D	231.42 ± 0.21	231.56 ± 0.23	231.45	231.47		239.48 ± 0.20	239.71 ± 0.23	239.61	239.64	
10	A	97.13 ± 1.10	97.55 ± 1.07	95.86	95.91	95.81	98.29 ± 1.05	99.50 ± 1.10	97.99	98.25	97.98
	B	96.23 ± 0.73	97.07 ± 0.78	95.86	95.88		98.58 ± 0.74	98.62 ± 0.78	98.01	98.78	
	C	95.65 ± 0.39	95.81 ± 0.42	95.87	95.89	± 0.08	98.12 ± 0.40	97.78 ± 0.43	98.01	98.01	± 0.09
	D	95.88 ± 0.23	95.84 ± 0.24	95.87	95.86		97.93 ± 0.23	97.95 ± 0.24	98.01	98.00	
20	A	31.84 ± 1.17	31.84 ± 1.23	30.71	30.68	30.57	31.39 ± 1.20	32.06 ± 1.27	31.18	31.26	31.05
	B	31.42 ± 0.84	31.78 ± 0.90	30.64	30.60		31.69 ± 0.85	31.12 ± 0.90	31.11	31.13	
	C	30.53 ± 0.45	30.99 ± 0.47	30.63	30.63	± 0.06	31.02 ± 0.45	31.04 ± 0.50	31.10	31.09	± 0.07
	D	30.73 ± 0.26	30.61 ± 0.27	30.63	30.63		31.47 ± 0.26	31.10 ± 0.27	31.09	31.08	

	HMC	SMC	HPDE	APDE
A	30 s	30 s	30 s	31 s
B	122 s	118 s	121 s	120 s
C	486 s	477 s	483 s	479 s
D	1951 s	1924 s	1956 s	1939 s

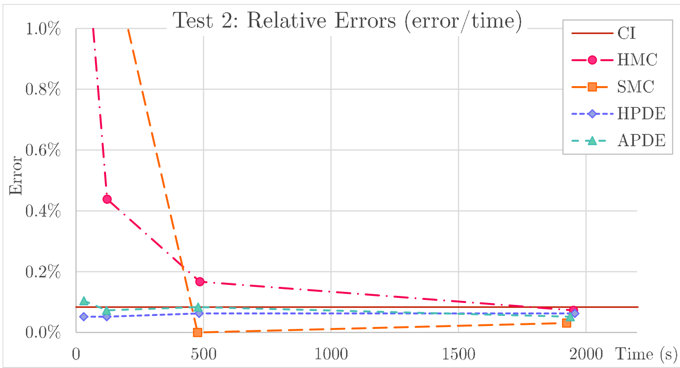

Table 4.18: Test 2. In the first Table, the fair fee α_g in bp for the Heston model, with annual or six-monthly withdrawal. In the second Table the run times for the case $WF = 1$, $T_2 = 10$. Finally, the plot of relative error (w.r.t. BM value) for the four methods in the same case of run-times Table. The parameters used for this test are available in Table 4.13 and in Table 4.17.

T_2	$WF = 1$	$WF = 2$
5	200	200
10	100	100
20	50	50

Table 4.19: The α_g values used for Delta calculation in the Static BS HW case (bp).

T_2		$WF = 1$					$WF = 2$				
		HMC	SMC	HPDE	APDE	BM	HMC	SMC	HPDE	APDE	BM
5	A	6212 ± 4	6214 ± 3	6213	6212	6213	6178 ± 4	6180 ± 4	6181	6180	6180
	B	6213 ± 3	6213 ± 1	6213	6213		6180 ± 3	6180 ± 2	6180	6180	
	C	6211 ± 1	6213 ± 1	6213	6213	±1	6179 ± 1	6180 ± 1	6180	6180	±1
	D	6213 ± 0	6213 ± 1	6213	6213		6179 ± 1	6180 ± 1	6180	6180	
10	A	7153 ± 7	7154 ± 6	7155	7153	7154	7138 ± 7	7129 ± 8	7133	7127	7132
	B	7155 ± 5	7152 ± 3	7154	7154		7134 ± 5	7132 ± 4	7132	7131	
	C	7152 ± 3	7153 ± 1	7154	7154	±1	7132 ± 3	7131 ± 2	7132	7131	±1
	D	7157 ± 2	7154 ± 1	7154	7154		7133 ± 2	7131 ± 1	7132	7131	
20	A	8018 ± 16	8010 ± 13	8017	8008	8016	8010 ± 20	8005 ± 20	8005	7995	8004
	B	8023 ± 11	8016 ± 7	8017	8014		8014 ± 14	8005 ± 10	8005	8002	
	C	8025 ± 6	8013 ± 3	8016	8015	±1	8013 ± 7	8002 ± 4	8004	8003	±1
	D	8020 ± 3	8015 ± 1	8016	8015		8007 ± 4	8001 ± 2	8004	8003	

Table 4.20: Test 3. Delta calculation for the Static BS HW case. All results must be multiplied by 10^{-4}. The parameters used for this test are available in Table 4.13, 4.17 and in Table 4.19 .

T_2	$WF = 1$	$WF = 2$
5	250	250
10	100	100
20	50	50

Table 4.21: The α_g values used for Delta calculation in the Static Heston case (bp).

T_2		$WF = 1$					$WF = 2$				
		HMC	SMC	HPDE	APDE	BM	HMC	SMC	HPDE	APDE	BM
5	A	6132 ± 4	6141 ± 5	6132	6131	6131 ±1	6101 ± 5	6107 ± 5	6099	6098	6098 ±1
	B	6134 ± 3	6136 ± 3	6131	6131		6101 ± 3	6104 ± 3	6098	6098	
	C	6131 ± 2	6131 ± 2	6131	6131		6099 ± 2	6097 ± 2	6098	6098	
	D	6131 ± 1	6131 ± 1	6131	6131		6098 ± 1	6098 ± 1	6098	6098	
10	A	7287 ± 8	7297 ± 9	7286	7284	7285 ±1	7277 ± 9	7273 ± 9	7263	7261	7262 ±1
	B	7289 ± 6	7287 ± 6	7285	7284		7266 ± 6	7269 ± 6	7262	7263	
	C	7287 ± 3	7287 ± 3	7284	7284		7264 ± 3	7262 ± 3	7262	7262	
	D	7285 ± 2	7287 ± 2	7284	7284		7263 ± 2	7264 ± 2	7262	7262	
20	A	8051 ± 19	8084 ± 19	8059	8058	8056 ±1	8048 ± 19	8053 ± 19	8048	8045	8047 ±1
	B	8067 ± 13	8074 ± 14	8058	8056		8055 ± 13	8072 ± 14	8047	8045	
	C	8060 ± 7	8068 ± 7	8057	8056		8050 ± 7	8047 ± 8	8046	8045	
	D	8060 ± 4	8063 ± 4	8057	8056		8051 ± 4	8048 ± 4	8046	8045	

Table 4.22: Test 3. Delta calculation for the Static Heston case. All results must be multiplied by 10^{-4}. The parameters used for this test are available in Table 4.13, 4.17 and in Table 4.21 .

T_2	$WF = 1$			$WF = 2$		
	PDE	MC	Ch.Fo.	PDE	MC	Ch.Fo.
5	248.33	247.75 ± 1.39	*n.c.*	258.20	257.32 ± 1.42	*n.c.*
10	129.18	128.58 ± 1.08	129.10	133.60	133.09 ± 1.11	133.52
20	66.42	66.20 ± 0.89	*n.c.*	68.59	68.52 ± 1.29	*n.c.*

Table 4.23: Fair bp values of α_g in Black Scholes model, for Dynamic GMWB-CF with the same parameters as in [13]. The values that aren't available in [13] (not computed) are denoted by *"n.c."* .

4.5.3 Dynamic Withdrawal for GMWB-CF

In the Dynamic withdrawal case we suppose the PH to chose at each event time how much withdraw, in order to maximize his (her) gain (optimal withdrawal).

The Static Tests 4 and 5 are inspired by [13]: in their article, Chen and Forsyth price a GMWB contract in a optimal withdrawal framework, under the Black Scholes model. First we priced their product for different maturities and withdrawal rates, assuming optimal withdrawals in Black and Scholes model to get a reference price in this model; we got the α value using both a Regression by Lines Monte Carlo method and a standard PDE method. As we got the good values for the simple Black-Scholes model, then we add stochastic interest rate and stochastic volatility. Model parameters are available in Table 4.13, and the values of α_g that we got are given in Table 4.23.

We remark that we used the Full Regression algorithm for the calculation of the MC prices (case A, B, C, D for SMC and HMC): this method is quite fast, however the results quality is low.

Conversely, we used the Regression by Lines algorithm to calculate the benchmarks (BM): this algorithm is much more time demanding than the Full Regression, but its results are higher, proving that the regression performs better and the PH, using this approach, can have a better payoff. Moreover, this method performed very well in the Black-Scholes model, and we used it to fill Table 4.23. We tried to use Regression by Lines algorithm also for cases A, B, C, D but we didn't get good results, because of the short run time available (max 30 mins).

For benchmarks calculation, we used 4 degree polynomials with 10^6 scenarios (doubled by the antithetic variables technique), excluding the case $T_2 = 20$, $WF = 2$ where we used half scenarios: the time needed to perform these calculations (two secant steps around the value of case D of HPDE) varies from 30 minutes (case $T_2 = 5$, $WF = 1$) to 38 hours (case $T_2 = 20$, $WF = 2$).

We would remark that, using PDE method for the Black Scholes model, we obtained the same values as in [13] (only two values are available in Chen and Forsyth's paper), but MC method (Regression by Lines) had a few problems (lower values): the least squares regression doesn't work very well and this problem is stiff for MC methods (see Table 4.23). We can therefore imagine that the MC methods will have difficulties also in the following tests, in which a dimension is added.

4.5.3.1 Test 4: Dynamic GMWB-CF in the Black-Scholes Hull-White Model

Test 4 is the Dynamic case of Test 1. Model parameters are shown in Table 4.15. Results are available in Table 4.24. In this test PDE methods proved to be much more efficient than MC ones. In fact MC methods use a least-squares regression approach to find the optimal withdrawal: this method needs a lot of scenarios to approximate through the regression the value of the policy for a given set of variable, and this is time demanding. Then, working at fixed time, we could perform fewer scenarios than the Static case. PDE methods felt another problem: the increase of problem dimension forced us to reduce the number of time steps wrt Static case. Using MC methods, we always got lower values with regard to PDE methods, and moreover MC values increased by several bps when moving from configuration A to D.

The two MC methods proved to be equivalent: the differences in scenarios generation runtime are negligible because most of the time is spent in finding the best withdrawal. The HPDE method gave good and stable results, while APDE had more troubles, with results floating around the good values. Then, the HPDE method proved to be the best one according to the results of this test.

The case $(T_2, WF) = (20, 2)$ proved to be very insidious: the long maturity and the large number of withdrawal dates (40 event times) made the problem hard also for PDE methods. In this case MC methods in configuration A also gave lower values than Static approach (18.64 bp vs 25.20 bp): due to the few scenarios considered, the least squares regression failed to increase PH's gain.

4.5.3.2 Test 5: Dynamic GMWB-CF in the Heston Model

Test 5 is the Dynamic case of Test 2. Model parameters are shown in Table 4.17. Results are available in Table 4.25.

In this test things are similar to Test 4, but the optimization problem seemed to be easier than in Test 4: MC methods converged better, especially when using high level configurations. PDE methods behaved good as usual, and in this case they proved to be almost equivalent: they both gave good results except for the case $(T_2, WF) = (20, 2)$ where the initial results of APDE were too high. The two MC methods proved to be equivalent. We note that, in the Heston model case, Dynamic strategy increased the value of α_g less than in BS HW case: probably, playing on interest rate lets the PH gain more than playing on volatility.

The case $(T_2, WF) = (20, 2)$ is still the most insidious, but this time we didn't get any value lower than the Static value of α_g.

4.5.3.3 Test 6: Hedging for Dynamic GMWB-CF

Test 6 is the Dynamic case of Test 3. Results are available in Table 4.27.

In this test we got good results with PDE methods: values of HPDE are very regular despite the high dimension of the problem. Results from APDE are good, but a bit worse than HPDE especially in BS HW case (see for example case $(T_2, WF) = (20, 2)$). Monte Carlo methods suffered the few scenarios performed and sometimes the confidence interval is very large. In the case $(T_2, WF) = (20, 2)$ we also got some convergence problems in the BS HW model.

4.5.3.4 Optimal Withdrawal Strategy Plots for Dynamic GMWB-CF

In Figure 4.5.1 and 4.5.2 we calculated the optimal withdrawal for the GMWB-CF product with $(T_2, WF) = (10, 1)$ for both the BS HW model and Heston model. We used HPDE methods to obtain these plots: we chose three nodes of the tree around the initial value at time $t = 1$ and we used the best withdrawals to get these plots.

We remark that these plots are very similar to those proposed in [13]: we note the same structure around the bisector and the wide region of regular withdrawal.

T_2		WF = 1 HMC	SMC	HPDE	APDE	BM	WF = 2 HMC	SMC	HPDE	APDE	BM
5	A	226.22 ± 3.04	223.88 ± 2.73	282.19	278.25	282.00	244.75 ± 5.85	242.09 ± 5.78	320.54	313.90	319.00
	B	255.42 ± 2.36	256.27 ± 2.25	282.24	276.29		277.53 ± 5.15	275.53 ± 5.12	320.44	320.59	
	C	266.97 ± 1.71	265.12 ± 1.65	282.28	280.55	±1.54	310.22 ± 3.58	308.76 ± 3.67	320.35	320.14	±1.69
	D	275.83 ± 1.31	272.62 ± 1.23	282.32	282.63		312.12 ± 2.82	311.16 ± 2.72	320.33	320.73	
10	A	128.40 ± 4.58	130.58 ± 4.01	163.54	160.38	162.51	142.01 ± 11.54	141.73 ± 10.55	194.56	192.57	186.42
	B	144.12 ± 3.81	145.35 ± 3.71	163.03	157.76		146.20 ± 5.09	149.63 ± 4.95	190.76	190.90	
	C	155.56 ± 2.72	155.54 ± 2.80	162.92	159.72	±1.23	165.58 ± 3.97	169.64 ± 3.87	189.66	188.87	±1.39
	D	156.97 ± 1.99	155.34 ± 2.02	162.86	157.37		182.37 ± 3.31	180.42 ± 3.16	189.47	188.24	
20	A	65.58 ± 4.45	65.77 ± 5.52	90.67	62.10	84.01	92.16 ± 24.50	18.64 ± 21.96	109.26	13.72	98.96
	B	65.50 ± 3.13	67.69 ± 3.57	86.92	87.53		80.34 ± 12.31	79.69 ± 19.48	106.77	31.67	
	C	75.87 ± 2.72	76.31 ± 2.67	86.11	86.42	±1.03	84.79 ± 6.22	83.28 ± 5.10	106.04	71.80	±1.64
	D	78.89 ± 2.14	81.15 ± 2.44	85.73	85.75		89.68 ± 4.30	92.43 ± 4.15	104.49	95.82	

	HMC	SMC	HPDE	APDE
A	29 s	31 s	31 s	30 s
B	121 s	124 s	121 s	123 s
C	484 s	484 s	489 s	474 s
D	1881 s	1927 s	1899 s	1901 s

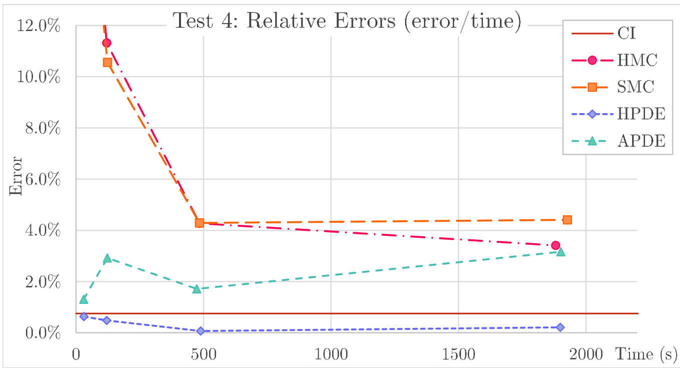

Table 4.24: Test 4. In the first Table, the fair fee α_g in bp for the Black-Scholes Hull-White model, with annual or six-monthly withdrawal. In the second Table the run times for the case $WF = 1$, $T_2 = 10$. Finally, the plot of relative error (w.r.t. BM value) for the four methods in the same case of run-times Table. The parameters used for this test are available in Table 4.13 and in Table 4.15.

T_2		$WF=1$ HMC	SMC	HPDE	APDE	BM	$WF=2$ HMC	SMC	HPDE	APDE	BM
5	A	238.42 ± 2.36	238.87 ± 2.38	246.27	246.33	246.45	242.33 ± 4.37	245.93 ± 4.46	255.51	256.10	255.20
	B	237.80 ± 1.49	240.29 ± 1.50	246.58	246.44		244.68 ± 2.93	246.49 ± 2.96	256.62	256.43	
	C	242.33 ± 1.09	242.95 ± 1.08	246.62	246.66	± 1.07	248.81 ± 2.21	250.94 ± 2.21	256.67	256.27	± 1.10
	D	243.42 ± 0.79	243.08 ± 0.77	246.64	246.68		253.06 ± 1.56	251.79 ± 1.58	256.70	256.31	
10	A	125.77 ± 4.03	123.42 ± 4.05	133.70	133.92	133.72	132.71 ± 12.0	119.84 ± 11.10	137.85	146.61	137.00
	B	126.16 ± 2.47	126.92 ± 2.51	133.89	133.91		118.80 ± 4.59	122.60 ± 4.57	138.12	139.39	
	C	129.47 ± 1.76	130.41 ± 1.73	133.98	133.96	± 0.86	125.17 ± 3.07	124.11 ± 2.97	138.29	138.36	± 0.87
	D	132.89 ± 1.21	132.36 ± 1.22	134.02	133.99		130.63 ± 2.25	130.27 ± 2.32	138.41	138.35	
20	A	37.68 ± 11.94	37.46 ± 9.74	72.25	74.11	69.35	82.48 ± 26.34	61.22 ± 28.49	74.54	99.35	71.82
	B	64.51 ± 3.84	69.22 ± 2.44	71.05	72.30		70.71 ± 7.75	67.08 ± 8.48	73.00	86.01	
	C	66.85 ± 3.19	66.13 ± 2.47	71.12	71.57	± 0.72	65.25 ± 3.40	66.50 ± 4.05	73.12	77.08	± 1.05
	D	64.18 ± 2.55	66.88 ± 2.18	71.15	71.69		61.27 ± 2.56	62.16 ± 2.77	73.24	74.60	

	HMC	SMC	HPDE	APDE
A	32 s	32 s	29 s	30 s
B	123 s	124 s	122 s	118 s
C	483 s	475 s	474 s	495 s
D	1903 s	1882 s	1923 s	1947 s

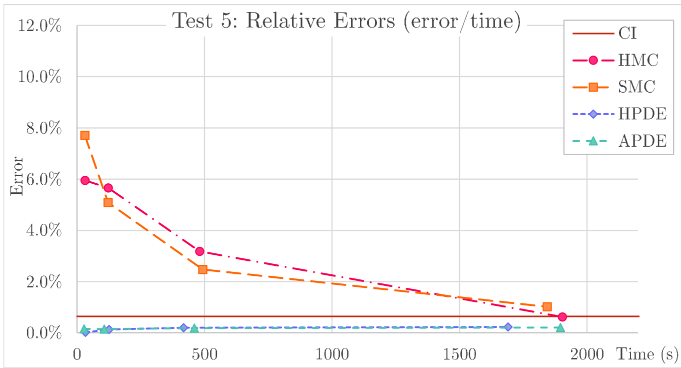

Table 4.25: Test 5. In the first Table, the fair fee α_g in bp for the Heston model, with annual or six-monthly withdrawal. In the second Table the run times for the case $WF = 1$, $T_2 = 10$. Finally, the plot of relative error (w.r.t. BM value) for the four methods in the same case of run-times Table. The parameters used for this test are available in Table 4.13 and in Table 4.17.

T_2	$WF = 1$	$WF = 2$
5	350	350
10	200	200
20	150	150

Table 4.26: The α_g values used for Delta calculation in the Dynamic BS HW case (bp).

T_2		$WF = 1$					$WF = 2$				
		HMC	SMC	HPDE	APDE	BM	HMC	SMC	HPDE	APDE	BM
5	A	4986 ± 311	4790 ± 310	4474	4465	4514	4753 ± 537	4117 ± 429	4191	4187	4181
	B	4455 ± 171	4443 ± 179	4477	4469		4220 ± 182	4362 ± 222	4196	4191	
	C	4385 ± 130	4420 ± 122	4478	4473	± 62	4057 ± 196	3987 ± 170	4198	4195	± 68
	D	4319 ± 103	4432 ± 94	4478	4476		4158 ± 167	4235 ± 170	4198	4198	
10	A	4734 ± 656	5152 ± 543	4630	4625	4593	4612 ± 946	3881 ± 817	4270	4325	4291
	B	4577 ± 320	4367 ± 307	4636	4616		4628 ± 460	3846 ± 362	4296	4316	
	C	4665 ± 259	4548 ± 240	4639	4631	± 112	3898 ± 310	4122 ± 303	4304	4300	± 123
	D	4517 ± 178	4537 ± 175	4639	4635		4492 ± 329	4201 ± 276	4306	4304	
20	A	4053 ± 302	4223 ± 112	4129	4149	4083	4152 ± 150	4037 ± 144	3639	4062	3857
	B	4370 ± 105	4253 ± 108	4153	4118		4095 ± 74	4039 ± 68	3752	3924	
	C	4046 ± 332	4011 ± 326	4157	4150	± 143	4078 ± 72	4049 ± 60	3766	3803	± 211
	D	3980 ± 268	3857 ± 242	4157	4145		3434 ± 294	3659 ± 332	3780	3798	

Table 4.27: Test 6. Delta calculation for the Dynamic BS HW case. All results must be multiplied by 10^{-4}. The parameters used for this test are available in Table 4.13, 4.26 and in Table 4.17.

T_2	$WF=1$	$WF=2$
5	300	300
10	150	150
20	100	100

Table 4.28: The α_g values used for Delta calculation in the Dynamic Heston case (bp).

T_2		$WF=1$					$WF=2$				
		HMC	SMC	HPDE	APDE	BM	HMC	SMC	HPDE	APDE	BM
5	A	5732 ± 40	5727 ± 26	5629	5631	5637	5878 ± 97	5803 ± 62	5571	5577	5599
	B	5715 ± 28	5699 ± 27	5628	5630		5678 ± 88	5603 ± 78	5570	5572	
	C	5614 ± 49	5668 ± 66	5628	5629	± 26	5674 ± 122	5635 ± 78	5570	5570	± 25
	D	5607 ± 44	5653 ± 55	5628	5628		5695 ± 103	5618 ± 63	5569	5570	
10	A	6082 ± 179	6103 ± 157	6007	6009	5983	6918 ± 712	6083 ± 392	5938	5915	5914
	B	5949 ± 133	5886 ± 125	6006	6007		6225 ± 263	6058 ± 142	5936	5942	
	C	5909 ± 117	6062 ± 136	6005	6006	± 58	5779 ± 151	6026 ± 116	5936	5939	± 55
	D	6008 ± 113	6059 ± 97	6004	6005		5980 ± 121	5840 ± 114	5936	5937	
20	A	5604 ± 480	5658 ± 383	5636	5644	5635	4162 ± 987	5886 ± 831	5540	5122	5343
	B	5428 ± 405	5855 ± 433	5635	5642		5056 ± 362	5543 ± 437	5540	5382	
	C	5410 ± 297	5571 ± 206	5635	5640	± 126	5421 ± 306	5297 ± 369	5542	5543	± 174
	D	5734 ± 213	5571 ± 199	5635	5638		5379 ± 226	5416 ± 198	5543	5550	

Table 4.29: Test 6. Delta calculation for the Dynamic Heston case. All results must be multiplied by 10^{-4}. The parameters used for this test are available in Table 4.13, 4.28 and in Table 4.17.

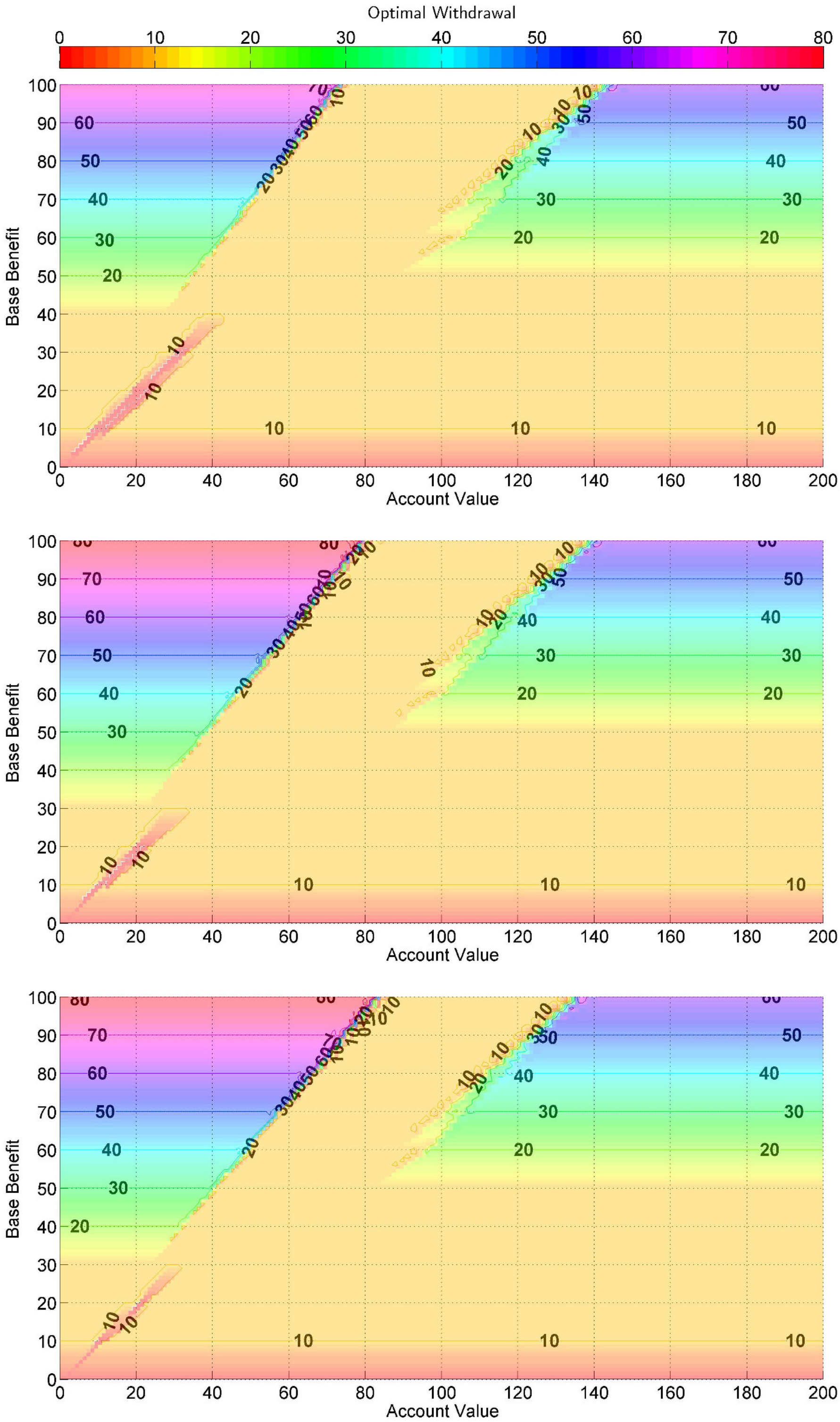

Figure 4.5.1: Plots of the optimal withdrawals at time $t = 1$ for the BS HW model according to different values of r_1: from the top to the bottom $r_1 = 0.03$, $r_1 = 0.05$ and $r_1 = 0.07$. The parameters used to obtain these plots are the same as for Delta calculation for case $T_2 = 10$, $WDF = 1$: see Tables 4.13, 4.15 and 4.26 .

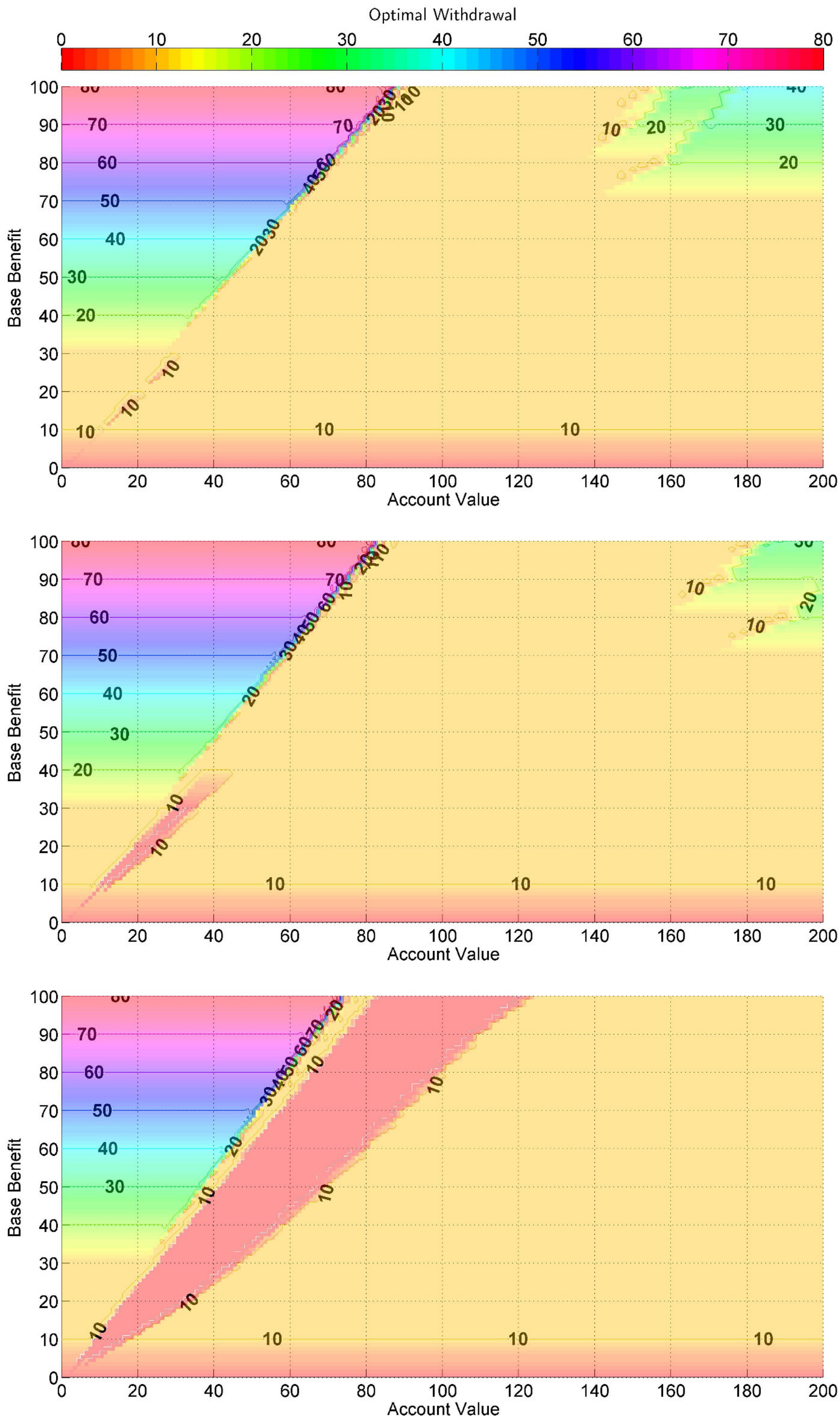

Figure 4.5.2: Plots of the optimal withdrawals at time $t = 1$ for the Heston model according to different values of the volatility v_1: from the top to the bottom $v_1 = 0$, $v_1 = 0.04$ and $v_1 = 0.16$. The parameters used to obtain these plots are the same as for Delta calculation for case $T_2 = 10$, $WDF = 1$: see Tables 4.13, 4.17 and 4.28.

Contract times (T_1, T_2)	$(0, 25)$ or $(10, 25)$	GMW G	$\frac{\max[P, W_{T_1}]}{T_2 - T_1}$
Withdrawal Frequency WF	1 Year	m	1.0
Initial account value A_0	100.0	S_0	100.0
Initial Premium	100.0	r	0.0325
Withdrawal penalty κ	0.10	σ	0.30
PH's behavior	Static or Surrendering	Mortality	OFF

Table 4.30: Parameters used by Yang and Dai in [46].

(T_1, T_2)	Static			surrendering		
	PDE	MC	YD	PDE	MC	YD
$(0, 25)$	102.02	101.95 ± 0.21	102	158.28	157.33 ± 0.41	158
$(10, 25)$	254.01	253.99 ± 0.16	170	305.35	305.26 ± 0.50	248

Table 4.31: Fair bp values of α_g in Black Scholes model, for GMWB-YD with the same parameters as in [13].

4.5.4 Static Withdrawal and Optimal surrender for GMWB-YD

In the Static Withdrawal case we suppose the PH to withdrawal exactly at the guaranteed rate, while in Optimal surrender case, the PH can stop the contract at each event time.

The Tests 7 and 8 are inspired by [46]: in their article, Yang and Dai price a GMWB contract both in Static and Dynamic (optimal surrender) framework, under the Black Scholes model. First we priced their products for different maturities and withdrawal rates, in Black and Scholes model to get a reference price in this model and to compare our results with the author's ones. We used a standard Monte Carlo method and a standard PDE method for the Black Scholes model. Then, we add stochastic volatility and stochastic interest rate. Model parameters are available in Table 4.30, and the values of α_g that we got are given in Table 4.31.

We dealt with four numerical cases: deferred or not and Static behavior or Surrendering.

We note that using different methods (a simple Monte Carlo approach, and a PDE method for the Black-Scholes model) we didn't obtain the same results of Yang and Dai in the case $(T_1, T_2) = (10, 25)$. Probably we misunderstood some contract specifications about the deferred case. We priced those products both using similarity reduction (see Section 4.2.5) and without, obtaining the same results. We would remark that Yang and Dai didn't use this technique for their product.

4.5.4.1 Test 7: GMWB-YD in the Black-Scholes Hull-White Model

In the conclusion of their paper [46], Yang and Dai proposed themselves to evaluate their contract including stochastic interest rate. That's what we do in this Chapter, and in Test 7 we present some numerical results about GMWB-YD pricing. Contract specifications are shown in Table 4.30, model parameters in Table 4.32 and the fair values of α_g in Table 4.33.

All four numerical methods behaved well in the Static case, but PDE methods outperformed the others. Thing are different in the surrendering case: the Longstaff Schwartz method showed its limits: in the BS HW model the underlying and thus the account value can diffuse so much in 25 years and the regression over such a wide set of values is stiff. PDE methods proved to be reliable and stable, especially in case $(T_1, T_2) = (0, 25)$ where 25 regressions are required.

4.5.4.2 Test 8: GMWB-YD in the Heston Model

After pricing the GMWB-CF product in the BS and BS HW model, then we did it in the Heston model. Contract specifications are shown in Table 4.30, model parameters in Table 4.34 and the fair values of α_g in Table 4.35.

Like the previous test, all four numerical methods behaved well in the Static case; HPDE and APDE outperformed the others and proved to be equivalent in that framework. In this test, numerical results of MC methods for the surrendering case are good: probably, the least square regression is easier in the Heston case. Moreover, results in the $(T_1, T_2) = (10, 25)$ case are very good: in this case, the Longstaff-Schwartz algorithm requires only 15 numerical regressions and we can simulate more scenarios than in the other case.

S_0	r	*curve*	k	ω	ρ	σ
100	0.0325	*flat*	1.0	0.2	-0.5	0.30

Table 4.32: The model parameters about Test 7.

	Static					*Surrendering*				
	HMC	SMC	HPDE	APDE	BM	HMC	SMC	HPDE	APDE	BM
	$(T_1, T_2) = (0, 25)$									
A	83.11 ± 3.55	81.30 ± 2.70	80.79	80.62	80.65	94.94 ± 5.28	89.98 ± 4.34	96.04	95.98	92.95
B	83.06 ± 2.48	80.05 ± 1.28	80.71	80.71		89.12 ± 2.20	91.75 ± 1.84	95.50	96.04	
C	82.49 ± 1.69	80.62 ± 0.57	80.71	80.72	± 0.20	89.55 ± 1.45	91.79 ± 1.34	95.52	96.08	± 0.78
D	81.48 ± 0.75	80.80 ± 0.28	80.70	80.72		89.60 ± 1.10	90.22 ± 1.11	95.53	96.09	
	$(T_1, T_2) = (10, 25)$									
A	213.24 ± 3.05	210.58 ± 2.28	210.40	210.91	210.76	242.15 ± 6.44	233.16 ± 5.58	242.38	242.86	241.41
B	212.68 ± 2.13	210.47 ± 1.11	210.67	210.99		244.06 ± 3.38	239.25 ± 2.86	242.83	243.12	
C	212.45 ± 1.44	210.72 ± 0.49	210.74	210.89	± 0.17	238.66 ± 2.02	239.37 ± 1.67	242.94	243.07	± 0.93
D	211.49 ± 0.66	210.73 ± 0.25	210.75	210.84		241.61 ± 1.29	239.75 ± 1.27	242.97	243.04	

Table 4.33: Test 7. The fair fee α_g in bp for the BS HW model, with Static withdrawal or Surrendering option. The parameters used for this test are available in Table 4.30 and in Table 4.32.

S_0	v_0	θ	k	ω	ρ	r
100	0.30^2	0.30^2	1.0	0.2	-0.5	0.0325

Table 4.34: The model parameters about Test 8.

	Static					*Surrendering*				
	HMC	SMC	HPDE	APDE	BM	HMC	SMC	HPDE	APDE	BM
	$(T_1, T_2) = (0, 25)$									
A	104.19 ± 3.43	104.49 ± 3.64	101.17	101.10	100.71	142.75 ± 4.67	140.41 ± 4.60	145.58	145.86	143.71
B	101.04 ± 2.36	102.43 ± 2.62	101.07	101.07		139.92 ± 2.97	138.92 ± 2.82	145.48	145.80	
C	101.49 ± 1.30	102.19 ± 1.42	101.07	101.08	± 0.52	141.57 ± 1.55	140.22 ± 1.61	145.61	145.78	± 0.57
D	101.45 ± 0.75	101.05 ± 0.80	101.07	101.08		142.04 ± 1.05	142.41 ± 1.08	145.62	145.77	
	$(T_1, T_2) = (10, 25)$									
A	246.57 ± 2.70	248.46 ± 2.90	244.67	244.45	244.52	280.93 ± 5.33	286.79 ± 5.31	286.20	286.11	286.39
B	245.51 ± 1.90	248.04 ± 2.11	244.76	244.68		285.72 ± 2.73	286.46 ± 2.88	286.46	286.42	
C	245.31 ± 1.03	245.75 ± 1.14	244.80	244.78	± 0.41	286.67 ± 1.70	285.03 ± 1.72	286.56	286.52	± 0.65
D	245.42 ± 0.60	245.18 ± 0.65	244.81	244.80		286.54 ± 1.01	286.41 ± 1.01	286.57	286.60	

Table 4.35: Test 8. The fair fee α_g in bp for the Heston model, with Static withdrawal or Surrendering option. The parameters used for this test are available in Table 4.30 and in Table 4.34.

4.6 Conclusions

In this Chapter we have developed four numerical methods to price two versions of GMWB contracts under different conditions. Regarding the stochastic model, both stochastic interest rate and stochastic volatility effects have been considered. Regarding the policy holder's behavior, both static and dynamic strategy have been considered.

Since GMWB variable annuities are such a long maturity products, the effects of stochastic interest rate and stochastic volatility cannot be overlook. In particular, the impact of stochastic interest rate seems to be more relevant.

All four methods gave compatible results both for pricing and delta calculation. The fair hedging fee (i.e. the cost of maintaining the replicating portfolio) is determined using a sequence of parameters refinements. The PDE methods proved to be not very expensive, while MC methods proved to be more expensive. The Hybrid PDE seemed to be the more performing than the others for its convergence speed and stability of results. Also ADI PDE behaved very well but the implementation was a little harder than Hybrid PDE one; moreover the choice of the good parameters for ADI PDE was a source of troubles. In the BS HW model case, Standard MC, thanks to its exact simulation, outperformed the hybrid method while, in the Heston model case, the MC methods proved to be roughly equivalent, even if the Hybrid MC was easier to be implemented.

As we said before, PDE methods proved to be much more efficient than MC methods, especially in Dynamic case where it's much more simple to implement the optimal withdrawal choice. In the GMWB-YD case, similarity reduction reduces the dimension of the problem to two and therefore PDE methods perform very well. In the GMWB-CF case similarity reduction cannot be applied and therefore pricing is an harder task, especially in the case of six-monthly withdrawal and 20 years maturity. Anyway, we have to remark that MC methods offer a confidence interval for results, they are useful in risk measures calculation (for example VAR or ES), and they are preferred by insurance companies because of their attachment to the concept of scenario.

The use of special numerical techniques (splines, improved LS convergence) allowed to ensure the convergence containing the computational time.

A future development that could be treated is to combine stochastic interest rate and stochastic volatility: the combined model could be an element of greater realism.

We conclude by pointing out that our methods are quite flexible in that they can accommodate a wide variety of policy holder withdrawal strategies such as ones derived from utility-based models.

Appendix A

Proof of the formulas in 3.4.1.2

We remark that DB and MF denote the average (w.r.t. the death year) value of the discounted death benefit and management fees paid in $[t_i, t_{i+1}]$, and discounted in t_i.

CASE 1: DB paid at the end, Fees withdrawn at the end The death benefit is paid at the end of the period and so it is equal to $A^{1+}_{t_{i+1}}$. The fraction of the original PHs who dies in $[t_i, t_{i+1}]$ is equal to $\mathcal{M}(t_i)$, so we get

$$DB = \mathcal{M}(t_i)\, e^{-\int_{t_i}^{t_{i+1}} r_s ds} A^{4+}_{t_i} \frac{S_{t_{i+1}}}{S_{t_i}} e^{-\alpha_{tot}\Delta t}$$

Management fees are paid only by PHs still alive at time t_i: this fraction is equal to $\mathcal{R}(t_i)$. These fees are calculated on the account value at the end of the period. Therefore we get

$$MF = \mathcal{R}(t_i)\, e^{-\int_{t_i}^{t_{i+1}} r_s ds} A^{4+}_{t_i} \frac{S_{t_{i+1}}}{S_{t_i}} \left(1 - e^{-\alpha_{tot}\Delta t}\right) \frac{\alpha_m}{\alpha_{tot}}$$

CASE 2: DB paid at the end, Fees withdrawn continuously The formula for death benefit is the same as before

$$DB = \mathcal{M}(t_i)\, e^{-\int_{t_i}^{t_{i+1}} r_s ds} A^{4+}_{t_i} \frac{S_{t_{i+1}}}{S_{t_i}} e^{-\alpha_{tot}\Delta t}$$

In this case the management fees are withdrawn continuously. We have shown that in this case the value of the management fees paid in $[t_i, t]$ follows the equation

$$F_t = \int_{t_i}^{t} e^{\int_s^t r_u du} \alpha_m A_s ds$$

then we can easily get

$$MF = \mathcal{R}(t_i)\, \alpha_m \frac{A^{4+}_{t_i}}{S_{t_i}} \int_{t_i}^{t_{i+1}} e^{-\int_{t_i}^{t} r_s ds} S_t e^{-\alpha_{tot}(t-t_i)} dt$$

CASE 3: DB paid immediately, Fees withdrawn at the end If the death benefit is paid immediately, then we can get its mean by integrating from t_i to t_{i+1}

$$DB = \mathcal{M}(t_i)\frac{A_{t_i}^{4+}}{S_{t_i}}\int_{t_i}^{t_{i+1}} e^{-\int_{t_i}^{t} r_s ds} S_t e^{-\alpha_{tot}(t-t_i)}dt$$

To calculate the value in t_i of the average management fees we have to distinguish two cases: if the PH has die in $[t_i, t_{i+1}]$ or not. If he has died before t_i , he pays no fees. If he is still alive in t_{i+1} the value of his fees is equal to $e^{-\int_{t_i}^{t_{i+1}} r_s ds} A_{t_i}^{4+}\frac{S_{t_{i+1}}}{S_{t_i}}\left(1-e^{-\alpha_{tot}\Delta t}\right)\frac{\alpha_m}{\alpha_{tot}}dt$, as we computed in case 1. Otherwise, if he has died in $[t_i, t_{i+1}]$, we have to integrate on the continuous density of the death random variable the benefit paid at the death. We remark that we suppose that, in this case, if the PH doesn't survive, the fees are taken before the payment of the death benefit. Therefore, in this case we get $\mathcal{M}(t_i)\frac{\alpha_m}{\alpha_{tot}}\frac{A_{t_i}^{4+}}{S_{t_i}}\int_{t_i}^{t_{i+1}} S_t\left(1-e^{-\alpha_{tot}(t-t_i)}\right)e^{-\int_{t_i}^{t} r_u du}dt$. Finally, adding together the two parts, we get

$$MF = \mathcal{M}(t_i)\frac{\alpha_m}{\alpha_{tot}}\frac{A_{t_i}^{4+}}{S_{t_i}}\int_{t_i}^{t_{i+1}} S_t\left(1-e^{-\alpha_{tot}(t-t_i)}\right)e^{-\int_{t_i}^{t} r_u du}dt +$$
$$+\mathcal{R}(t_{i+1})\, e^{-\int_{t_i}^{t_{i+1}} r_s ds} A_{t_i}^{4+}\frac{S_{t_{i+1}}}{S_{t_i}}\left(1-e^{-\alpha_{tot}\Delta t}\right)\frac{\alpha_m}{\alpha_{tot}}$$

CASE 4: DB paid immediately, Fees withdrawn continuously The formula for death benefit is the same as before

$$DB = \mathcal{M}(t_i)\frac{A_{t_i}^{4+}}{S_{t_i}}\int_{t_i}^{t_{i+1}} e^{-\int_{t_i}^{t} r_s ds} S_t e^{-\alpha_{tot}(t-t_i)}dt$$

To calculate the value in t_i of the average management fees we have to distinguish two cases: if the PH has die in $[t_i, t_{i+1}]$ or not. If he has died before t_i , he pays no fees. If he is still alive in t_{i+1} the value of his fees is equal to $\alpha_m\frac{A_{t_i}^{4+}}{S_{t_i}}\int_{t_i}^{t_{i+1}} e^{-\int_{t_i}^{t} r_s ds} S_t e^{-\alpha_{tot}(t-t_i)}dt$, as we computed in case 2. Otherwise, if he has died in $[t_i, t_{i+1}]$, we have to integrate on the continuous density of the death random variable the benefit paid at the death.We remark that we suppose that, in this case, if the PH doesn't survive, the fees are taken up to his death, and then no more. The contribution from this case is equal to

$$\int_{t_i}^{t_{i+1}} S_t e^{-\int_{t_i}^{t} r_u du}\mathcal{M}(t_i)\, dt =$$

$$= \int_{t_i}^{t_{i+1}}\left(\alpha_m\frac{A_{t_i}^{4+}}{S_{t_i}}\int_{t_i}^{t} e^{\int_v^t r_u du} S_v e^{-\alpha_{tot}(v-t_i)}dv\right)e^{-\int_{t_i}^{t} r_u du}\mathcal{M}(t_i)\, dt$$
$$= \alpha_m\frac{A_{t_i}^{+}}{S_{t_i}}\mathcal{M}(t_i)\int_{t_i}^{t_{i+1}}\int_{t_i}^{t} S_v e^{-\alpha_{tot}(v-t_i)}e^{-\int_{t_i}^{v} r_u du}dv dt$$

This is an integral of the following type

$$\int_a^b \int_a^x f(y)\, dy\, dx$$

and it can be rewritten as

$$\int_a^b \int_y^b f(y)\, dx\, dy = \int_a^b f(y)\,(b-y)\; dy$$

Therefore

$$\alpha_m \frac{A_{t_i}^{+}}{S_{t_i}} \mathcal{M}(t_i) \int_{t_i}^{t_{i+1}} S_s e^{-\alpha_{tot}(s-t_i)} e^{-\int_{t_i}^{s} r_u du} (t_{i+1}-s)\, ds$$

Finally we get

$$MF = \mathcal{M}(t_i)\, \alpha_m \frac{A_{t_i}^{4+}}{S_{t_i}} \int_{t_i}^{t_{i+1}} S_t e^{-\alpha_{tot}(t-t_i)} e^{-\int_{t_i}^{t} r_u du} (t_{i+1}-t)\, dt+$$
$$+\mathcal{R}(t_{i+1})\, \alpha_m \frac{A_{t_i}^{4+}}{S_{t_i}} \int_{t_i}^{t_{i+1}} e^{-\int_{t_i}^{t} r_s ds} S_t e^{-\alpha_{tot}(t-t_i)} dt$$

Bibliography

[1] Y. Achdoul, O. Bokanowski, T. Lelievre, (2007). Partial differential equations in finance, *CERMICS*, Paris.

[2] A. Alfonsi, (2010). High order discretization schemes for the CIR process: application to Affine Term Structure and Heston models. *Mathematics of Computation*, Vol. 79, No. 269, pp. 209-237.

[3] A. D. Andricopoulos, M. Widdicks, P. W. Duck, D. P. Newton, (2004). Curtailing the Range for Lattice and Grid Methods. *The Journal of Derivatives* Vol. 11 Issue 4, Pag 55-61.

[4] E. Appolloni, L. Caramellino A. Zanette, (2014). A robust tree method for pricing American options with the Cox-Ingersoll-Ross interest rate model. *IMA J Management Math* first published online January 15, 2014 doi:10.1093/imaman/dpt030.

[5] A. R. Bacinello, P. Millossovich, A. Olivieri, E. Pitacco, (2011). Variable annuities: A unifying valuation approach. *Insurance: Mathematics and Economics* 49, pp. 285-297.

[6] A. Belanger, P. Forsyth, G. Labahn, (2009). Valuing the guaranteed minimum death benefit clause with partial withdrawals. *Applied Mathematical Finance* 16, pp. 451-496.

[7] F. Black, M. Scholes, (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, Vol 81, No. 3, pp. 637-654.

[8] A. Bonafede, (2008). Variable Annuity: l'uovo di Colombo. *Affari e Finanza*, 04/07/2008 page 22.

[9] P. Brandimarte, (2002). Numerical methods in finance, Wiley-Interscience,Politecnico di Torino, Torino.

[10] M. Briani, L. Caramellino, A. Zanette, (2015). A hybrid approach for the implementation of the Heston model. Accepted for pubblication in *IMA Journal of Management Mathematics*, DOI 10.1093/imaman/dpv032

[11] M. Briani, L. Caramellino, A. Zanette, (2015). Numerical approximations for Heston-Hull-White type models. *The Journal of Computational Finance*, forthcoming.

[12] D. Brigo, F. Mercurio, (2006). Interest rate models-Theory and practice. *Springer*, Berlin.

[13] Z. Chen, P. A. Forsyth, (2007). A Numerical Scheme for the Impulse Control Formulation for Pricing Variable Annuities with a Guaranteed Minimum Withdrawal Benefit (GMWB). *Numerische Mathematik*, Vol. 109, pp. 535-569.

[14] Z. Chen, K. Vetzal, P. Forsyth, (2008). The effect of modelling parameters on the value of GMWB guarantees. *Insurance: Mathematics and Economics* 43, pp. 165-173.

[15] E. Clément, D. Lamberton, P. Protter, (2011). An analysis of the Longstaff-Schwartz algorithm for American option pricing.

[16] R. Courant, K Friedrichs, H. Lewy, (1967). On the partial difference equations of mathematical physics. *IBM J. Res. Develop.*, 11:215-234.

[17] J. C. Cox, S. Ross, M. Rubinstein, (1976). The Valuation of options for alternative stochastic processes. *Journal of Financial economics* 3, 145-166.

[18] J. C. Cox, M. Rubinstein, (1985). Options Markets, *Prentice-Hall*, Englewood Cliffs, NJ.

[19] M. Dai, Y. K. Kwok, J. Zong, (2007). Guaranteed minimum withdrawal benefit in variable annuities. *Mathematical Finance*, forthcoming.

[20] L. Devroye, (1985). Non Uniform Random Variate Generation. *Springer-Verlag.*

[21] P. Forsyth, K.Vetzal, (2014). An optimal stochastic control framework for determining the cost of hedging of variable annuities. *Journal of Economic Dynamics and Control* 44 (2014), pp. 29-53.

[22] P. Gaillardetz, J. Lakhmiri, (2011). A new premium principle for equity indexed annuities. *Journal of Risk and Insurance* 78, 245-265.

[23] L. Goudenege, A. Molent, A. Zanette, (2015). Pricing and Hedging GLWB in the Heston and in the Black-Scholes with Stochastic Interest Rate Models, Working paper, http://arxiv.org/abs/1509.02686.

[24] L. Goudenege, A. Molent, A. Zanette, (2016). Pricing and Hedging GMWB in the Heston and in the Black-Scholes with Stochastic Interest Rate Models, Working paper, http://arxiv.org/abs/1602.09078.

[25] T. Haentjens, K. J. In 't Hout, (2012): Alternating direction implicit finite difference schemes for the Heston-Hull-White partial differential equation. *The Journal of Computation Finance* (83-110), Vol. 16, No. 1, Fall 2012.

[26] S. Heston, (1993): A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, Vol. 6, No. 2, pp. 327-343.

[27] D. Holz, A. Kling, J. Russ, (2007). GMWB for life: an analysis of lifelong withdrawal guarantees. Working paper.

[28] J. Hull, A. White, (1994). Numerical procedures for implementing term structure models I: single factor models. *The Journal of Derivatives Fall*, 716.

[29] B. Jourdain, (2013). Méthodes de Monte Carlo pour les processus financiers. ENPC, Marne-la-Vallée.

[30] A. Kling, F. Ruez, J. Russ, (2014). The impact of stochastic volatility on pricing, hedging, and hedge efficiency of variable annuity guarantees. *European Actuarial Journal*, Vol. 4, No. 2, pp. 281-314.

[31] B. Lapeyre, (2007). Introduction to Monte-Carlo Methods, Halmstad.

[32] P. L'Ecuyer, (1990). Random numbers for simulation. *Communications of the ACM*, 33(10).

[33] F. A. Longstaff, E. S. Schwartz, (2001). Valuing american options by simulation: a simple least-squares approach. *The Review of Financial Studies* Spring 2001, Vol. 14, No. 1, pp. 113-147.

[34] M. A. Milevsky, T. S. Salisbury, (2006). Financial valuation of guaranteed minimum withdrawal benefits. *Insurance: Mathematics and Economics* 38, pp. 21-38.

[35] D. B. Nelson, K. Ramaswamy, (1990). Simple binomial processes as diffusion approximations in financial models. *The Review of Financial Studies*, Vol. 3, No. 3, pp. 393-430.

[36] J. Neveu, (1975). Discrete-parameter Martingales. *North Holland*, Amsterdam.

[37] H. Niederreiter, (1995). New developments in uniform pseudorandom number and vector generation. *P.J.-S. Shiue and H. Niederreiter, editeurs*, Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing, volume 106 of Lecture Notes in Statistics, pages 87-120, Heidelberg, New York. Springer-Verlag.

[38] S. Ninomiya, N. Victoir, (2008). Weak approximation of stochastic differential equations and application to derivative pricing, *Applied Mathematical Finance*, Vol. 15, No. 2, pp. 107-121.

[39] V. Ostrovski, (2013). Efficient and exact simulation of the Hull-White model. Available at SSRN: http://ssrn.com/abstract=2304848 or http://dx.doi.org/10.2139/ssrn.2304848.

[40] D. W. Peaceman, H. H. Rachford, (1955). The numerical solution of parabolic and elliptic differential equations. *J. Soc. Indust. Appl. Math.*, 3 , pp. 28-41.

[41] B. D. Ripley, (1987). Stochastic Simulation. *Wiley* New York, Vol. 11, No. 237, pp. 31-85.

[42] R. Y. Rubinstein, (1981). Simulation and the Monte Carlo Method. *Wiley Series in Probabilities and Mathematical Statistics.*

[43] P. Shah, D. Bertsimas, (2008). An analysis of the guaranteed withdrawal benefits for life option. Working paper, Sloan School of Management, MIT.

[44] D. W. Stroock, S. R. Srinivasa Varadhan, (1979): Multidimensional Diffusion Processes, *Springer*, Berlin.

[45] D. Talay, L. Tubaro, (1990): Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Analalysis and Applications* Vol. 8 No.4 pp. 94-120.

[46] S. S. Yang, T. S. Dai, (2013). A flexible tree for evaluating guaranteed minimum withdrawal benefits under deferred life annuity contracts with various provisions. *Insurance: Mathematics and Economics*, Vol. 52, pp. 231-242.

[47] Information package of the product "Formula accumulation" of AXA, page 6. `http://www.axa.it/patrimonio/risparmio.aspx`

[48] AXA Life invest Website
`https://www.axa-lifeinvest.fr/fr`

[49] Borsa Italiana,
`http://www.borsaitaliana.it/notizie/finanza-personale/assicurazioni /dettaglio/assicurazioni-renditavariabile.htm`

[50] Think Advisor, (2014)
`http://www.thinkadvisor.com/2014/03/25/variable-annuity-sales-to-hit-22-billion-by-2018-c`