

Elsevier Editorial System(tm) for Journal of Computational and Applied Mathematics
Manuscript Draft

Manuscript Number:

Title: PDE formulation and resolution of a single factor Heath-Jarrow-Morton model with stochastic volatility

Article Type: Research Paper

Section/Category: 91-xx

Keywords: Quantitative finance, Computational finance, PDE stochastic formulation

Corresponding Author: Dr. eusebio valero, Ph.D.

Corresponding Author's Institution: Universidad Politécnica Madrid

First Author: eusebio valero, Ph.D

Order of Authors: eusebio valero, Ph.D; Manuel Torrealba, Ph.D; Lucas Lacasa, Ph.D; Francois Fraysse, Graduate

PDE formulation and resolution of a single factor Heath-Jarrow-Morton model with stochastic volatility

E. Valero¹, M. Torrealba², L. Lacasa³ and F. Fraysse¹

¹*Dpto. de Matemática Aplicada, ETSI Aeronáuticos, Universidad Politécnica de Madrid.*

²*Grupo BBVA, Mercados Globales y Distribución, Desarrollo de Nuevos Productos.*

³*IFISC, Instituto de Física Interdisciplinar y Sistemas Complejos, (CSIC-UIB), E-07122 Palma de Mallorca, Spain.*

Abstract

This paper considers the single factor Heath-Jarrow-Morton model for the interest rate curve with stochastic volatility. While from a financial point of view, the natural formulation of this model is in terms of stochastic differential equations, the Feynman-Kac theorem allows us to map the stochastic problem into a partial differential system. Accordingly, we numerically solve a three dimensional version of the HJM model, making use of finite difference schemes for the spatial discretization and Crank-Nicholson schemes for the temporal discretization, as well as Alternating Direction Implicit schemes for increasing the computational efficiency. Several numerical considerations such as convergence criteria or computation time are analyzed and discussed.

Keywords: Quantitative finance, Computational finance, PDE stochastic formulation

1. Introduction

The interest rate curve has been intensely studied and modelled through different techniques so far, assuming for this curve a temporal evolution which satisfies the no existence of arbitrage opportunities in a complete and efficient market. A market model of the interest rates curve must calibrate market prices of the most representative market instruments, these latter being the ones quoted more liquidly in the market. The simplest one is the *zero coupon bond* that, roughly speaking, is a financial instrument that pays one unit of currency at a certain future date (maturity). Another relevant set of instruments are the *caps* and the *floors*, which define implicitly the probability distribution of the rates. A cap pays the difference between a certain rate and a certain preespecified level (strike) if this difference is positive, while the floor pays off the difference between the strike and the value of the rate, if positive. The relation that links the cap or floor premium with the arbitrage-free probability distribution function of the rates is the well known Black-Scholes formula [3].

Focused in the calibration of zero coupon bond prices together with cap/floor prices, in this paper we address the Heath-Jarrow-Morton (HJM). Any model of the interest rate curve admits a representation in stochastic differential equations. Among others, a well known approximation to this problem is the so called HJM [1]. The main interest of the HJM approach is set on the fact that it provides a broad mathematical formulation [2], where most of market

observed features can be taken into account. This wasn't the case of preceding seminal models such as Black-Derman-Toy model or Hull-White models [3], where the model parameters were very difficult to calibrate in terms of market observed patterns. Those models only incorporated one source of stochasticity, and therefore the only possible mode in the interest rate curve was the parallel movement. Only lognormal or normal statistical distributions of the short rate were possible. These distributions for the short rate arise in distributions for different maturities of the libor rates, which are far away from those implied in the markets. Conversely, HJM models are not afflicted by these drawbacks. The HJM framework is general, in the sense that many previous models modelling the term structure of interest rates can be understood as particular cases of a HJM model, that may incorporate as many risk factors as needed in order to accurately describe the evolution of the rate's curves. The formulation of the HJM can be extended to several stochastic factors; then, the simulated interest rates curve movements could include deformation modes that changes the initial slope and convexity of this curve, in order to describe the covariance and autocorrelation structures present in the time behaviour of the interest rate curve. Additionally, the probability distribution function of interest rates can be exogeneously defined by means of local volatility functions or even stochastic processes in order to match option prices as quoted in the market.

The major drawback of HJM models is set on the continuous nature of its state variables (the continuous time structure of forward rates), what leads to an infinite amount of state variables. Therefore, in general, this model is non-Markovian in finite dimension (and thus can only be solved by using fitted Monte Carlo techniques which are eventually slow and computationally delicate). Quite interestingly, HJM models can be in some particular situations transformed into a low order Markovian system [4, 5, 6, 7], and can be mapped to a partial differential system due to the well known Feynman-Kac formula [9]. Numerical methods for solving partial differential systems can consequently apply [13].

The preliminary HJM models used to model the stochastic behavior of the interest rate market by a single Wiener process that drives the forward rate processes (that is to say, there was a single source of stochasticity in the models). In the last years, some authors have introduced, according to empirical evidence, a new source of stochasticity in the description of the volatility evolution, leading to the so called stochastic volatility HJM models [8]. In this paper, we address a stochastic volatility HJM model in its three dimensional version, where the Feynman-Kac formula applies. We make use of PDE-computational techniques to efficiently solve the latter model. We will present a complete numerical methodology based on finite differences schemes, temporal integrators and Alternating Direct Implicit (ADI) schemes that will solve the PDE system efficiently and we will analyze the computational performance accordingly.

The rest of the paper goes as follows: Section 2 is focused on the interest rate market, defining the zero coupon bond as the trading derivative. Then, section 3 presents the HJM framework, in the stochastic volatility model and in the PDE hallmark. Section 4 and 5 describe the numerical method. The numerical validation is depicted in section 6. Finally, after some additional remarks concerning the numerical method (section 7), in section 8 we will point out some conclusions.

2. Interest rate market: the zero coupon bond

In the interest rate market, the zero coupon bonds are taken as the most basic market instruments in the sense that any other financial quantity related to interest rates can be derived from them. These assets merely pay a monetary unit in a given future time that is called the expiration date, and the current price as a function of the expiration date is determined by the so called zero coupon bond curve (ZCBC). We will formally denote this curve as it follows:

$$p(t, T), t \in \mathbb{R}^+, T \geq t : T \rightarrow p(t, \cdot) \quad (1)$$

where t stands for the present time (valuation date), T stands for the bond expiration date and $p(t, T)$ is merely the zero coupon bond price (ZCBP). We will also assume from now on that ZCBC fulfills the necessary regularity conditions.

By definition, ZCBC is such that:

- the ZCBP that expires at the present time is 1 (trivially): $p(T, T) = 1$.
- ZCBP $\in (0, 1]$ and is a monotonically decreasing function (this property is assumed in order to avoid the existence of arbitrage).

It is worth saying that in exceptional macroeconomic situations, it is actually possible that the latter property doesn't hold, due to the intervention of central banks: this was indeed the case of Switzerland, in order to avoid the revaluation of the Swiss franc.

Provided the preceding properties, there always exists a function $f(t, s), s \geq t$ such that

$$\begin{aligned} p(t, T) &= e^{-\int_t^T f(t, s) ds}, t \in \mathbb{R}^+, T \geq t \\ f(t, s) &\geq 0, \forall s \geq t \end{aligned} \quad (2)$$

and consequently, f fulfills:

$$f(t, T) = -\frac{\partial \ln p(t, T)}{\partial T}, \forall T \geq t. \quad (3)$$

From a financial point of view, $f(t, T)$ is interpreted in terms of the interest rate that an investor would receive if he sells in t a zero coupon bond with expiration date T and buys another with expiration date $T + dT$. This bond is the so called forward rate of interest. In particular, the short term rate of interest r fulfills:

$$r(t) = f(t, t) = -\frac{\partial \ln p(t, T)}{\partial t} \Big|_{T=t}, \quad (4)$$

and is of special interest because it quotes the return of an investment of one monetary unit in the present time, t , that is redeemed an infinitesimal time later, $t + dt$.

3. HJM framework

3.1. The model

In a single factor HJM model of the interest rate curve, it is possible to write the evolution of the ZCBP as a function of one stochastic process. In the case of the single factor HJM model, the ZCBP takes the form [2]:

$$\begin{aligned} p(t, T) &= \frac{p(0, T)}{p(0, t)} e^{-G(T-t)x(t) - \frac{1}{2}G(T-t)^2 y(t)} \\ G(s) &= \frac{1 - e^{-\kappa s}}{\kappa} \end{aligned} \quad (5)$$

where κ is a positive constant, $x(t)$ is a stochastic process and $y(t)$ a deterministic one. These two processes are modelled in terms of the following set of differential equations

$$\begin{aligned} dx(t) &= [-\kappa x(t) + y(t)]dt + \eta(t, x(t))dW(t) \\ dy(t) &= [\eta(t, x(t))^2 - 2\kappa y(t)]dt \end{aligned} \quad (6)$$

where $dW(t)$ is a Wiener process. The volatility function $\eta(t, x(t))$ fully describes the model and it will be properly defined further in the text.

Following the notation, the short rate r would fulfill:

$$r(t) = -\left. \frac{\partial \ln p(t, T)}{\partial T} \right|_{T=t} = f(0, t) + x(t), \quad (7)$$

and consequently would fulfill the following stochastic differential equation (SDE):

$$dr(t) = \left(\frac{\partial f(0, t)}{\partial t} - \kappa(r(t) - f(0, t)) + y(t) \right) dt + \eta(t, r(t))dW(t). \quad (8)$$

The volatility function $\eta(t, r(t))$ can thus be defined as the one that fulfills the following SDE:

$$\begin{aligned} \eta(t, r(t)) &= \sqrt{v(t)}\lambda(t)r(t)^{\gamma(t)} \\ dv(t) &= \theta(1 - v(t))dt + \epsilon(t)\sqrt{v(t)}dZ(t) \\ dZ(t) \cdot dW(t) &= \rho dt, \end{aligned} \quad (9)$$

where $dZ(t)$ is another Wiener process, $\lambda(t)$ and $\gamma(t)$ are deterministic functions of time ($\gamma(t) \in (0, 1]$), $v(t)$ is a stochastic variable that drives the rate variance, θ is a constant that estimates the mean reversion speed of the process $v(t)$, ρ is a constant that estimates the correlation between the short rate and the volatility (hence $\rho \in [-1, 1]$), and $\epsilon(t)$ is the volatility associated to $v(t)$, which in this case is a deterministic function of time.

The HJM model is thereby fully characterized. The state variables belong to the following range:

$$r(t) \in [0, \infty), \quad v(t) \in [0, \infty), \quad y(t) \in [0, \infty)$$

3.2. Model reformulation in terms of Partial Differential Equations

Given the model formulation, now we would like to price contracts whose future payoffs depend on the evolution of the yield curve, that is to say, those payoffs are deterministic functions of the ZCBP at certain times. Let $C(t)$ be the price in a given time t , of a contract that pays in $T > t$ in terms of $F(T, p(T, s))$, $s \geq T$, where P denotes the payoff function. The simplest example of this situation is the zero coupon bond whose payoff function is a constant function, $F(T, p(T, s)) = 1$. The so called Caplet [3] is another example of a payoff function, frequently traded in the market:

$$F(T, p(T, T_M)) = \max(1 - \Delta_M p(T, T_M), 0) \quad (10)$$

where T is the contract's expiration date, T_M is the contract's payment date and $\Delta_M = 1 + (T_M - T)K$, where K stands for the strike and is a positive constant. In this paper we focus on both the Zero Coupon Bond and the Caplet as the derivatives under study. In these terms, the price of an interest rate derivative depends on time and space $C(t, x, y, v)$ and fulfills the following partial differential equation:

$$\frac{\partial C}{\partial t} + \zeta_{rr} \frac{\partial^2 C}{\partial r^2} + \zeta_{vv} \frac{\partial^2 C}{\partial v^2} + \zeta_{rv} \frac{\partial^2 C}{\partial r \partial v} + \mu_r \frac{\partial C}{\partial r} + \mu_v \frac{\partial C}{\partial v} + \mu_y \frac{\partial C}{\partial y} = rC \quad (11)$$

where the Feynman-Kac formula [9] is applied to map the stochastic problem into a PDE one. Note that the coefficients ζ_{rr} , ζ_{rv} , ζ_{vv} , μ_r , μ_v , μ_y are functions of the space variables r, v, y and time, but their dependence has been omitted for clarity. These coefficients are given by

$$\begin{aligned} \zeta_{rr} &\equiv \frac{1}{2} \lambda(t)^2 r^{2\zeta(t)} v \\ \zeta_{vv} &\equiv \frac{1}{2} \epsilon(t)^2 v \\ \zeta_{rv} &\equiv \lambda(t) r^{\zeta(t)} \epsilon(t) \rho v \\ \mu_r &\equiv \frac{\partial f(0, t)}{\partial t} - \kappa(r - f(0, t)) + y \\ \mu_v &\equiv \theta(1 - v) \\ \mu_y &\equiv \lambda(t)^2 r^{2\zeta(t)} v - 2\kappa y \end{aligned} \quad (12)$$

The boundary/terminal conditions of the problem are different depending on the payoff function. For the cases of the zero coupon bond and the caplet, these are given by:

- *Zero coupon bond with expiration date T* : The terminal condition simply reads

$$C(T, r, y, v) = 1 \quad (13)$$

In the limit $r \rightarrow \infty$ the payoff is zero, and consequently one boundary condition is:

$$C(t, r \rightarrow \infty, y, v) = 0, \quad t < T \quad (14)$$

Now, when $r = 0$ the PDE reduces to

$$\begin{aligned} \frac{\partial C}{\partial t} + \zeta vv|_{r=0} \frac{\partial^2 C}{\partial v^2} + \mu_r|_{r=0} \frac{\partial C}{\partial r} + \mu_v|_{r=0} \frac{\partial C}{\partial v} + \mu_y|_{r=0} \frac{\partial C}{\partial y} &= 0 \\ C(T, r \rightarrow 0, y, v) &= 1, \end{aligned} \quad (15)$$

that can be solved analytically checking solutions of the shape $C(t, x, y, v) = c(t)e^{a(t)r+b(t)y}$ or using numerical methods alternatively.

In the boundary $y \rightarrow \infty$ the PDE reads

$$\begin{aligned} \frac{\partial C}{\partial t} - \kappa y \frac{\partial C}{\partial r} - 2\kappa y \frac{\partial C}{\partial y} &= 0 \\ C(T, r, y \rightarrow \infty, v) &= 1 \end{aligned} \quad (16)$$

that is also solvable. Finally, note that the boundaries $v \rightarrow \infty$ and $v \rightarrow 0$ are not relevant in this case as long as the zero coupon bond price is not a function of v .

- *Caplet*: The terminal condition is given by

$$C(T, r, y, v) = \max(1 - \Delta_M p(T, T_M; r, y, v), 0) \quad (17)$$

Just as in the case of the zero coupon PDE, the boundary condition for $r \rightarrow \infty$ is

$$C(t, r \rightarrow \infty, y, v) = 0, \quad t < T. \quad (18)$$

When $r = 0$ the PDE reduces to:

$$\begin{aligned} \frac{\partial C}{\partial t} + \zeta vv|_{r=0} \frac{\partial^2 C}{\partial v^2} + \mu_r \frac{\partial C}{\partial r} + \mu_v|_{r=0} \frac{\partial C}{\partial v} + \mu_y|_{r=0} \frac{\partial C}{\partial y} &= 0 \\ C(T, 0, y, v) &= \text{Max}(1 - \Delta_M p(T, T_M; 0, y, v), 0) \end{aligned} \quad (19)$$

In the limit $y \rightarrow \infty$, the PDE and its boundary condition have the following shape:

$$\begin{aligned} \frac{\partial C}{\partial t} + y \frac{\partial C}{\partial r} - 2\kappa y \frac{\partial C}{\partial y} &= 0 \\ C(T, r, y \rightarrow \infty, v) &= 0 \end{aligned} \quad (20)$$

Thereby we have:

$$C(t, r, y \rightarrow \infty, v) = 0, \quad t < T. \quad (21)$$

In the case $v \rightarrow \infty$ we will have

$$C(T, r, y, v \rightarrow \infty) = p(t, T) - p(t, T_M), \quad (22)$$

and when $v = 0$ the PDE reduces to

$$\begin{aligned} \frac{\partial C}{\partial t} + \mu_r|_{v=0} \frac{\partial C}{\partial r} + \theta \frac{\partial C}{\partial v} - 2\kappa y \frac{\partial C}{\partial y} &= rC \\ C(T, r, y, v \rightarrow 0) &= \text{Max}(1 - \Delta_M p(T, T_M; r, y, 0), 0) \end{aligned} \quad (23)$$

that in this case can only be solved numerically.

Additionally, when $v = 0$ the following identity is commonly assumed to hold:

$$\left. \frac{\partial^2 C}{\partial v^2} \right|_{v \rightarrow 0} = 0,$$

and will be taken into account in the numerical development.

3.3. Model parameters

In order to fully specify the preceding PDE, the constants and functions implicitly defined in the HJM model have to be defined. As a reference guide, and for the sake of an order of magnitude estimation, the function's variation range and the set of parameters are set as:

-Initial zero coupon curve:

- $p(0, T) = 1.04^{-T}$

-Volatility function:

- $\kappa = 0.001$

- $\lambda(t) = 0.15$

- $\gamma(t) = 0.9$

- $\epsilon(t) = 1.5$

- $\theta = 0.25$

- $\rho = -0.75$

4. Numerical schemes: preliminary insight

Before solving numerically equation (11) we need to make some preliminary analysis, in order to optimize the numerical approach and reduce its computational cost. Three points are of major importance, namely (i) the study of some analytical solutions will provide information about the system's solution itself, (ii) a detailed scale analysis of the problem and (iii) some considerations regarding the metric need also be addressed.

4.1. Particular solutions

The first case considers the solution of the zero coupon bond curve with expiration date T . As far as this solution does not depend explicitly on v , we assume $v = 0$ without lack of generality in order to simplify the system of equations. Furthermore, constant κ is generally small (volatility function depends on κ) and we have thus assumed that it is also null. Hence, the partial differential equation reduces to:

$$\frac{\partial C}{\partial t} + \left(\frac{\partial f(0, t)}{\partial t} + y \right) \frac{\partial C}{\partial r} = rC(t, r, y) \quad (24)$$

with the additional condition $C(T, r, y) = 1$. We can trivially map this PDE into a system of ordinary differential equations of the following shape:

$$\frac{dt}{d\tau} = 1, \quad \frac{dr}{d\tau} = \frac{\partial f(0, \tau)}{\partial \tau} + y, \quad \frac{dC}{d\tau} = rC. \quad (25)$$

with initial conditions $\tau = 0$, $t = T$, $r = r_1$, $C = 1$, and y being a parameter. It is indeed easy to check that its solution is:

$$C(t, r, y) = e^{-\int_t^T f(0, s) ds} e^{-(T-t)(r-f(0,t)) - \frac{(T-t)^2}{2} y}. \quad (26)$$

Note that not assuming a null value for κ is equivalent to substitute $T - t$ by the function $G(T - t)$, which is defined in (5).

The usefulness of this analytical solution is twofold: first, it will serve to validate the numerical method, and second, it will stand as a boundary condition for the (more general) Caplet problem.

4.2. Scale analysis

In the numerical realm, reducing the computing time as well as the computational cost (in terms of memory resource, for instance) is fundamental. An adequate temporal and spatial scale analysis will enable us to increase the mesh resolution only where/when needed, what leads to a saving of computational resources. Typically, this scale analysis is realized by adimensionalizing the equations under study and consequently comparing the relevance of the respective terms (this technique is broadly used in fluid mechanics when performing the scale analysis of the Navier-Stokes equations, for instance [17]).

Scales are indeed determined by the variables characteristic values, as well as by the boundary conditions. In the case of the Cap problem, the variable of reference is the interest $r(t)$ ($r \in [0, 100]$, that is to say, a percentage). We can rescale this variable as:

$$\tilde{r} = r/r_0, \quad r_0 \simeq 10^{-2}$$

in such a way that the characteristic value of \tilde{r} is the unity (r_0 is usually the forward rate of interest observed at value date). Now, having in mind that the Cap's boundary condition reads

$$C(T, 0, y, v) = \max(1 - (1 + (T_M - T)r_0\tilde{K})p(T, T_M; 0, y, v), 0),$$

where $\tilde{K} = K/r_0$, we may define the characteristic time \tilde{t} as:

$$\tilde{t} = r_0 t.$$

Finally, taking into account the relation between r and y (equation 25):

$$\Delta r \simeq y \Delta t, \rightarrow \tilde{y} = y/r_0^2.$$

With these rescaled variables, the original equation takes the following shape:

$$\begin{aligned} r_0 \frac{\partial C}{\partial \tilde{t}} + f_1(t, r, v) \frac{\partial^2 C}{\partial r^2} + f_2(t, v) \frac{\partial^2 C}{\partial v^2} + f_3(t, r, v) \frac{\partial^2 C}{\partial r \partial v} + \\ + f_4(t, r, v, y) \frac{\partial C}{\partial r} + f_5(v) \frac{\partial C}{\partial v} + f_6(t, r, v, y) \frac{\partial C}{\partial y} = r r_0 C \end{aligned} \quad (27)$$

where

$$f_1(t, r, v) = \frac{1}{2}\lambda(t)^2 r^{2\gamma(t)} r_0^{2(\gamma(t)-1)} v, \quad f_2(t, v) = \frac{1}{2}\epsilon(t)^2 v,$$

$$f_3(t, r, v) = \lambda(t) r^{\gamma(t)} \epsilon(t) \rho v r_0^{\gamma(t)-1},$$

$$f_4(t, r, v, y) = \left(r_0 \frac{\partial f(0, t)}{\partial t} - \kappa(r - f(0, t)) + r_0 y \right),$$

$$f_5(v) = \theta(1 - v), \quad f_6(t, r, v, y) = (\lambda(t)^2 r^{2\gamma(t)} v - 2\kappa y).$$

(Note that the $\tilde{\sim}$ marks have been eliminated for notation simplicity). No rescaling has been applied to both v and $C(t, r, v, y)$ as long as there's no dominant scale defined (the latter rescaling wouldn't affect the resultant equation).

4.3. Metrics

By introducing metrics in the independent variables (r, v, y) , we can transform the problem's domain into a computational domain which is usually simpler, and consequently concentrate the mesh points in the areas under study. In our case of study (Cap), provided that the solution is likely to live in the strike's neighborhood (K), this should be the most dense zone. Following Tavella [16], we have used an hyperbolic-like metric generally defined as:

$$z = K + \alpha \sinh(c_2 x + c_1(1 - x)),$$

$$c_1 = \arg \sinh \left(\frac{(z_0 - K)}{\alpha} \right), \quad c_2 = \arg \sinh \left(\frac{(z_\infty - K)}{\alpha} \right),$$

$$x \in [0, 1], z \in [z_0, z_\infty], \text{ being } z = (r, v, \text{ or } y). \quad (28)$$

In order to get a mesh that is accurate enough in the zone under study, the parameters α, K, z_0, z_∞ have to be correctly chosen for each dimension of the problem.

It is worth saying that the introduction of these kind of metrics doesn't modify the system of equations in a substantial manner. Note that the derivatives with respect of the new variables can be expressed as:

$$\frac{\partial C}{\partial r} = \frac{\partial C}{\partial x} \frac{1}{\partial r / \partial x}$$

and

$$\frac{\partial^2 C}{\partial r^2} = \frac{\partial^2 C}{\partial x^2} \frac{1}{(\partial r / \partial x)^2} - \frac{\partial C}{\partial x} \frac{\partial^2 r / \partial x^2}{(\partial r / \partial x)^3}$$

Thereby, if we denote $J_z = \partial z / \partial x$, and $J_{2z} = \partial^2 z / \partial x^2$, where z is a generic variable from (r, v, y) and x its respective transform, the original system of equations (27) will only differ from the new one in the substitution of functions f_i , $i = 1..5$ by:

$$g_1(t, x_r, x_v) = f_1(t, x_r, x_v) / J_r^2, \quad g_2(t, x_v) = f_2(t, x_v) / J_v^2,$$

$$g_3(t, x_r, x_v) = f_3(t, x_r, x_v)/J_r/J_v, \quad g_4(t, x_r, x_v, x_y) = f_4(t, x_r, x_v, x_y)/J_r - f_1(t, x_r, x_v)J_{2r}/J_r^3, \\ g_5(t, x_v) = f_5(x_v)/J_r - f_2(t, x_v)J_{2v}/J_v^3, \quad g_6(t, x_r, x_v, x_y) = f_6(t, x_r, x_v, x_y)/J_y.$$

While up to now we have only defined metrics in the direction of the independent coordinates, it is actually possible to employ more complex transformations that involve several variables ($\xi = \xi(r, v, y)$). However, it is likely that the integration domain wouldn't in that case be cartesian anymore, and consequently the finite difference scheme wouldn't apply. Furthermore, while these types of metrics would eventually enable us to eliminate the cross derivative terms in (27) (transforming the original equation into its canonical form), this transformation would on the other side modify the frontiers of the problem from straight to curve lines, something that not desirable in any case.

Once the preliminary insights have been put forward, we will describe in the next sections the numerical method employed to integrate equation (27) as well as the results that we have obtained.

5. Numerical methods

As commented above, the equation under hands is (27):

$$r_0 \frac{\partial C}{\partial t} + g_1(t, r, v) \frac{\partial^2 C}{\partial r^2} + g_2(t, v) \frac{\partial^2 C}{\partial v^2} + g_3(t, r, v) \frac{\partial^2 C}{\partial r \partial v} + \\ + g_4(t, r, v, y) \frac{\partial C}{\partial r} + g_5(v) \frac{\partial C}{\partial v} + g_6(t, r, v, y) \frac{\partial C}{\partial y} = r r_0 C. \quad (29)$$

Let us define a new temporal variable

$$t = T - t$$

where T stands for the maturity. The integration should then be done for $t \in [0, T]$. Note that equation (27) is parabolic for r and v and hyperbolic for y .

In the particular problems concerning the estimation of financial derivatives, the execution time of numerical tools is an issue of fundamental importance. According to this fact, it seems suitable to apply second order schemes for the discretization of both temporal and spatial partial derivatives, as far as these schemes show optimal computational cost and adequate precision. In a second step, one has to decide whether to apply explicit or implicit schemes. Naturally, the simplest option is always to tackle explicit schemes, which are fast and easy to implement. However, it is easy to check that due to the second derivatives, the following relation holds:

$$\Delta t \simeq (\Delta r)^2.$$

This relation implies that in order to achieve a precision of say 10^{-4} in the solution, a time step would need 10^4 iterations. Moreover, the coefficients of the derivatives are powers of r , v or y , and the integration domain ranges to the infinite. Since the time step is inversely proportional to those coefficients, the problem comes to be even more delicate. We can thus conclude that explicit temporal schemes won't fit in this case due to their inevitably lengthy

behavior. Thereby, we will have to choose implicit schemes for the temporal integration. These have the following general expression:

$$\frac{dC}{dt} = F(C) \rightarrow C^{n+1} - C^n = \Delta t (\theta F(C^{n+1}) + (1 - \theta)F(C^n)),$$

where θ stands for an explicit scheme for $\theta = 0$ (Euler scheme) while it stands for implicit schemes when $\theta \neq 0$. More concretely, $\theta = 1$ characterizes the so called Euler implicit scheme and finally $\theta = 1/2$ characterizes the second order Crank-Nicholson scheme. We will use the latter one as the temporal integrator since they are adequate to use with ADI schemes (this will be explained further in the text).

The spatial derivatives (first and second order) will be discretized by centered finite difference schemes as it follows:

$$\delta_{xx}u \equiv \frac{\partial^2 u}{\partial x^2} = \frac{u_{i+1} - 2u_i + u_{i-1}}{(\Delta x)^2}, \quad \delta_x u \equiv \frac{\partial u}{\partial x} = \frac{u_{i+1} - u_{i-1}}{(\Delta x)},$$

where x is a generic variable that stands for r , v or y . The notation $\delta_{xx}u$ and $\delta_x u$ describes a difference scheme. Gathering both spatial and temporal schemes, we come to a final discretization of the following kind:

$$U^{n+1} - U^n = \Delta t(\theta F(U^{n+1}) + (1 - \theta)F(U^n)) \quad (30)$$

where

$$F(U) = g_1(t, r, v)\delta_{rr}U + g_2(t, v)\delta_{vv}U + g_3(t, r, v)\delta_{rv}U \\ + g_4(t, r, v, y)\delta_r U + g_5(v)\delta_v U + g_6(t, r, v, y)\delta_y U - rr_0 U,$$

$$\text{and } U = \{C_{ijk}, i = 0..nr, j = 0..nv, k = 0..ny\}, \quad (31)$$

is the discretized solution vector in a structured mesh of dimension (nr, nv, ny) .

It is worth saying that the use of centered difference schemes allows us to obtain a compact stencil. For instance, note that the discretized equation in the point C_{ijk} only contains information of $\{C_{i+1,j,k}, C_{ijk}, C_{i-1,j,k}\}$. Focusing on variable r , equation (30) would adopt the shape:

$$U^{n+1} - U^n = \Delta t((\theta L_r U^{n+1} + b) + (1 - \theta)(L_r U^n + b)),$$

where L_r is a tridiagonal matrix [13]. Tridiagonal systems are indeed quite easy to implement and solve (for instance, the Thomas algorithm [18] solves a tridiagonal system in $7N$ operations, where N is the order of the system). This goodness will enable the use of Alternating Direction Implicit schemes (ADI) [10, 13] as will be shown further in the text.

Finally, taking into account that the equation is indeed linear, it can be written as:

$$U^{n+1} - U^n = \theta \Delta t [L_r(t^n, r_i, v_j, y_k)U^{n+1} + L_v(t^n, v_j)U^{n+1} + L_y(t^n, r_i, v_j, y_k)U^{n+1}] \\ + (1 - \theta) \Delta t [L_r(t^n, r_i, v_j, y_k)U^n + L_v(t^n, v_j)U^n + L_y(t^n, r_i, v_j, y_k)U^n] \\ + \Delta t L_{rv}(t^n, r_i, v_j)U^n + \Delta t B(t^n, r_i, v_j),$$

and realigning,

$$\begin{aligned}
& [I - \theta\Delta t(L_r(t^{n+1/2}, r_i, v_j, y_k) + L_v(t^{n+1/2}, v_j) + L_y(t^{n+1/2}, r_i, v_j, y_k))] (U^{n+1} - U^n) \\
= \Delta t [L_r(t^n, r_i, v_j, y_k) + L_v(t^n, v_j) + L_y(t^n, r_i, v_j, y_k) + L_{rv}(t^n, r_i, v_j)] U^n + \Delta t B(t^n, r_i, v_j) \\
& = \Delta t F(U^n, t^n, r_i, v_j, y_k). \quad (32)
\end{aligned}$$

Note that the operators L_r, L_v, L_y include the terms related to the spatial discretization. These operators, treated implicitly, give rise to a system of equations $(I - \theta(L_r + L_v + L_y))U = f$ which in general has 7 diagonals, and whose resolution can be performed applying either direct or iterative methods. However, each one of them treated separately can be rewritten as a tridiagonal matrix, whose resolution is trivial as commented above. Note also that the operators L_r, L_v, L_y have Neumann boundary conditions and consequently do not include any Dirichlet-like information.

Finally, note that the mixed derivative term is only treated explicitly, because otherwise its inclusion in the implicit scheme would eliminate the tridiagonal structures, and would consequently avoid the use of ADI schemes that will be described in what follows. This fact does not affect in any case neither to the convergence nor the precision of the numerical solution.

5.1. Alternating Direction Implicit (ADI) schemes

The ADI schemes belong to the category of Splitting methods [10, 11, 13, 12, 14, 15], used in the resolution of multidimensional PDE systems. The key idea behind these methods is to separate the original multidimensional problem in several unidimensional split problems. Then, each split problem can be under certain conditions reduced to the resolution of a tridiagonal system of equations. These conditions are related to the use of centered spatial operators in structured meshes, which is our case.

ADI schemes were initially introduced by Douglas, Peaceman and Rachford [10, 11] in order to integrate, using finite difference schemes, the well known Navier-Stokes equations describing the fluid motion. Some modifications have been put forward so far (see for instance [12, 14, 15, 13], in order to apply these schemes to either stationary or non stationary problems. In this work we will use an ADI scheme recently put forward by Hout & Welfert [15], called the Douglas scheme.

Consider equation (32), this one can be formally written as:

$$[I - \theta\Delta t(L_r + L_v + L_y)]\Delta U^n = \Delta t F(U^n)$$

The Douglas scheme applies thus in the following way:

$$\begin{aligned}
\Delta U^0 &= \Delta t F(U^n) \\
[I - \Delta t \quad L_r \quad \Delta \tilde{U}^1] &= \Delta U^0 \\
[I - \Delta t \quad L_v \quad \Delta \tilde{U}^2] &= \Delta \tilde{U}^1 \\
[I - \Delta t \quad L_y \quad \Delta \tilde{U}^n] &= \Delta \tilde{U}^2
\end{aligned}$$

$$U^{n+1} = U^n + \Delta \tilde{U}^n$$

Note that each step only requires the resolution of a tridiagonal system of dimension nr , nv , or ny . The unconditional convergence [13] of this scheme has been proved for $\theta = 0.5$ (Crank-Nicholson) in 2-dimensional systems with constant coefficients. However no similar study has been performed so far in the 3-dimensional case with variable coefficients [15], which is nonetheless our case. Special attention will be thus paid to the convergence behavior of the solution.

It is worth saying at this point that the Craig & Sneyd method is an apparent improvement to the Douglas scheme (in terms of the solution precision), while being more expensive computationally speaking. We actually have also tackled this ADI scheme, but given that no such improvement has been observed, we will only focus on the Douglas scheme.

6. Validation

We have done two kind of studies in order to validate the numerical methods:

- (i) first, we have compared the numerical solution of the zero coupon curve or deduction curve with its analytical solution, in two different situations, and
- (ii) second, we have compared the numerical solution of a Caplet with the one obtained by a 2D-Heston model [13].

6.1. Zero coupon curves

In the first study we consider the parameters and the initial zero coupon curve depicted in section 3.3. Thereby, the forward rate of interest is constant $f(0, t) = \log(1.04)$ and it has a null derivative. In figures (1-2) we compare the theoretical zero coupon curve with the one obtained through numerical simulations with $r = \log(1.04)$, $y = 0$, and $t = T$. Concretely, figure (1) shows the error for different meshes. Note that this one is always below 10^{-5} even for coarse meshes. As a result, we have set the mesh reference values to 100×40 . In figure (2) we plot the convergence of the solution as a function of the number of time steps per year. Notice that from 12 steps per year, in a given mesh the variations are quite small ($O(10^{-7})$).

In figures (3) and (4) we show the error's spatial distribution for a given set of parameters whether assuming $y_\infty = 25$ (figure (3)) or $y_\infty = 250$ (figure (4)). Note that in the former case, some non desirable errors take place in the infinite boundary, which can actually propagate into the zone under study $y \simeq 0$ (figure (3)).

Up to know the numerical method is validated, as far as the solution's error is confined, in the zone under study, around 10^{-7} . However, as long as the analytical solution strongly depends on the initial curve, it is necessary to check whether if the precision of the numerical method holds for more realistic curves (with non null forward rate interest curve derivative). For that task, in a second example we tackle a new initial zero coupon curve, which is not anymore a continuous curve but a discrete valued one (figure (5)). Its derivative is plotted in figure (6) and stands for the forward rate of interest curve, and its second derivative is plotted in figure (7).

As long as the curve is expressed in terms of discrete values, we need to perform a smoothing approximation in order to introduce it in the simulation. Notice that the solution's smoothness will strongly depend with the smoothness of this initial curve (this is due to the

fact that the temporal derivative of the solution is related to the derivative of the forward rate of interest curve). An appropriate solution to this problem is to approximate the initial curve with splines [18], in order to have a piecewise function with continuous second derivative $df(0, T)/dT$, and consequently have a solution with continuous temporal derivative.

According to this approximation, we have performed the same simulations and analysis as for the first study. Conclusions are plotted in figures (8-10). As expected, the fact that the derivative of the forward rate of interest curve is non null has a net effect in the precision of the solution. While it is quite easy to achieve convergence of order $O(10^{-5})$, it comes necessary to overrefine the mesh (figure (8)) or alternatively increase the number of time steps (figure (9)) in order to go beyond 10^{-6} . Nevertheless, as is shown in figure (10), the error's spatial distribution is quite similar to the one found in the first study: we can conclude that the numerical method correctly reproduces the expected results.

6.2. Caplet

The second validation test consists in making a comparison between the results obtained with several Caplets and those obtained by the Heston model, which is an already validated model [13].

In order to optimize the mesh's size, we have performed a previous analysis of the numerical scheme's convergence (both in spatial and temporal discretizations). Some of the results are plotted in figure (11), where we represent the evolution of a generic Caplet's prime as a function of nr , nv , ny and nt . Notice that we need at least a mesh size of $100 \times 40 \times 40$ if we seek variations of the prime below 10^{-5} . With a reference mesh of $100 \times 40 \times 40$, the number of time steps does not affect practically the results.

The difference between the solutions of the two models are plotted in figures (12) and (13), both for the premium and for the volatility.

7. Some additional numerical aspects

7.1. Metric choice

As commented above, it is highly recommendable to introduce a metric layer in the numerical method, such that the domain under study transforms into a computational domain which is typically easier to handle (in most cases, this one is the unity cube), as long as this domain enables the use of structures equispaced meshes, where one can concentrate the mesh points wherever needed. The mesh that has been used in this work is hyperbolic (see eq. 28), following Tavella [16]. One of the main properties of these meshes is that one can concentrate as many points as needed in the inner regions of the zone under study, in order to achieve a better resolution. It is thus convenient to fix the parameters related to the domain transformation. For instance, r will transform according to:

$$r = K_r + \alpha_r \sinh(c_{2r}x_r + c_{1r}(1 - x_r)),$$

$$c_{1r} = \operatorname{asinh}\left(\frac{(r_0 - K_r)}{\alpha_r}\right), \quad c_{2r} = \operatorname{asinh}\left(\frac{(r_\infty - K_r)}{\alpha_r}\right),$$

where the jacobian of the transformation reads:

$$\frac{dr}{dx_r} = \alpha_r \sinh(c_2 x_r + c_1(1 - x_r))(c_{2r} - c_{1r}).$$

We have then four parameters $K_r, \alpha_r, r_0, r_\infty$ to fix:

- r_0, r_∞ define the real domain of study. Obviously $r_0 = 0$. On the other side, r_∞ must be such that his values doesn't modify the solution in the zone under study (that is, close to the strike). There is no recipe in order to find the adequate value, but after some preliminary estimations and taking into account the boundary conditions, we have set $r_\infty = 250$.
- K_r defines the region under study, that is, a neighborhood of the strike.
- α_r This parameter provides a measure of the mesh's stretching, i.e. the number of points that will be concentrated in the zone of interest -close to the strike-. Concretely, the smaller α_r , the larger concentration. Given that its value also affects the jacobian of the transformation, it is desirable that α_r is such that the jacobian be close to 1. In figure (14) we plot this dependence, for $r = 1(r_0)$. Note that for $\alpha_r \simeq 0.05 - 0.1$, the jacobian reaches the unity. A similar study for v e y lead us to fix $\alpha_v \simeq 0.5$ y $\alpha_y \simeq 0.05$.

As a summary, the metric's characteristic values are:

$$r_0 = v_0 = y_0 = 0, \quad r_\infty = y_\infty = 250, v_\infty = 30,$$

$$\alpha_r = \alpha_y = 0.05, \quad K_r = \text{Strike}, K_y = 0, K_v = \text{input}(\simeq 0.5)$$

7.2. Softening of the initial condition

Following Tavella [16], as far as the payoff is typically a discontinuous function, small variations in the strike lead to a non smooth behavior of the solution. This is not desirable and therefore some numerical techniques should be applied in order to soften it:

- Perform a dynamical modification of the mesh, related to the payoff's shape. This is an elegant solution, however for practical purposes this technique is not well fitted as long as it usually leaves to mesh interpolation.
- Soften the initial/final conditions. In order to do so, one can define an average initial/final condition in the following terms:

$$C_{ijk}(t, r, v, y) = \frac{1}{\omega} \int_{\omega} C(t, r, v, y) d\omega,$$

where ω is a control element centered in the mesh point r_i, v_j, y_k . The effect of this average is represented in figures (15-16).

7.3. Boundary condition for variable y

Note that the HJM model under hands is only convective for variable y (first derivatives are null for every variable but y). When $y = 0$, the solution's characteristic crosses the domain, what indicates that the boundary takes some information from inside. Consequently, the discretization of both the interior and the boundary should be consistent, and then a second order scheme should be applied to the boundary discretization. We have implemented two different possibilities in the numerical scheme:

- Advanced first order differences:

$$\frac{\partial C}{\partial y} \simeq \frac{C_{i,j,1}^n - C_{i,j,0}^n}{\Delta y}$$

- Advanced second order differences:

$$\frac{\partial C}{\partial y} \simeq \frac{-C_{i,j,2}^n + 4C_{i,j,1}^n - 3C_{i,j,0}^n}{2\Delta y}.$$

and quite surprisingly, no significative differences have been found between both schemes results.

7.4. Computational efficiency

In the following table we have represented, as a reference guide, the required computation time for different Caplets. Simulations have been run in a mesh of 100x50x50, with nt=12 steps per year, in a Pentium(R)IV processor (3.2 GHz, 1Gb RAM).

TMc	Tc	Ntotal	CpuTime
2	1	12	0.9 s
11	10	120	7.3 s
20	19	228	14.0 s

8. Conclusions

In this paper we have solved a stochastic volatility 3D HJM model. For that task we have reformulated the model in terms of a PDE system, and have solved it numerically. In particular, we have applied finite difference schemes: the temporal integrator used has been a θ scheme in its Crank-Nicholson version and we have used centered finite difference schemes for the spatial discretization. In order to decrease the computing time without losing precision, we have successfully applied ADI schemes. The validation of the numerical schemes has been realized comparing the numerical solution of some test curves (zero coupon bond, Caplet) with analytical models and a Heston model respectively. We can conclude that the results are quite satisfactory, as long as the typical errors obtained either for artificial or quite realistic forward interest rate curves haven't gone beyond 10^{-4} in any case.

References

- [1] Heath, Jarrow, Morton, *Bond pricing and the Term structure of interest rates: a new methodology for contingent claim valuation*, *Econometrica* **60**, 1 (1992).
- [2] M.W. Baxter, *General interest rate models and the Universality of HJM*, in *Mathematics of Derivative Securities*.
- [3] John C. Hull, *Options, Futures, and Other Derivatives*, 4th Edition, Prentice Hall, 1999.
- [4] A. Caverhill, *When is the short rate markovian?*, *Mathematical Finance* **4**, 4 (1994).
- [5] P. Ritchken, L. Sankarasubramanian, *Volatility structures of forward rates and the dynamics of the term structure*, *Mathematical Finance* **5**, 1 (1995).
- [6] K. Inui, M. Kiyima, *A Markovian framework in multi-factor Heath-Jarrow-Morton models*, *Journal of Financial and Quantitative Analysis* **33**, 3 (1998).
- [7] C. Chiarella, O.K. Kwon, *Forward rate dependent Markovian transformations of the Heath-Jarrow-Morton term structure model*, *Finance and Stochastics* **5**, 2 (2001).
- [8] C. Chiarella, O.K. Kwon, *A complete Markovian Stochastic Volatility model in the HJM framework*, *Asia-Pacific Financial Markets* **7**, 4 (2000).
- [9] I. Karatzas, S.E. Shreve, *Brownian motion and stochastic calculus*, Graduate texts in Mathematics, 2ed, Springer (1991).
- [10] D.W. Peaceman, H.H. Rachford, *The numerical solution of parabolic and elliptic differential equations*, *SIAM* **3**, (1955) pp. 28-41.
- [11] J. Douglas, H.H. Rachford, *On the numerical solution of heat conduction problems in two and three space variables*, *Trans. Amer. Math. Soc.* **82**, (1956) pp. 421-439.
- [12] J.L. Steger, *Implicit Finite Difference Simulation Flow About Arbitrary Geometries with Applications to Airfoils*, AIAA Paper 77-665, 1977.
- [13] D.J. Duffy, *Finite Difference methods in financial engineering: a partial differential equation approach*, Wiley (2006).
- [14] S. McKee, D.P. Wall, and S.K. Wilson, *An Alternating Direction Implicit Scheme for Parabolic Equations with Mixed Derivative and Convective Terms*, *Journal of Comput. Physics.* **126**, 64-76, 1996.
- [15] K.J. Jout and B.D. Welfert *Stability of ADI Schemes Applied to Convection-Diffusion Equations with Mixed Derivative Terms.*, *Applied Numerical Mathematics*, **57**, 1,19-35, 2007.

- [16] D. Tavella, C. Randall, *Pricing Financial Instruments: The Finite Difference Method*, Wiley, 2000.
- [17] Munson, *Fundamentals Of Fluid Mechanics*, Wiley 2007.
- [18] *The book Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, 2007.

9. Figures

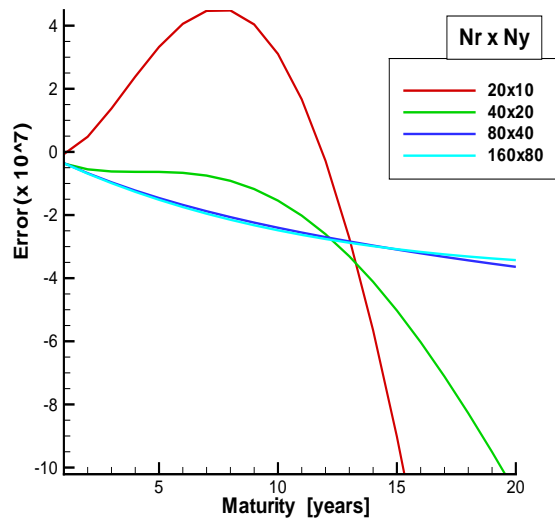


Figure 1: Plot of the numerical solution's error as a function of the mesh size, for the initial zero coupon curve. The particular values of the metric are $\alpha_r = 0.05$, $\alpha_y = 0.5$, $r_\infty = 25$, $y_\infty = 250$.

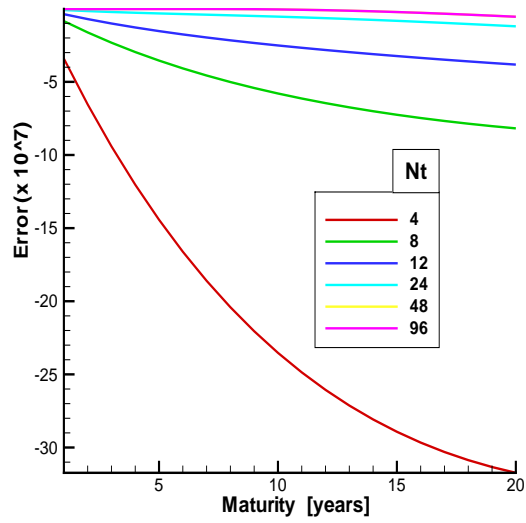


Figure 2: Plot of the numerical solution's error (convergence) as a function of the number of time steps per year, for the reference mesh of size 100x40. The metric parameters are the same as for figure 1.

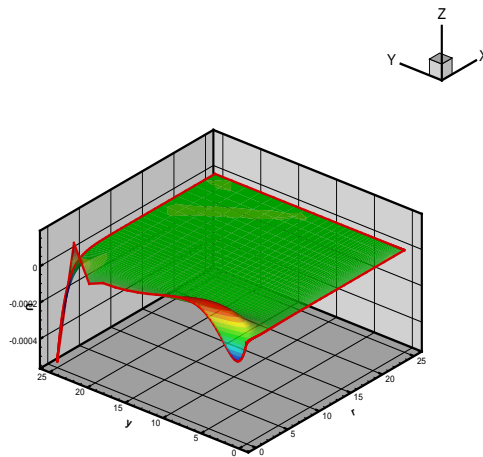


Figure 3: Error spatial distribution obtained in $t=T=20$ years for the zero coupon curve. The metric parameters are $\alpha_r = 0.05, \alpha_y = 0.5, r_\infty = 25, y_\infty = 25.$, the mesh is the reference one and the temporal discretization assumes 12 time steps per year.

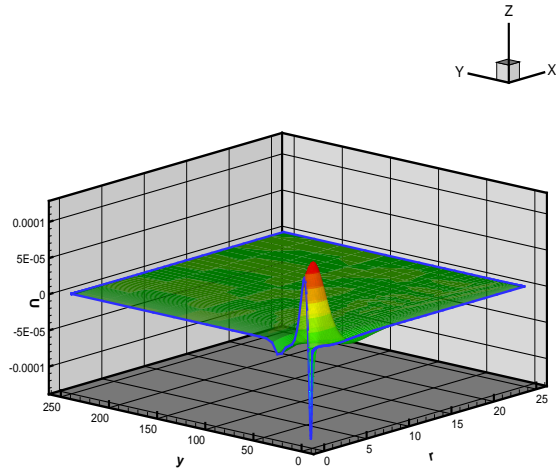


Figure 4: Error spatial distribution obtained in $t=T=20$ years for the zero coupon curve. The metric parameters are $\alpha_r = 0.05, \alpha_y = 0.5, r_\infty = 25, y_\infty = 250.$, the mesh is the reference one and the temporal discretization assumes 12 time steps per year.

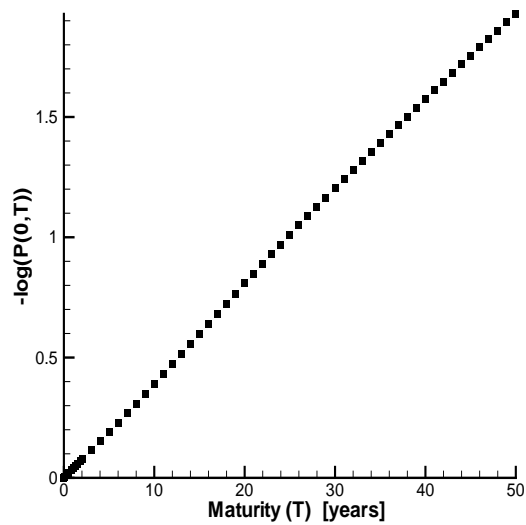


Figure 5: Initial zero coupon curve (discrete curve). Note that the plot is in semilog.

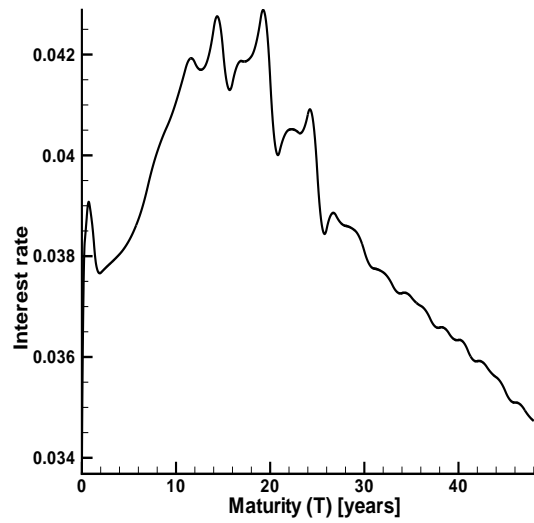


Figure 6: Approximation of the forward rate of interest curve, obtained with splines.

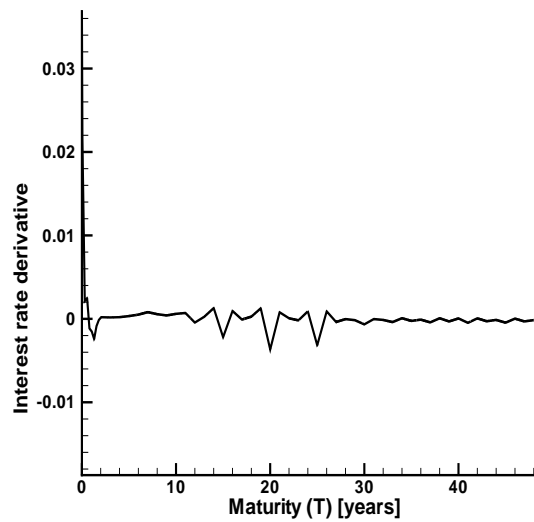


Figure 7: Approximation of the forward rate of interest curve derivative, obtained with splines.

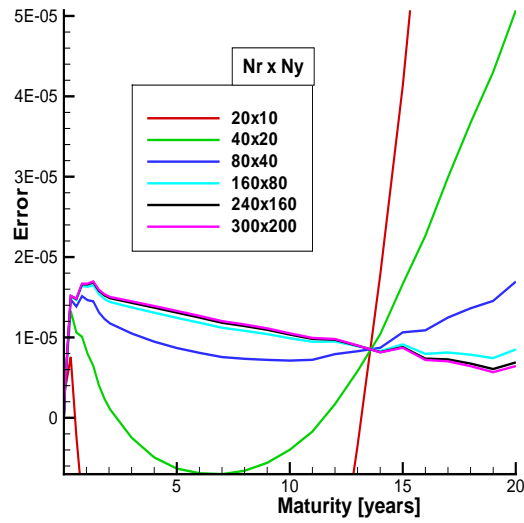


Figure 8: Plot of the numerical solution's error as a function of the mesh size $N_r \times N_y$, for the initial zero coupon curve. The particular values of the metric are $\alpha_r = 0.05, \alpha_y = 0.5, r_\infty = 25, y_\infty = 250$.

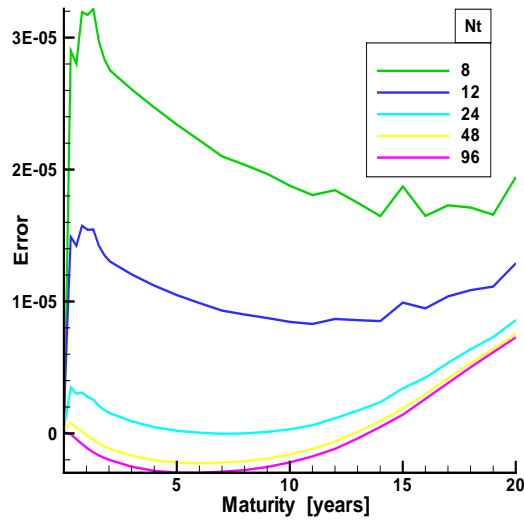


Figure 9: Plot of the numerical solution's error (convergence) as a function of the number of time steps per year, for the reference mesh of size 100×40 .

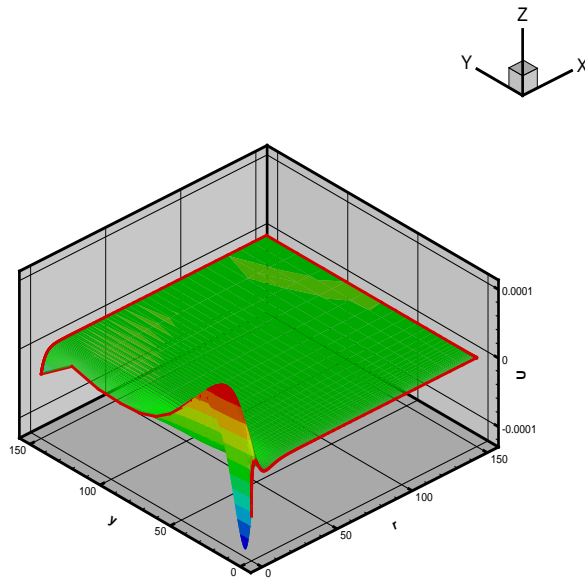


Figure 10: Error spatial distribution obtained in $t=T=20$ years for the zero coupon curve. The metric parameters are $\alpha_r = 0.05, \alpha_y = 0.5, r_\infty = 150, y_\infty = 150.$, the mesh is the reference one and the temporal discretization assumes 12 time steps per year.

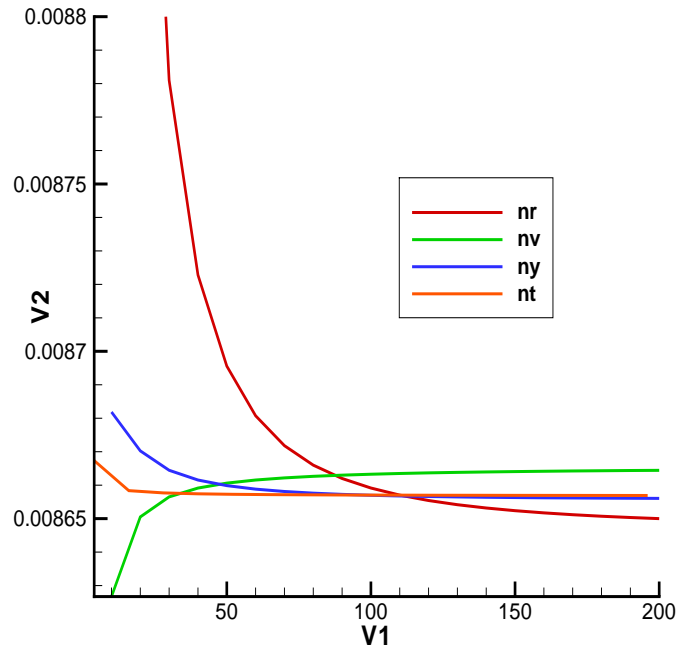


Figure 11: Premium variation for a generic Caplet as a function of nr, nv, ny, nt.

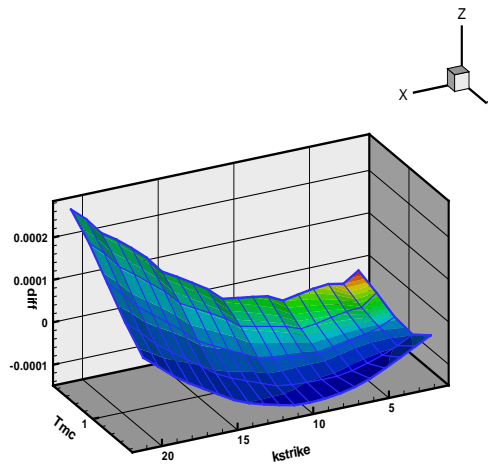


Figure 12: Difference between the Heston model and the HJM model in the assessment (premium) of several Caplets. Metric parameters: $\alpha_r = 0.5, \alpha_v = 0.5, \alpha_y = 0.5, r_\infty = 250, v_\infty = 30, y_\infty = 250$. and 12 time steps per year.

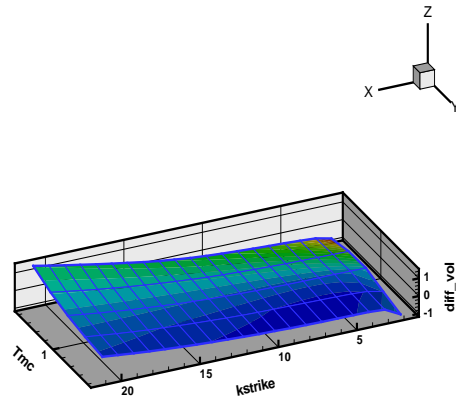


Figure 13: Difference between the Heston model and the HJM model in the assessment (volatility) of several Caplets. Metric parameters: $\alpha_r = 0.5, \alpha_v = 0.5, \alpha_y = 0.5, r_\infty = 250, v_\infty = 30, y_\infty = 250$. and 12 time steps per year.

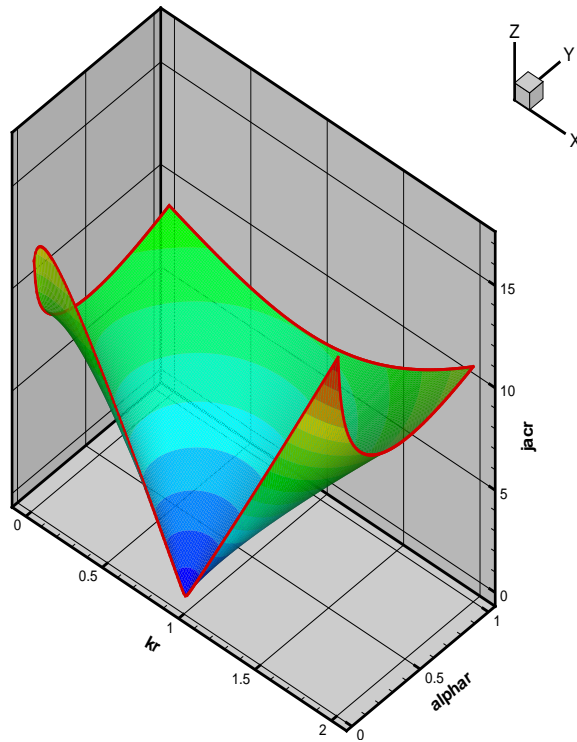


Figure 14: Variation of the transformation's Jacobian in $r = 1$ as a function of K_r and α_r .

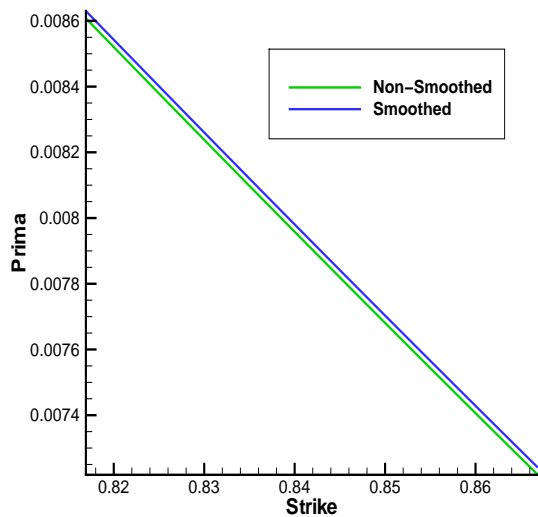


Figure 15: Premium variation for a generic Caplet and different strikes. Simulations have been realized in a single mesh of 100x40x40 with nt=12 steps per year.

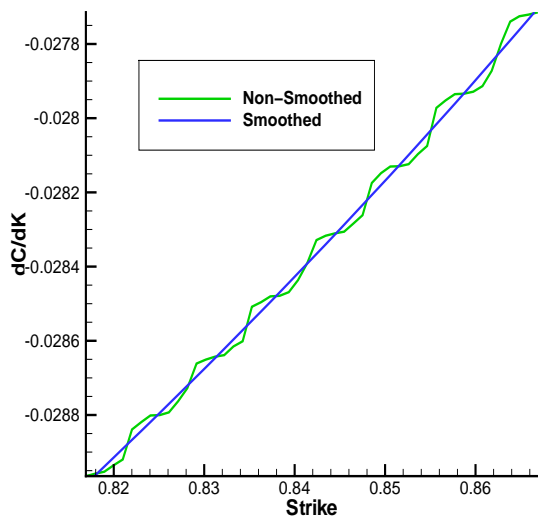


Figure 16: Variation of the premium derivative for a generic Caplet and different strikes. Simulations have been realized in a single mesh of 100x40x40 with nt=12 steps per year.