
Hatem Ben Ameur Eymen Errais
HEC Montreal Stanford University

August 31, 2004

Abstract

In this paper, we devise a numerical method to value bonds with embedded options. The term structure is driven by the short interest rate which we modeled as an extended Vasicek process and extended Cox Ingersoll Ross process. The numerical method proposed is based on neuro-dynamic programming. The approximation architecture chosen is a piecewise linear function involving a number of free parameters which we determine by neural network training. Models are calibrated to swaptions and caplet markets using analytical formulas which enhances the calibration tractability.

1 Introduction

In the last twenty years, interest rate-sensitive contingent claims have become increasingly popular. According to the Bank for International Settlements, the total notional amount outstanding in excess was about $60 trillion at the end of 1999. The International Swaps and Derivatives Association (ISDA) estimated that the notional amount of interest rate swaps and swaption at end of 1997 were respectiveley $22.3 trillion and $1.7 trillion.

In a series of recent papers, Brace Gatarek and Musiela (1997), Miltersen Sandmann and Sondermann (1997) and Jamshidian (1997) have independently proposed what is generally known as "Libor Market Models" (LMM), a model that is easily calibrated to caps and swaptions market quotes. These models have successfully introduced a theoretical framework to the already established market caps and swaptions Black based formulas. While for these makets, LMM became a standard, there is no such agreement in bonds embedded options market for example. One of the oldest approaches is based on modelling the evolution of short rates. Different short rate models were proposed starting with the important contribution of Vasicek (1977) who proposed an Ornstein-Uhlenbeck process. Cox Ingersoll and Ross (1985) (CIR) proposed instead a noncentral chi-square distribution to preclude negative interest rates. These "endogeneous" models have the major drawback of not fitting the initial term structures (interest rate and volatility). Hull and White (1990) tried to overcome this problem by substituting the constant parameters of Vasicek (1977) and CIR
(1985) with time dependent functions defined on the basis of the initial curve. While the extended Vasicek was adopted in market practices due to its analytical tractability\(^1\), the extended CIR (ECIR) model as proposed by Hull and White (1990) was not adopted due to its implementation problems. Rogers (1995) showed that to price bonds in the ECIR model one has to solve a Riccati equation which makes the model less tractable. Jamshidian (1995), Maghsoodi (1996) and Rogers (1995) proposed more tractable extensions of CIR. Still, in these models, pricing bonds amounts in solving particular “families” of Riccati equations. Hull and White (1994) and Errais (2003) developed some numerical approximation approaches to overcome this intractability problem. More recently, Brigo and Mercurio (2001) proposed a further extension of CIR (CIR++) which, to our knowledge, the most tractable extension and the most likely to be adopted by practitioners\(^2\). This model presents the big advantage of having analytical formulas for bond prices, bond option prices, swaption and cap prices. In the remainder of the paper, we will adopt this model in our pricing along with the extend Vasicek model.

Our focus in this paper is on bonds with embedded options. Several bonds contain one or several options coming in various flavours. First, the issuer can have the right to buy back the bond for a fixed price, called the “call price”, during the bond’s life. On the other hand, the investor may have the right to sell the bond for a fixed price, put price. These options (call and put) are American options, i.e. could be exercised any time during the bond’s life. Unfortunately, there are no analytical formulas for valuing American options, not even in the basic Black and Scholes model(1973). Numerical methods, essentially trees, finite-differences and more recently simulation, are generally used to price these claims using a backward induction framework. Particulary, the pricing of bonds with embedded options can be traced back to Brennan and Schwartz (1977). Hull and White (1990) have proposed a tree based method. Trees crudely approximate the dynamics of the underlying asset and convergence is not easy to attain. Büttler and Waldvogel (1996) priced callable bonds using Green’s function. This approach suffers from the drawback that it is limited to constant parameter models. d’Halluin and al. (2001) generalised their work using flux limiters. More recently, Ben-Ameur and al. (2004) proposed a dynamic programming (DP) approach in pricing options embedded in bonds. They proposed an efficient DP approximation to price callable and puttable bonds in the case where short rates are driven by Vasicek(1977) and CIR models (1985).

In this paper, we follow the approach proposed by Ben-Ameur and al.(2004) and extend it to the CIR++ and extended Vasicek model. The numerical method proposed is based on neuro-dynamic programming. The value function (bond price) is computed by approximation. The first step in the development of an

\(^1\)In market practices, corporate bonds are priced using a gaussian model (Hull and White (1990)) and a lognormal model (Black and Krazinski (1990)). This practice allows traders to have a range of prices which will encompass the “true” bond price.

\(^2\)It has been, actually, adopted and implemented by some interest rates and credit risk practitioners.
approximate representation is to choose an approximation architecture, that is, a certain functional form, a piecewise linear function in our case, involving a number of free parameters. These parameters are to be tuned, what we call neural network training, so it can provide a best fit of the function to be approximated. The neural network training, i.e. the calibration, is performed through the swaptions and caplets market quotes since they are by far the most liquid and important interest rate derivatives.

The remainder of this paper is organized as follows. Section 2 sets up the characteristics of the extended Vasicek and CIR++. Section 3 describes in detail our approach to approximating the value function. Section 4 addresses the calibration issue. Section 5 discusses some numerical experiments.

2 Short rate models and their statistical properties

2.1 Extended Vasicek Model

The extended Vasicek model for interest rates was first proposed by Hull and White (1990b). On a filtered probability space \((\Omega, \mathcal{F}, \mathbb{F}, Q)\) the stochastic differential equation of this process could be written as

\[
dr_t = \kappa (\theta_t - r_t) dt + \sigma_t dW_t
\]

Integrating, we get

\[
r_t' = r_t e^{-\kappa (t' - t)} + \kappa \int_t^{t'} e^{-\kappa (t' - u)} \theta_u du + \int_t^{t'} e^{-\kappa (t' - u)} \sigma_u dW_u
\]

for \(t' > t \geq 0\).

Hence, the mean of this process is

\[
\mu_1(r_t, t') = r_t e^{-\kappa (t' - t)} + \kappa \int_t^{t'} e^{-\kappa (t' - u)} \theta_u du
\]

and covariance function.

\[
\rho_r(s, t') = e^{-\kappa (t' + s)} \int_s^{t'} e^{2\kappa u} \sigma^2(u) du
\]

and therefore

\[
\sigma^2_1(t') = \rho_r(t', t') = e^{-2\kappa t'} \int_t^{t'} e^{2\kappa u} \sigma^2(u) du
\]

We will now review the statistical properties of \(r_u\) and \(\int_{t_1}^{t_2} r_u du\)

**Proposition 1** The mean of \(\int_{t_1}^{t_2} r_u du\) is:

\[
\mu_2(r_{t_1}, t_1, t_2) = \int_{t_1}^{t_2} e^{-\kappa (u - t_1)} r_{t_1} du + \kappa \int_{t_1}^{t_2} \int_{t_1}^{t_2} e^{-\kappa (u - v)} \theta_u dv du dt
\]
And the variance is

$$\sigma^2(t_1, t_2) = \int_{t_1}^{t_2} e^{\kappa(v-t_1)} \sigma_v^2 \left( \int_v^{t_2} e^{-\kappa(u-v)} du \right)^2 dv$$

**Proposition 2** \((r_{t_2}; \int_{t_1}^{t_2} r_u du)\) is jointly normal with mean \(\mu = (\mu_1(r_{t_1}, t_2); \mu_2(r_{t_1}, t_1, t_2))\) and covariance

$$\text{cov} \left( r_{t_2}; \int_{t_1}^{t_2} r_u du \right) = \int_{t_1}^{t_2} \rho_r(u, t_2) du$$

where \(\rho_r(u, t_2)\) as defined by (3). Also define the correlation coefficient :

$$\rho(t_1, t_2) = \frac{\int_{t_1}^{t_2} \rho_r(u, t_2) du}{\sigma_2(t_1, t_2) \sigma_1(t_2)}$$

**Proof.**

$$\text{cov} \left( r_{t_2}; \int_{t_1}^{t_2} r_u du \right) = E \left( (r_{t_2} - \mu_1(r_{t_1}, t_2)) \int_{t_1}^{t_2} (r_u - \mu_2(r_{t_1}, t_2)) du \right)$$

$$= E \left( \int_{t_1}^{t_2} (r_u - \mu_2(r_{t_1}, t_2)) (r_{t_2} - \mu_1(r_{t_1}, t_2)) du \right)$$

$$= \int_{t_1}^{t_2} \rho_r(u, t_2) du$$

\[ \blacksquare \]

### 2.2 The extended Cox Ingersoll Ross model: CIR++

The CIR++ model was proposed Brigo and Mercurio (2001). However, similar idea has been developed independently by Dybvig (1997) and by Avellaneda and Newman (1998).

On a filtered probability space \((\Omega^x, \mathcal{F}^x, \mathbb{F}^x, Q^x)\), we consider a given time-homogeneous stochastic process, whose dynamics is expressed by :

$$dx_t = \kappa (\theta - x_t) dt + \sigma \sqrt{x_t} dW_t \quad \text{(4)}$$

where \(\kappa, \theta\) and \(\sigma\) are positive constants satisfying \(2\kappa\theta > \sigma^2\). This condition ensures that the origin is inaccessible.

The process \((x_t)_t\) as defined by (4) is a non-central \(X^2\) process. Recall that \(X\) follows a noncentral \(X^2\) distribution means that

$$X = \sum_{j=1}^{d} (Z_j + \delta_j)^2$$

where \(Z_j\)'s are independent unit normal variables, and \(\delta_j\)'s are constants. \(X\) is called the noncentral \(X^2\) distribution with \(d\) degrees of freedom and noncentrality parameter \(\lambda = \sum_{j=1}^{d} \delta_j^2\). It is denoted by the symbol \(X^2_d(\lambda)\) (as opposed to \(X^2_d\), the central \(X^2\) with \(d\) degrees of freedom).
Proposition 3 For \( t^* > t' \geq 0 \), the Laplace transform of the couple \( \left( x_{t^*}, \int_{t'}^{t^*} x^s ds \right) \) under measure \( Q^x \) is given by:

\[
\mathcal{L}_{Q^x} \left( t^*, t', \beta, \alpha \right) = \mathbb{E}_{Q^x} \left[ e^{-\beta \int_{t'}^{t^*} x^s ds - \alpha x_{t^*}} \right] = A(t', t^*, \beta, \alpha)e^{-B(t', t^*, \beta, \alpha) x_{t^*}}
\]

where

\[
A(t', t^*, \beta, \alpha) = \left( \frac{2h(h(\beta) - \kappa)}{(\sigma^2 \alpha + h(\beta) + \kappa) (e^{h(\beta)(t'-t)} - 1)} + 2h(\beta) \right)^{\frac{2\alpha}{\sigma^2}}
\]

and

\[
B(t', t^*, \beta, \alpha) = \frac{\alpha \left( h(\beta) + \kappa + e^{h(\beta) + \kappa}(h(\beta) - \kappa) + 2h(\beta) - 1 \right)}{(\sigma^2 \alpha + h(\beta) + \kappa) (e^{h(\beta)(t'-t)} - 1) + 2h(\beta) \beta}
\]

where \( h(\beta) = \sqrt{\kappa^2 + 2\sigma^2 \beta} \). For a detailed proof of this result we refer the reader to Feller (1951).

Theorem 4 Under the \( Q^x \) probability measure, the distribution of \( \frac{Z_{t+}}{\eta} \), conditioning to \( x_{t_1} = a \), is a non-central \( \chi^2 \) distribution with noncentrality parameter \( \lambda \) and \( d \) degrees of freedom, where \( \lambda = \frac{4\alpha a}{\sigma^2 (e^{\alpha(t_2-t_1)} - 1)} \) with \( d = \frac{4\alpha}{\sigma^2} \) and \( \eta = \frac{\sigma^2 (1 - e^{\alpha(t_2-t_1)})}{4\kappa} \).

After defining the time-homogeneous spot-rate model \( x_t \), we can define the instantaneous short rate under the new filtered probability space \( (\Omega, \mathcal{F}, \mathbb{F}, Q) \) by:

\[
r_t = x_t + \varphi_t
\]

\[
dx_t = \kappa (\theta - x_t) dt + \sigma \sqrt{T} dW_t
\]

where \( x_0, \kappa, \theta \) and \( \sigma \) are positive constants satisfying \( 2\kappa \theta > \sigma^2 \). Notice that \( x_0 \) is one more parameter at our disposal. We are free to select its value as long as \( \varphi_0 = r_0 - x_0 \).

Proposition 5 The price at time \( t \) of a zero-coupon bond maturing at time \( T \) is:

\[
P(t, T, r_t) = \bar{A}(t, T) \exp(-B(t, T) r_t)
\]

where \( \bar{A}(t, T) = \frac{P^M(0, T) A(0, t) e^{-B(0, t)x_0}}{P^M(0, T) A(0, T) e^{-B(0, T)x_0}} A(t, T) \exp(B(t, T) \varphi_t) \) with \( A(t, T) = \left( \frac{2h(h(\beta) + \kappa)}{(h(\beta) + \kappa) (e^{h(\beta)(T-t)} - 1) + 2h(\beta)} \right)^{\frac{2\alpha}{\sigma^2}} \),

\[
B(t, T) = \frac{2h(e^{h(\beta)(T-t)} - 1)}{(h(\beta) + \kappa) (e^{h(\beta)(T-t)} - 1) + 2h(\beta} \right), \ h = \sqrt{\kappa^2 + \sigma^2} \) and \( P^M(0, T) \) the time 0 market price of a bond maturing at \( T \).
Proposition 6 The price at time $t$ of a European call option with maturity $T > t$ and strike price $K$ on a zero-coupon bond maturing at $\tau > T$ is:

$$ZBC(t, T, \tau, K) = P(t, \tau) \mathcal{X}^2 \left( \frac{2\hat{r} [\rho + \psi + B(T, \tau)]}{\hat{r}^2 \rho \exp(h(T-t))}, \frac{4\sqrt{\rho}}{\sigma}, \frac{2\rho^2 \tau \exp(h(T-t))}{\rho + \psi} \right) - K P(t, T) \mathcal{X}^2 \left( \frac{2\hat{r} [\rho + \psi + B(T, \tau)]}{\hat{r}^2 \rho \exp(h(T-t))}, \frac{4\sqrt{\rho}}{\sigma}, \frac{2\rho^2 \tau \exp(h(T-t))}{\rho + \psi} \right),$$

where $
\hat{r} = \frac{1}{n(T, \tau)} \left[ \ln \frac{A(T, \tau)}{K} - \ln \frac{P_M(0, T)}{A(0, \tau)} \exp(-B(0, \tau) x_0) \right]$, $\mathcal{X}^2(x, d, \lambda)$ is the cumulative noncentral $\mathcal{X}^2$ distribution evaluated at $x$ having $d$ degrees of freedom and noncentral parameter $\lambda$, $\rho = \frac{2h^2}{\sigma^2}$, and $\psi = \kappa + \frac{h^2}{\sigma^2}$.

The price of a European put option is given by put-call parity:

$$ZBP(t, T, \tau, K) = ZBC(t, T, \tau, K) - P(t, \tau) + K P(t, T)$$

We refer the reader to Brigo and Mercurio (2001) for a formal proof of these propositions.

3 The Dynamic Programming Problem

3.1 Notation

We introduce in this section some notation which will allow us to better understand the approximation to solve the DP problem as proposed by Ben Ameur and al. (2004). Note that we will have the same setting as in Büttler and Waldvogel (1996) and. d’Halluin and al. (2001) so that we can perform a comparison. In their setting, the exercise of the embedded options is limited to the coupon dates posterior to a known protection period and that there is a notice period of fixed duration $\Delta t$. The benefits of an exercise decision are obtained at the coupon date.

Let $t_{c_0}, t_{c_1}, ..., t_{c_n}$ be a sequence of date, where $t_{c_0}$ is the initial time, $t_{c_1}, ..., t_{c_{n-1}}$ are the coupon dates, and $t_{c_n} = T$ is the bond maturity. For simplicity, we assume the principal to be equal to 1 and we denote the coupon value by $c$. Let $t_{n+1}, ..., t_n$ be a sequence of date where we can exercise the embedded options. Since there is a notice period $\Delta t$, $t_m = t_{c_m} - \Delta t > t_{m-1}$ for $m = n^*, ..., n$. The time increment $t_{n^*} - t_0$ is the protection period against early exercising. We adopt the convention that $t_{n+1} = T$ and $t_m = t_{c_0}$ for $m = 0, ..., n^* - 1$.

Let $C_m$ and $P_m$ be the call and the put prices at time $t_m$. Assume that $C_m \geq P_m \geq 0$. Let $v_m(r)$ be the value of the bond and $v_m^h(r)$ its holding value at time $t_m$, where $r$ is the interest rate at that time. Also, define $D(r, t, \delta) = \mathbb{E}^Q \left[ \exp \left( - \int_t^{t+\delta} r_s \, ds \right) | r_t = r \right]$ where $Q$ is the risk neutral measure. For ease of notation, we denote $D(r, t_m, \Delta t)$ by $D_m(r)$.
The optimal exercise strategy is given by:

\[ v_m (r) = \begin{cases} 
D_m (r) (C_m + c) & \text{if } v_m^h (r) > D_m (r) C_m \\
D_m (r) (P_m + c) & \text{if } v_m^h (r) \leq D_m (r) P_m
\end{cases} \]

(7)

Where \( v_m^h \) is the holding value of the bond at time \( t_m \). The holding value is computed by risk neutral pricing as follows:

\[ v_m^h (r) = \mathbb{E} \left[ e^{-\int_{t_m}^{t_{m+1}} r(t) dt} v_{m+1} (r_{t_{m+1}}) \mid \mathcal{F}_{t_m}, r_{t_m} = r \right] \]

(8)

for \( m = n^*, \ldots, n \) and

\[ v_0^h (r) = \mathbb{E}_{0,r} \left[ e^{-\int_{t_0}^{t_{n^*}} r(t) dt} v_{n^*} (r_{t_{n^*}}) \right] \]

(9)

Therefore, the bond price is given by

\[ v_{n+1} (r) = 1 + c \]

(10)

\[ v_m (r) = \max \{ D_m (r) P_m, \min (D_m (r) C_m v_m^h (r)) \} + cD_m (r), \text{ for } m = n^*, \ldots, n \]

(11)

\[ v_0 (r) = v_0^h (r) + c \sum_{m=1}^{n^*} D_m (r) \]

(12)

The equations (8) to (12) define the dynamic programming problem we have to solve to value the bond. This can not be done analytically and in the next section we device a numerical method to solve this DP problem.

### 3.2 Approximation of the Value Function

To solve the dynamic programming problem presented above, we will resort to the finite element method based on a piecewise linear interpolation. Let us give the basic steps of this approach. First, we start by discretizing the state space. Let \( a_0 = -\infty < a_1 < \ldots < a_{p+1} = +\infty \) be a set of points, so that we come up with a sequence of intervals \( R_i = [a_i, a_{i+1}] \) for \( i = 1, \ldots, p \), and \( R_1 = (-\infty, a_1] \).

Given a continuous approximation \( \hat{v} \) of the value function at time \( t_m \), fully determined at each point of our grid, we do a piecewise linear interpolation as follows:

\[ \hat{v}_m (a) = \sum_{i=1}^{p+1} (a_i^m + \beta_i^m a) I_i (a) \]
where \( I_i \) is an indicator function i.e.
\[
I_i(a) = \begin{cases} 
1 & \text{if } a \in R_i \\
0 & \text{elsewhere}
\end{cases}
\]

The coefficients \( \alpha_i^m \) and \( \beta_i^m \) are obtained by solving the following equation at each time step, note that this equation is the result of the continuity assumption imposed on \( \hat{v}_m \).
\[
\hat{v}_m(a_i) = \tilde{v}_m(a_i)
\]

and therefore
\[
\begin{align*}
\alpha_i^m &= \frac{a_{i+1} \tilde{v}_m(a_i) - a_i \bar{v}_m(a_{i+1})}{a_{i+1} - a_i} \\
\beta_i^m &= \frac{\tilde{v}_m(a_{i+1}) - \bar{v}_m(a_i)}{a_{i+1} - a_i}
\end{align*}
\]

Assume now that \( \hat{v}_{m+1} \) is known, therefore the holding value at time \( t_m \) is computed as follows
\[
\hat{v}_m^h(a_k) = \mathbb{E}_{m,a_k} e^{-\int_{t_m}^{t_{m+1}} r(t) dt} \hat{v}_{m+1}(r_{t_{m+1}})
\]

\[
= \sum_{i=1}^{p+1} (\alpha_i^{m+1} A_{k,i}(t_m, t_{m+1}) + \beta_i^{m+1} B_{k,i}(t_m, t_{m+1}))
\]

\[
A_{k,i}(t_m, t_{m+1}) = \mathbb{E}_{m,a_k} e^{-\int_{t_m}^{t_{m+1}} r(t) dt} I_i(r_{t_{m+1}})
\]

\[
B_{k,i}(t_m, t_{m+1}) = \mathbb{E}_{m,a_k} e^{-\int_{t_m}^{t_{m+1}} r(t) dt} r_{t_{m+1}} I_i(r_{t_{m+1}})
\]

for \( m = n^*, ..., n \) and
\[
A_{k,i}(t_0, t_{n^*}) = \mathbb{E}_{m,a_k} e^{-\int_{t_0}^{t_{n^*}} r(t) dt} I_i(r_{t_{n^*}})
\]

\[
B_{k,i}(t_0, t_{n^*}) = \mathbb{E}_{m,a_k} e^{-\int_{t_0}^{t_{n^*}} r(t) dt} r_{t_{n^*}} I_i(r_{t_{n^*}})
\]

We will now compute the values of \( A_{k,i} \) and \( B_{k,i} \).

**Theorem 7** The coefficients \( A \) and \( B \) for the extended Vasicek model are given by
\[
A_{k,i}(t_m, t_{m+1}) = E_{m,a_k} \left[ e^{-\int_{t_m}^{t_{m+1}} r(t) dt} I_i(r_{t_{m+1}}) \right]
\]

\[
= e^{-\mu_2(a_k,t_m,t_{m+1}) + \frac{1}{2} \sigma_2^2(t_m,t_{m+1})} \left[ \Phi \left( \hat{a}_{k,i} (t_m, t_{m+1}) \right) - \Phi \left( \hat{a}_{k,i-1} (t_m, t_{m+1}) \right) \right]
\]

\[
B_{k,i}(t_m, t_{m+1}) = E_{m,a_k} \left[ e^{-\int_{t_m}^{t_{m+1}} r(t) dt} r_{t_{m+1}} I_i(r_{t_{m+1}}) \right]
\]

\[
= e^{-\mu_2(a_k,t_m,t_{m+1}) + \frac{1}{2} \sigma_2^2(t_m,t_{m+1})} \left[ (\mu_1(a_k,t_m) - \sigma_1(t_m,t_{m+1})) \Phi \left( \hat{a}_{k,i} \right) - \Phi \left( \hat{a}_{k,i-1} \right) \right]
\]

\[
- \frac{\sigma_1(t_m)}{\sqrt{2\pi}} \left[ e^{-\hat{a}_{k,i}^2(t_m,t_{m+1})} - e^{-\hat{a}_{k,i-1}^2(t_m,t_{m+1})} \right]
\]

where \( \Phi \) is the cumulative normal distribution, \( \hat{a}_{k,i} (t_m, t_{m+1}) = \frac{a_i - \mu_2(a_k,t_m) + \sigma_2(t_m,t_{m+1})}{\sigma_1(t_m)} \). and \( \rho = \rho(t_m, t_{m+1}) \)
Theorem 8 The coefficients A and B for the CIR++ model are given by

\[
A_{k,i}(t_m, t_{m+1}) = P(t_m, t_{m+1}, a_k) \times \left( \sum_{n=0}^{\infty} e^{-\frac{\left(\lambda_k/2\right)^2}{n!}} \frac{(\lambda_k/2)^n}{n!} \times \left( F_{d+2n} \left( \frac{a_{i+1} - \varphi_{t_{m+1}}}{\eta} \right) - F_{d+2n} \left( \frac{a_i - \varphi_{t_{m+1}}}{\eta} \right) \right) \right)
\]

where

\[
\eta = \frac{\sigma^2 \left( e^{h(t_{m+1}-t_m)} - 1 \right)}{2(h + \kappa) \left( e^{h(t_{m+1}-t_m)} - 1 \right) + 2h}
\]

\[
\lambda_k = \frac{8h^2 e^{h(t_{m+1}-t_m)} a_k}{\sigma^2 \left( (h + \kappa) \left( e^{h(t_{m+1}-t_m)} - 1 \right) + 2h \right)}
\]

and \( P(t_m, t_{m+1}, a_k) \) is the price at time \( t_m \) of a bond maturing \( t_{m+1} \) with initial rate \( r_{t_m} = a_k \) given by equation (5)

\[
B_{k,i}(t_m, t_{m+1}) = P(t_m, t_{m+1}, a_k) \times \eta \times \left( \sum_{n=0}^{\infty} e^{-\frac{\left(\lambda_k/2\right)^2}{n!}} \frac{(\lambda_k/2)^n}{n!} \times \left( F_{d+2n} \left( \frac{a_{i+1} - \varphi_{t_{m+1}}}{\eta} \right) - F_{d+2n} \left( \frac{a_i - \varphi_{t_{m+1}}}{\eta} \right) \right) \right)
\]

where

\[
Q_n (a_{i+1}, a_i) = -2 \left( a_{i+1} f_{d+2n} \left( \frac{a_{i+1}}{\eta} \right) - a_i f_{d+2n} \left( \frac{a_i}{\eta} \right) \right) + (d + 2n) \times \left( F_{d+2n} \left( \frac{a_{i+1}}{\eta} \right) - F_{d+2n} \left( \frac{a_i}{\eta} \right) \right)
\]

with \( h = \sqrt{\kappa^2 + 2\sigma^2} \) and \( d = \frac{4\kappa}{\sigma} \). \( F_{d+2n} \) and \( f_{d+2n} \) are respectively the cumulative and density functions of a khi-2 with \( d + 2n \) degrees of freedom model

3.3 Pseudo-code of the dynamic programming problem

Now that we computed the coefficients A and B for each time and state step, let us give the big picture of how the pseudo-code of the dynamic programming problem works:

Equation (10) gives the terminal value of the value function.

For \( m = n : -1 : n^* \)

For \( k = 1 : p, k + + \)

Compute \( \tilde{v}_{m}^{h}(a_k) \) by (14)

Compute \( \tilde{v}_{m}(a_k) \) by (11)

Next k

Compute the coefficients of \( \tilde{v}_{m} \) at step \( m \) by (13)
Next m
Set \( m = 0 \)
For \( k = 1 : p, k + + \)
    Compute \( \tilde{v}_0^k (a_k) \) by (9)
    Compute \( \tilde{v}_0 (a_k) \) by (12)
Next \( k \)

3.4 Calibration in the extended Vasicek framework

In the last decade, swaptions became one of the most important and liquid instruments in the fixed income security market. Hence, it is important that we calibrate the model to this instrument. Notice that in the extended Vasicek model, the calibration of \( \sigma \)'s and \( \kappa \) is performed independently from \( \theta \)'s. Therefore, we calibrate the \( \sigma \)'s and \( \kappa \) to swaptions, then we calibrate the \( \phi \)'s to the zero-coupon curve. To do so, first we need to compute caplet prices since, as it will be clearer later, we will show that pricing a swaption in a one factor model amounts to pricing a portfolio of caplets. To perform these calculations, we will use the approach adopted by Mercurio and Brigo (2001) and Cousot and Pain (2003).

3.4.1 Caplet prices

The value of the discounted payoff, at time \( t \), of a caplet starting at \( T_{i-1} \) and maturing at \( T_i \) with a strike \( K \) is:

\[
\text{PayOff}_{\text{cap}}(t, T_{i-1}, T_i) = e^{-\int_{T_{i-1}}^{T_i} r_s ds} \left( (L(T_{i-1}, T_i) - K)^+ \right)
\]

Therefore the caplet can be seen as a put on a bond, and its price at time \( t \) under the \( T_{i-1} \) forward measure is

\[
P_{\text{caplet}}(t, T_{i-1}, T_i) = (1 + \tau_i K)E^Q \left[ e^{-\int_{T_{i-1}}^{T_i} r_s ds} P(T_{i-1}, T_i) \left( \frac{1}{1 + \tau_i K} - P(T_{i-1}, T_i) \right)^+ \mid \mathcal{F}_{T_{i-1}} \right] | \mathcal{F}_t
\]

where \( P(t, T_{i-1}) \) is the price at \( t \) of a zero-coupon bond maturing at time \( T_{i-1} \), \( Q^{T_{i-1}} \) is the \( T_{i-1} \)-forward measure.
Applying Black(1976) formula to \( P(T_{i-1}, T_i) \) we obtain

\[
P_{\text{caplet}}(t, T_{i-1}, T_i) = P(t, T_{i-1}) \left( 1 + \tau_i K \right) B(t, t - T_{i-1}, T_{i-1}) \frac{1}{(1 + \tau_i K)} P(T_{i-1}, T_i) \Sigma_{T_{i-1}}^2 \]

(16)

where \( \Sigma_{T_{i-1}}^2 = \left( e^{-k(T_{i-1}-t)} - e^{-kT_{i-1}} \right)^2 \int_t^{T_{i-1}} \sigma^2 u du \) (since the diffusion of \( P(T_{i-1}, T_i) \) is equal to \( \sigma e^{-k(T_{i-1}-t)} \)).

3.4.2 Calibration to the swaption volatility structure

The main difference between swaptions and caps, as far as pricing is concerned, is that the payoff of the former is not additively "separable" with respect to the different rates. However, as we will see later, in the one factor framework, this additivity featured could be retrieved.

Define \( T_\alpha \) the maturity of the swaption and \( T_\beta \) the maturity of the underlying swap. Therefore the discounted payoff of a payer swaption could be written as

\[
\text{Payoff}_{\text{swaption}} = e^{-\int_{T_\alpha}^{T_{\alpha}} \sigma^2 \sigma u du} \left( \sum_{i=\alpha+1}^\beta \tau_i e^{-\int_{T_\alpha}^{T_{\alpha}} \sigma^2 \sigma u du} \left( F_{T_\alpha}(T_i - 1, T_i) - K \right) \right) +
\]

where \( F_{T_\alpha}(T_{i-1}, T_i) \) is the forward rate prevailing at time \( T_\alpha \) applied to the period \( (T_{i-1}, T_i) \).

Now, note that the floating cash flow stream is exactly the same as that generated by a floating rate bond except that no final principal payment is made. We know that the initial value of a floating rate bond is par; hence the value of the floating rate portion of the swap is par minus the present value of the principal received at \( T_\beta \).

\[
\text{Payoff}_{\text{swaption}} = \left( 1 - P(T_\alpha, T_\beta) - \sum_{i=\alpha+1}^\beta \tau_i K P(T_{T_\alpha}, T_i) \right) +
\]

\[
= \left( 1 - \sum_{i=\alpha+1}^\beta c_i P(T_{T_\alpha}, T_i) \right) +
\]

where \( c_i = \tau_i K \) for \( i = \alpha, ..., \beta - 1 \) and \( c_\beta = 1 + \tau_\beta K \).

Then swaptions could be priced as a put on coupon bearing bond. For that purpose we use the decomposition of Jamshidian (1989) to find the exact value.

Following Jamshidian and since our model is one factor model we can find \( r^* \) such that,

\[
1 = \sum_{i=\alpha+1}^\beta c_i P(T_{\alpha}, T_i, r^*)
\]

(17)

where \( P(T_{\alpha}, T_i, r^*) \) is determined by (19).

Therefore we can write the payoff in the following form:

\[
\text{Payoff}_{\text{swaption}} = e^{-\int_{T_\alpha}^{T_{\alpha}} \sigma^2 \sigma u du} \left( \sum_{i=\alpha+1}^\beta c_i P(T_{T_\alpha}, T_i, r^*) - c_i P(T_{T_\alpha}, T_i, r_\alpha) \right) +
\]

11
Now, since $P(T_\alpha, T_i, r)$ is a decreasing function of $r$, if $r^*$ is higher than $r_{T_\alpha}$ the cash flows are negative and the payoff is zero, if not they are all positive. Therefore, we can put the positive sign inside the sum. Note that this result, which seems to be surprising at a first glance, is due to the fact that we are working in one factor framework. This analysis leads us to the following expression:

$$\text{Payoff}_{\text{swaption}} = e^{-\int_0^{T_\alpha} r_t ds} \sum_{i=\alpha+1}^{\beta} \left( c_i P(T_\alpha, T_i, r^*) - c_i P(T_\alpha, T_i, r_{T_\alpha}) \right)^+$$

Therefore the swaption can be priced as a sum of caplets:

$$P_{\text{swaption}}(t, T_\alpha, T_\beta) = P(t, T_\alpha) \sum_{i=\alpha+1}^{\beta} c_i \text{Bl}(T_{i-1} - t, K^i, F(t, T_\alpha, T_i), \Sigma_{T_{i-1}}^2)$$

where $F(t, T_\alpha, T_i) = \frac{P(t, T_i)}{P(t, T_\alpha)}$, $\Sigma_{T_{i-1}}^2 = \frac{(e^{-r_{T_i}} - e^{-r_{T_{i-1}}})^2}{k^2 (T_{i-1} - t)} \int_t^{T_\alpha} \sigma^2(u) du$ and $K^i = P(T_\alpha, T_i, r^*) = e^{-r^* (T_i - T_\alpha)}$

Now, recall that the payer swaption is priced in the market by Black’s formula as following:

$$P_{\text{swaptionMkt}}(t, T_\alpha, T_\beta) = \text{Bl}(K, S_{\alpha,\beta}(t), \sigma_{\alpha,\beta} \sqrt{T_\alpha - t}) \sum_{i=\alpha+1}^{\beta} P(t, T_\alpha) \tau_i$$

where $\sigma_{\alpha,\beta}$ is the implied volatility quoted in the market, and $S_{\alpha,\beta}(t)$ is the swap rate between maturities $T_\alpha$ and $T_\beta$ and computed as $S_{\alpha,\beta}(t) = \frac{P(t, T_\alpha) - P(t, T_\beta)}{\sum_{i=\alpha+1}^{\beta} \tau_i P(0, T_i)}$.

So now, to calibrate the volatility term structure, all that we need is to compute the value market and theoretical price of swaptions using Black (1976) and minimize the difference.

### 3.4.3 Calibration to the term structure curve.

In this section, we follow the approach proposed by Mercurio and Brigo (2001) to calibrate the extended Vasicek, even though, as it will be clear later on, their model is equivalent to the Hull and White (1990) model. However their approach is more tractable and the calibration is done in more intuitive way. The advantage of their approach will be clearer when we calibrate the extended CIR model.

Recall that the extended Vasicek process could be written as

$$dr_t = \kappa (\theta - r_t) dt + \sigma_t dW_t$$

Following Mercurio and Brigo(2001) we define the spot rate as the sum of a deterministic fudge factor, which allow us later to fit the term structure curve, and a mean reverting stochastic process:

$$r_t = x_t + \phi(t) \quad (18)$$

$$dx_t = \kappa (\theta - x_t) dt + \sigma_t dW_t$$

Now, recall that the payer swaption is priced in the market by Black’s formula as following:

$$P_{\text{swaptionMkt}}(t, T_\alpha, T_\beta) = \text{Bl}(K, S_{\alpha,\beta}(t), \sigma_{\alpha,\beta} \sqrt{T_\alpha - t}) \sum_{i=\alpha+1}^{\beta} P(t, T_\alpha) \tau_i$$

where $\sigma_{\alpha,\beta}$ is the implied volatility quoted in the market, and $S_{\alpha,\beta}(t)$ is the swap rate between maturities $T_\alpha$ and $T_\beta$ and computed as $S_{\alpha,\beta}(t) = \frac{P(t, T_\alpha) - P(t, T_\beta)}{\sum_{i=\alpha+1}^{\beta} \tau_i P(0, T_i)}$. So now, to calibrate the volatility term structure, all that we need is to compute the value market and theoretical price of swaptions using Black (1976) and minimize the difference.
The price at time $t$ of a zero-coupon bond paying 1 dollar at time $T$ is:

$$P(t, T, r_t) = E_t \left[ e^{-\int_r^T r_s ds} \right]$$

$$= e^{-\int_r^T \phi_s ds E_t \left[ e^{-\int_r^T x_s ds} \right]}$$

$$= e^{-\int_r^T \phi_s ds Z(t, T)}$$

where $Z(t, T) = e^{-\frac{\theta - \mu}{2\sigma^2} \left( e^{-\kappa(T-t)} - e^{-\kappa t} \right) - \theta(T-t) + \frac{\sigma^2}{2} \int_r^T \kappa^2 \left( 1 - e^{-\kappa(T-u)} \right)^2 du}$

$E_t \left[ e^{-\int_r^T x_s ds} \right]$ could be interpreted as the price of the bond in the not shifted model.

Now the calibration to the zero-coupon curve is straightforward through:

$$\int_{t_{i-1}}^{t_i} \phi_s ds = \log \left( \frac{P(t, t_{i-1})/Z(t, t_{i-1})}{P(t, t_i)/Z(t, t_i)} \right)$$

So the calibration of $\phi$’s is done independently and automatically in this framework. Of course, this means that we already calibrated the $\sigma$’s and $\kappa$. This was done through a calibration to swaptions.

Now, let us see the relation between $\phi$’s and $\theta$’s. From equation (22) we have an analytical formula for $\phi_t$

$$\phi_t = f^M (0, t) - f^x (0, t)$$

where $f^M$ and $f^x$ denotes respectively the market forward rate and the process $x$ forward rate at time $t$.

More explicitly

$$\phi_t = f^M (0, t) - \theta \left( 1 - e^{-\kappa t} \right) - \int_0^T \sigma^2 e^{- \kappa (T-u)} \left[ 1 - e^{-\kappa(T-u)} \right] dW_u$$

Hence we have the following relation

$$\theta_t = \theta + \phi_t + \frac{1}{\kappa} \frac{d\phi_t}{dt}$$

Therefore the Brigo and Mercurio (2001) extension of the Vasicek Model is perfectly equivalent to that of Hull and White (1990).

3.4.4 Calibration to caplet prices.

In practice, many traders avoid using time dependent $\sigma$ since they have a better grasp on time-homogeneous parameters. The calibration of $\sigma$ is better done with caplet quotes since the caplet implied volatility captures only information about the underlying volatility whereas swaptions volatility captures the correlation information as well.

Recall from equation (16) that the price of a caplet could be written as:

$$P_{\text{caplet}} (t, T_{i-1}, T_i, K) = (1 + \tau_i K) E^Q \left[ \frac{e^{-\int_r^{T_i} r_s ds}}{P(T_{i-1}, T_i) \left( 1 + \tau_i K \right)} - \frac{1}{P(T_{i-1}, T_i)} \right]^{+} | \mathcal{F}_t$$
Following Brigo and Mercurio (2001) decomposition as defined by (18), this could be written as:

\[ P_{\text{caplet}}^{\text{HW}}(t, T_i, K) = (1 + \tau_i K) P(t, T_i) \Phi(h) - P(t, T_i-1) \Phi(h - \bar{\sigma}) \]

where \( h = \frac{1}{2} \ln \left( \frac{A(t, T_i) \exp(-B(t, T_i) x_0)}{X A(t, T_{i-1}) \exp(-B(t, T_{i-1}) x_0)} \right) + \frac{\bar{\sigma}}{2} \) with \( X = \frac{1}{1 + \tau_i K} \frac{P_{\text{Mkt}}^M(0, T_i-1) A(0, T_i) \exp(-B(t, T_i) x_0) - \bar{\sigma}^2 B(t, T_i)^2}{2 \kappa} \), \( \bar{\sigma} = \sigma B(t, T_{i-1}) \sqrt{\frac{1}{1 - \exp(-2 \kappa (T - t))}} \), \( A(t, T) = \exp \left[ \frac{(B(t, T) - T + t) \left( \kappa^2 \theta - \frac{\sigma^2}{2} \right) - \sigma^2 B(t, T)^2}{4 \kappa} \right] \), \( B(t, T) = \frac{1 - \exp(-\kappa (T - t))}{\kappa} \)

and \( \Phi \) denotes the cumulative normal distribution.

Define \( \zeta = (\kappa, \theta, \sigma) \) the set of variable we would like to find.

\[ \zeta = \arg \min \left( \sum_i \left( P_{\text{Mkt}}^\text{caplet} (t, T_{i-1}, T_i) - P_{\text{caplet}}^{\text{HW}}(t, T_i, T_i-1) \right)^2 \right) \]

### 3.5 Calibration in the CIR++ framework

Besides having analytical formulas for zero-coupon bonds, caplets and swaptions, the main advantage of using CIR++ is that it allows us to do the calibration in two steps. As, it will be clearer later on, the parameters \( \varphi \)'s will be found analytically through a calibration to the initial term structure curve. Then these parameters are plugged in the caplet formula so we can fit the caplet volatility term structure. Notice that we did not mention the swaption calibration. Indeed, in practice traders avoid pricing swaptions with CIR models since it is really cumbersome to provide a meaningful correlation structure in the noncentral chi-square distribution functions. However, we provide at the end of this section a way how to calibrate the model to swaption prices.

#### 3.5.1 Calibration to initial term structure curve

Following the same procedure as in the extended Vasicek case, the price of the bond in CIR++ framework could be written as:

\[
P(t, T, r_t) = E_t \left[ e^{-\int_t^T r_s ds} \right] = e^{-\int_t^T \varphi_s ds} E_t \left[ e^{-\int_t^T x_s ds} \right] = e^{-\int_t^T \varphi_s ds} A(t, T) \exp(-B(t, T) x_t)
\]

where \( A(t, T) = \left( \frac{2 h e^{(h + \kappa) (T-t) / 2} \sqrt{\pi}}{(h + \kappa)(e^{(h + \kappa) (T-t) / 2} - 1) + 2 h} \right)^{\frac{1}{2}} \), \( B(t, T) = \frac{2 \beta \left( e^{h(T-t) - 1} \right)}{(h + \kappa)(e^{(h + \kappa) (T-t) / 2} - 1) + 2 h} \) and \( h = \sqrt{\kappa^2 + \sigma^2} \).

Therefore
\[ \phi_t = f^M(0, t) - f^x(0, t) \]
\[ = f^M(0, t) - \frac{\kappa \theta (e^{\frac{\theta t}{h}} - 1)}{(h + \kappa) (e^{ht} - 1) + 2h} - x_0 \frac{4h^2 e^{\frac{\theta t}{2h}}}{[(h + \kappa) (e^{ht} - 1) + 2h]^2} \]

### 3.5.2 Caplet prices

As in the case of extended Vasicek the price of a caplet is given by (15)

\[ P_{\text{caplet}}(t, T_{i-1}, T_i) = P(t, T_{i-1}) (1 + \tau_i K) E^{Q^T_{i-1}} \left[ \left( \frac{1}{1 + \tau_i K} - P(T_{i-1}, T_i) \right)^+ | F_t \right] \]
\[ = (1 + \tau_i K) \text{ZBP} \left( t, T_{i-1}, T_i, \frac{1}{1 + K \tau_i} \right) \]

where \( \text{ZBP} \left( t, T_{i-1}, T_i, \frac{1}{1 + K \tau_i} \right) \) is the price of a put on a zero-coupon bond as defined by (6)

Define \( \zeta = (\kappa, \theta, \sigma) \) the set of variable we would like to find.

\[ \zeta = \arg \min \left( \sum_i \left( P_{\text{Mkt}} \left( t, T_{i-1}, T_i, r^* \right) - P_{\text{CIR++}} \left( t, T_{i-1}, T_i, K_i \right) \right)^2 \right) \]

where \( P_{\text{Mkt}} \left( t, T_{i-1}, T_i, r^* \right) \) is the time \( t \) market price of a caplet from \( T_{i-1} \) to \( T_i \), and \( P_{\text{CIR++}} \left( t, T_{i-1}, T_i, K_i \right) \) is the corresponding theoretical price.

### 3.5.3 Calibration to the swaption volatility structure

Similarly to the extended Vasicek case and since the extended CIR is a one factor model, the swaption could be priced as a sum of caplets through Jamshidian’s (1989) decomposition. First, we need to find \( r^* \) satisfying

\[ 1 = \sum_{i=\alpha+1}^{\beta} c_i P(T_{\alpha}, T_i, r^*) \]

where \( P(T_{\alpha}, T_i, r^*) \) is given by (5).

Now the swaption price is given by:

\[ P_{\text{swaption}}(t, T_{\alpha}, T_{\beta}) = \sum_{i=\alpha+1}^{\beta} c_i P_{\text{caplet}}(t, T_{\alpha}, T_i, K^i) \]

where \( c_i = \tau_i K \) for \( i = \alpha, ..., \beta - 1 \) and \( c_{\beta} = 1 + \tau_\beta K \), \( P_{\text{caplet}}(t, T_{\alpha}, T_i, K^i) \) is the price of a caplet with strike price \( K^i = P(T_{\alpha}, T_i, r^*) \).

Now, recall that the payer swaption is priced in the market by black’s formula as following

\[ P_{\text{swaptionMkt}}(t, T_{\alpha}, T_{\beta}) = Bl(K, S_{\alpha, \beta}(t), \sigma_{\alpha, \beta} \sqrt{T_{\alpha} - t}) \sum_{i=\alpha+1}^{\beta} P(t, T_\alpha) \tau_i \]
where $\sigma_{\alpha,\beta}$ is the implied volatility quoted in the market, and $S_{\alpha,\beta}(t)$ is the swap rate between maturities $T_\alpha$ and $T_\beta$ and computed as $S_{\alpha,\beta}(t) = \frac{P(t,T_\alpha) - P(t,T_\beta)}{\sum_{i=\alpha+1}^{\beta} \tau_i P(0,T_i)}$.

So now, the calibration amounts in minimizing the error between the theoretical and the market price.
4 Numerical experiments

4.1 Calibration

4.1.1 Extended Vasicek

We first calibrated the parameters of the extended Vasicek model to caplet prices where we supposed that the parameter \( \sigma \) is constant. The data used is 6 months at the money (ATM) caplets. Table (3) shows the caplet data used as quoted on 08/18/2004. Table (4) shows the parameters values obtained. Notice as far as order of magnitude is concerned, \( \sigma \) should be compared “roughly” to implied volatility time the short rate (since the market quotes caplet with Black’s formula i.e with a geometric brownian motion model).

Second, we calibrated the parameters of the extended Vasicek to the swaption prices. More precisely, we calibrated the model to swaptions on 5 years swap with maturities going from 1 year to 5 years quoted on 08/18/2004. We made this choice because, according to our knowledge, these are the most traded swaptions, maturities 7 and 10 lacking liquidity. The swaptions volatility quotes are shown in Table (1). The calibration results are presented in Table (2). Notice swaptions are quoted at the money, so our parameters do not capture any smile.

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Forward Swap Rate</th>
<th>Implied Volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.038776721</td>
<td>0.18</td>
</tr>
<tr>
<td>2</td>
<td>0.042744584</td>
<td>0.163</td>
</tr>
<tr>
<td>3</td>
<td>0.045702462</td>
<td>0.147</td>
</tr>
<tr>
<td>4</td>
<td>0.048003028</td>
<td>0.135</td>
</tr>
<tr>
<td>5</td>
<td>0.04976836</td>
<td>0.127</td>
</tr>
</tbody>
</table>

Table 2: HW parameters with sigma time dependent

<table>
<thead>
<tr>
<th>Maturity</th>
<th>( \sigma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.012284</td>
</tr>
<tr>
<td>2</td>
<td>0.0178</td>
</tr>
<tr>
<td>3</td>
<td>0.022894</td>
</tr>
<tr>
<td>4</td>
<td>0.031231</td>
</tr>
<tr>
<td>5</td>
<td>0.039647</td>
</tr>
<tr>
<td>( \kappa )</td>
<td>0.2125</td>
</tr>
</tbody>
</table>

4.1.2 CIR++

We calibrated the CIR++ to caplet prices. We used the same data as in the extended Vasicek case (Table (1)). The calibration results are presented in Table (5).
Table 3: ATM 6 months caplet implied volatility

<table>
<thead>
<tr>
<th>Forward rates</th>
<th>Implied Vol</th>
<th>Maturities (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.033237034</td>
<td>24.01</td>
<td>1.5</td>
</tr>
<tr>
<td>0.036523267</td>
<td>22.4</td>
<td>2</td>
</tr>
<tr>
<td>0.039450056</td>
<td>20.42</td>
<td>2.5</td>
</tr>
<tr>
<td>0.041441985</td>
<td>19.25</td>
<td>3</td>
</tr>
<tr>
<td>0.043223798</td>
<td>18.03</td>
<td>3.5</td>
</tr>
<tr>
<td>0.044960109</td>
<td>17.91</td>
<td>4</td>
</tr>
<tr>
<td>0.046513407</td>
<td>17.71</td>
<td>4.5</td>
</tr>
<tr>
<td>0.047864621</td>
<td>17.9</td>
<td>5</td>
</tr>
<tr>
<td>0.049133937</td>
<td>17.21</td>
<td>5.5</td>
</tr>
<tr>
<td>0.050309566</td>
<td>16.53</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 4: Extended Vasicek parameters values

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\kappa$</td>
<td>0.475384</td>
</tr>
<tr>
<td>$\theta$</td>
<td>0.0540725</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.0190654</td>
</tr>
</tbody>
</table>

Table 5: CIR++ parameters values

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\kappa$</td>
<td>0.256299</td>
</tr>
<tr>
<td>$\theta$</td>
<td>0.0587918</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.0632355</td>
</tr>
</tbody>
</table>

4.2 Convergence analysis

4.3 Price of the embedded options

5 Conclusion
References


6 Appendix

Proof of Computation of the 2 first moments of $\int_{t_1}^{t_2} r_s ds$.

Define $X(t_1, t') = \int_{t_1}^{t'} e^{K(t,s)} \sigma(u) dW_u$ and $Y(t_1, t_2) = \int_{t_1}^{t_2} e^{-K(t_1,t')} X(t_1, t') dt'$. Then the equation 2 could be written:

$$r' = e^{-K(t_1,t')} \left[ r_{t_1} + \int_{t_1}^{t'} e^{K(t_1,u)} \alpha_u du + X(t_1, t') \right]$$

and hence,

$$\int_{t_1}^{t_2} r_s ds = \int_{t_1}^{t_2} e^{-K(t_1,s)} \left[ r_{t_1} + \int_{t_1}^{s} e^{K(t_1,u)} \alpha_u du + X(t_1, s) \right] ds$$

Therefore

$$\mu_2(r_{t_1}, t_1, t_2) = \int_{t_1}^{t_2} e^{-K(t_1,s)} \left[ r_{t_1} + \int_{t_1}^{s} e^{K(t_1,u)} \alpha_u du \right] ds$$

Now the computation of the variance is a little bit trickier. Let’s compute first the general form of the covariance

$$\rho_Y(t_2, t_3) = \text{E} \left[ Y(t_1, t_2) Y(t_1, t_3) \right]$$

$$= \text{E} \left[ \int_{t_1}^{t_2} e^{-K(t_1,s)} X(t_1, s) ds \int_{t_1}^{t_3} e^{-K(t_1,z)} X(t_1, z) dz \right]$$

$$= \int_{t_1}^{t_2} \int_{t_1}^{t_3} e^{-K(t_1,x)} e^{-K(t_1,s)} \text{E} [X(t_1, s) X(t_1, z)] dz ds$$

$$= \int_{t_1}^{t_2} \int_{t_1}^{t_3} e^{-K(t_1,s)} e^{-K(t_1,z)} e^{2K(t_1,v)} \sigma^2(v) dv dz ds$$

$$= \int_{t_1}^{t_2} \int_{t_1}^{t_3} e^{-K(t_1,s)} e^{-K(t_1,z)} e^{2K(t_1,v)} \sigma^2(v) dv dz ds$$

$$+ \int_{t_1}^{t_2} \int_{t_1}^{t_3} e^{-K(t_1,s)} e^{-K(t_1,z)} \int_{t_1}^{t_3} e^{2K(t_1,v)} v \left( \int_{t_1}^{t_3} e^{-K(t_1,z)} dz \right) dv dz ds$$

$$= \int_{t_1}^{t_2} e^{2K(t_1,v)} \sigma^2(v) \left( \int_{t_1}^{t_3} e^{-K(t_1,z)} e^{-K(t_1,s)} dz ds \right) dv$$

$$+ \int_{t_1}^{t_2} e^{2K(t_1,v)} \sigma^2(v) \left( \int_{t_1}^{t_3} e^{-K(t_1,z)} e^{-K(t_1,s)} ds dz \right) dv$$

$$= \int_{t_1}^{t_2} e^{2K(t_1,v)} \sigma^2(v) \left( \int_{t_1}^{t_3} e^{-K(t_1,z)} dz \right) \left( \int_{t_1}^{t_2} e^{-K(t_1,s)} ds \right) dv$$

$$= \int_{t_1}^{t_2} e^{2K(t_1,v)} \sigma^2(v) \left( \int_{t_1}^{t_3} e^{-K(t_1,z)} dz \right) \left( \int_{t_1}^{t_2} e^{-K(t_1,s)} ds \right) dv$$

21
Therefore,

\[ \sigma_2^2(t_1, t_2) = \rho_Y(t_2, t_2) \]

\[ = \int_{t_1}^{t_2} e^{2K(t_1, v)}\sigma_2^2(v) \left( \int_{v}^{t_2} e^{-K(t_1, u)} du \right)^2 dv \]

**Proof of Theorem 4.** The Laplace transform of a noncentral \( X^2 \) random variable \( X \) is defined by

\[ \mathcal{L}_X(\eta) = E \left[ e^{-\eta X} \right] = (1 + 2\eta)^{-\frac{d}{2}} e^{-\frac{\eta^2}{\sigma^2}} \text{ for all } \eta \geq 0 \]

Now, from proposition 3, \( \mathcal{L}_X \left( t', t^*, 0, 1 \right) = E^Q \left[ e^{-r'v} \mid r' = a \right] = \left( \frac{2\sigma^2 + \Delta t}{(\sigma^2 + 2\kappa)(\sigma^2 + 1) + 2\kappa} \right)^{\frac{2\kappa}{\sigma^2}} \times \exp \left[ \frac{\kappa a}{(\sigma^2 + 2\kappa)(\sigma^2 + 1) + 2\kappa} \right] \]

where \( \Delta t = t^* - t' \)

Now, by identification and solving the following equations

\[
\begin{align*}
\frac{2\sigma^2 + \Delta t}{(\sigma^2 + 2\kappa)(\sigma^2 + 1) + 2\kappa} &= \frac{1}{\left(1 + 2\eta\right)} \\
\exp \left[ \frac{2\kappa a}{(\sigma^2 + 2\kappa)(\sigma^2 + 1) + 2\kappa} \right] &= \exp \left[ \frac{\lambda}{\pi(1 + 2\eta)} \right]
\end{align*}
\]

we obtain

\[ \lambda = \frac{4\kappa a}{\sigma^2 (e^{\kappa \Delta t} - 1)} \]

\[ d = \frac{4\kappa \theta}{\sigma^2} \]

\[ \eta = \frac{\sigma^2 (1 - e^{\kappa \Delta t})}{4\kappa} \]

**Proof of Theorem 5.** Recall that \( \left( r_{t_{m+1}}, \int_{t_m}^{t_{m+1}} r_t dt \right) \) is a gaussian vector with \( \mu \) and \( \Sigma \) as defined above.

For ease of notation we define \( X = r_{t_{m+1}}, Y = \int_{t_m}^{t_{m+1}} r_t dt, \mu_1 (a_k) = \mu_1 (a_k, t_m), \mu_2 (a_k) = \mu_2 (a_k, t_m, t_{m+1}), \sigma_1 = \sigma_1 (t_m), \sigma_2 = \sigma_2 (t_m, t_{m+1}) \) and \( \rho = \rho(t_m, t_{m+1}). \)

\[
A_{k,i} (t_m, t_{m+1}) = E_{a, a_k} \left[ e^{-Y} I_i (X) \right]
\]

\[
= \int_{R} \int_{R} \frac{1}{2\pi \sigma_1 \sigma_2 \sqrt{1 - \rho^2}} e^{-\frac{1}{2(1-\rho^2)} \left[ (x - \mu_1 (a_k))^2 + \frac{2\sigma_1^2}{\sigma_2^2} (y - \mu_2 (a_k))^2 \right]} e^{-y + \mu_2 (a_k)} dy dx
\]

\[
= e^{\mu_2 (a_k) + \frac{1}{2} \sigma_2^2} \int_{R} \int_{a_{k,i-1}}^{a_k} \frac{1}{2\pi \sigma_1 \sigma_2 \sqrt{1 - \rho^2}} e^{-\frac{1}{2(1-\rho^2)} \left[ (x - \mu_1 (a_k))^2 + \frac{2\sigma_1^2}{\sigma_2^2} (y - \mu_2 (a_k))^2 \right]} dy dx
\]

\[
= e^{\mu_2 (a_k) + \frac{1}{2} \sigma_2^2} \int_{R} \int_{a_{k,i-1}}^{\infty} \frac{1}{2\pi \sigma_1 \sigma_2 \sqrt{1 - \rho^2}} e^{-\frac{1}{2(1-\rho^2)} \left[ (x - \mu_1 (a_k))^2 + \frac{2\sigma_1^2}{\sigma_2^2} (y - \mu_2 (a_k))^2 \right]} dy dx
\]

\[
= e^{\mu_2 (a_k) + \frac{1}{2} \sigma_2^2} \left[ \Phi (\tilde{a}_{k,i}) - \Phi (\tilde{a}_{k,i-1}) \right] \text{ where } \tilde{a}_{k,i} = \frac{a_i - \mu_1 (a_k) - \sigma_1^2}{\sigma_2^2}
\]

\]

22
Similarly,

\[ B_{k,i}(t_m, t_{m+1}) = E_{t_m, ak} \left[ e^{-Y X_1(X)} \right] \]

\[ = \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{1}{2\pi \sigma_1 \sigma_2 \sqrt{1 - \rho^2}} e^{-\frac{1}{2(1-\rho^2)} \left\{ \frac{(x - \mu_1(ak))^2}{\sigma_1^2} + \frac{2(x - \mu_1(ak))(y - \mu_2(ak))}{\sigma_1 \sigma_2} + \frac{(y - \mu_2(ak))^2}{\sigma_2^2} \right\}} e^{-\mu_2(ak)} \left( x + \mu_1(ak) \right) dy dx \]

\[ = \left( \int_{\mathbb{R}} \int_{a_{k-i-1}}^\infty \frac{1}{2\pi \sigma_1 \sigma_2 \sqrt{1 - \rho^2}} e^{-\frac{1}{2(1-\rho^2)} \left\{ \frac{(x - \mu_1(ak))^2}{\sigma_1^2} + \frac{2(x - \mu_1(ak))(y - \mu_2(ak))}{\sigma_1 \sigma_2} + \frac{(y - \mu_2(ak))^2}{\sigma_2^2} \right\}} e^{-\mu_2(ak)} \left( x + \mu_1(ak) \right) dy dx \right) \]

\[ - \mu_2(ak) + \frac{1}{2} \sigma_2^2 \int_{a'_{k-i-1}}^\infty \frac{x e^{-\frac{(x - \mu_1(ak))^2}{2\sigma_1^2}}}{\sqrt{2\pi} \sigma_1} e^{-\frac{1}{2(1-\rho^2)} \left\{ \frac{(x - \mu_1(ak))^2}{\sigma_1^2} + \frac{2(x - \mu_1(ak))(y - \mu_2(ak))}{\sigma_1 \sigma_2} + \frac{(y - \mu_2(ak))^2}{\sigma_2^2} \right\}} \frac{dy}{\sigma_1^2} \]

where \( y' = \rho(x - \sigma_1) + \frac{y + \sigma_2^2}{\sigma_1^2} \)

\[ = \mu_2(ak) + \frac{1}{2} \sigma_2^2 \int_{a'_{k-i-1}}^\infty \frac{\sigma_1 x' + \mu_1(ak) + \frac{\sigma_1 e^{-\frac{(x - \mu_1(ak))^2}{2\sigma_1^2}}}{\sqrt{2\pi} \sigma_1} e^{-\frac{1}{2(1-\rho^2)} \left\{ \frac{(x - \mu_1(ak))^2}{\sigma_1^2} + \frac{2(x - \mu_1(ak))(y - \mu_2(ak))}{\sigma_1 \sigma_2} + \frac{(y - \mu_2(ak))^2}{\sigma_2^2} \right\}}}{\sigma_1^2} dx' \]

\[ - \mu_2(ak) + \frac{1}{2} \sigma_2^2 \int_{a'_{k-i-1}}^\infty \frac{\sigma_1 x' + \mu_1(ak)}{\sqrt{2\pi} \sigma_1} e^{-\frac{1}{2(1-\rho^2)} \left\{ \frac{(x - \mu_1(ak))^2}{\sigma_1^2} + \frac{2(x - \mu_1(ak))(y - \mu_2(ak))}{\sigma_1 \sigma_2} + \frac{(y - \mu_2(ak))^2}{\sigma_2^2} \right\}} dx' \]

\[ = \mu_2(ak) + \frac{1}{2} \sigma_2^2 \left[ \frac{\sigma_1}{\sqrt{2\pi}} \left( e^{-\frac{(x_{k-1})^2}{2}} + e^{-\mu_2(ak)} \int_{a'_{k-i-1}}^\infty \frac{\sigma_1 x' + \mu_1(ak) + \sigma_1 e^{-\frac{(x - \mu_1(ak))^2}{2\sigma_1^2}}}{\sqrt{2\pi} \sigma_1} e^{-\frac{1}{2(1-\rho^2)} \left\{ \frac{(x - \mu_1(ak))^2}{\sigma_1^2} + \frac{2(x - \mu_1(ak))(y - \mu_2(ak))}{\sigma_1 \sigma_2} + \frac{(y - \mu_2(ak))^2}{\sigma_2^2} \right\}} dx' \right) \]

\[ - \mu_2(ak) + \frac{1}{2} \sigma_2^2 \left[ \frac{\sigma_1}{\sqrt{2\pi}} \left( e^{-\frac{(x_{k-1})^2}{2}} + e^{-\mu_2(ak)} \int_{a'_{k-i-1}}^\infty \frac{\sigma_1 x' + \mu_1(ak) + \sigma_1 e^{-\frac{(x - \mu_1(ak))^2}{2\sigma_1^2}}}{\sqrt{2\pi} \sigma_1} e^{-\frac{1}{2(1-\rho^2)} \left\{ \frac{(x - \mu_1(ak))^2}{\sigma_1^2} + \frac{2(x - \mu_1(ak))(y - \mu_2(ak))}{\sigma_1 \sigma_2} + \frac{(y - \mu_2(ak))^2}{\sigma_2^2} \right\}} dx' \right) \]

\[ = \mu_2(ak) + \frac{1}{2} \sigma_2^2 \left[ \frac{\sigma_1}{\sqrt{2\pi}} \left( e^{-\frac{(x_{k-1})^2}{2}} + e^{-\mu_2(ak)} \int_{a'_{k-i-1}}^\infty \frac{\sigma_1 x' + \mu_1(ak) + \sigma_1 e^{-\frac{(x - \mu_1(ak))^2}{2\sigma_1^2}}}{\sqrt{2\pi} \sigma_1} e^{-\frac{1}{2(1-\rho^2)} \left\{ \frac{(x - \mu_1(ak))^2}{\sigma_1^2} + \frac{2(x - \mu_1(ak))(y - \mu_2(ak))}{\sigma_1 \sigma_2} + \frac{(y - \mu_2(ak))^2}{\sigma_2^2} \right\}} dx' \right) \]

\[ - \mu_2(ak) + \frac{1}{2} \sigma_2^2 \left[ \frac{\sigma_1}{\sqrt{2\pi}} \left( e^{-\frac{(x_{k-1})^2}{2}} + e^{-\mu_2(ak)} \int_{a'_{k-i-1}}^\infty \frac{\sigma_1 x' + \mu_1(ak) + \sigma_1 e^{-\frac{(x - \mu_1(ak))^2}{2\sigma_1^2}}}{\sqrt{2\pi} \sigma_1} e^{-\frac{1}{2(1-\rho^2)} \left\{ \frac{(x - \mu_1(ak))^2}{\sigma_1^2} + \frac{2(x - \mu_1(ak))(y - \mu_2(ak))}{\sigma_1 \sigma_2} + \frac{(y - \mu_2(ak))^2}{\sigma_2^2} \right\}} dx' \right) \]

\[ \right] \]
Proof of Theorem 8. Under the forward measure $\tilde{Q}$, $\frac{x_{m+\Delta t}^n}{\eta}$ has the non-central chi-square distribution with the non-centrality parameter $\tilde{\lambda}$ and with $d$ degrees of freedom where

$$\tilde{\lambda} = \frac{8h^2 e^{(t_{m+1}-t_m)\gamma}}{\sigma^2 ((h + \kappa) (e^{h(t_{m+1}-t_m)} - 1) + 2h}$$

where $h$, $d$ and $\eta$ are defined as in the theorem’s statement.

Therefore

$$A_{k,i}(t_m,t_{m+1}) = E_{m,a_k}\left[e^{-\int_{t_m}^{t_{m+1}} r_s ds} I\left(a_i \leq r_{t_m+1} \leq a_i+1\right)\right]$$

$$= e^{-\int_{t_m}^{t_{m+1}} \varphi_s ds} E_{m,a_k}\left[e^{-\int_{t_{m+1}}^{t_{m+1}} (r_s - \varphi_s) ds} I\left(a_i - \varphi_{t_{m+1}} \leq r_{t_{m+1}} - \varphi_{t_{m+1}} \leq a_i+1 - \varphi_{t_{m+1}}\right)\right]$$

$$= e^{-\int_{t_m}^{t_{m+1}} \varphi_s ds} E_{m,a_k}\left[e^{-\int_{t_{m+1}}^{t_{m+1}} x_s ds} I\left(a_i - \varphi_{t_{m+1}} \leq x_{t_{m+1}} \leq a_i+1 - \varphi_{t_{m+1}}\right)\right]$$

$$= P(t_m,t_{m+1},a_k) E_{m,a_k}\left[e^{-\int_{t_m}^{t_{m+1}} r_s ds} I\left(a_i \leq r_{t_m+1} \leq a_i+1\right)\right]$$

$$= P(t_m,t_{m+1},a_k) \times \sum_{n=0}^{\infty} e^{-\frac{(\lambda h/2)^n}{n!}} \left(\frac{\lambda h/2}{n!}\right)^n \left(F_{d+2n} \left(\frac{a_i+1 - \varphi_{t_{m+1}}}{\eta}\right) - F_{d+2n} \left(\frac{a_i - \varphi_{t_{m+1}}}{\eta}\right)\right)$$

$$B_{k,i}(t_m,t_{m+1}) = E_{m,a_k}\left[e^{-\int_{t_m}^{t_{m+1}} r_s ds} \varphi_{t_{m+1}} I\left(a_i \leq r_{t_m+1} \leq a_i+1\right)\right]$$

$$= P(t_m,t_{m+1},a_k) \times \sum_{n=0}^{\infty} e^{-\frac{(\lambda h/2)^n}{n!}} \left(\frac{\lambda h/2}{n!}\right)^n \left(F_{d+2n} \left(\frac{a_i+1 - \varphi_{t_{m+1}}}{\eta}\right) - F_{d+2n} \left(\frac{a_i - \varphi_{t_{m+1}}}{\eta}\right)\right)$$

Notice that the properties of the non-central $X^2$ distribution

$$E_{m,a_k}\left[\left(\frac{a_i - \varphi_{t_{m+1}}}{\eta}\right) \leq \frac{x_{t_{m+1}}}{\eta} \leq \frac{a_i+1 - \varphi_{t_{m+1}}}{\eta}\right]\right] = \sum_{n=0}^{\infty} e^{-\frac{(\lambda h/2)^n}{n!}} \left(\frac{\lambda h/2}{n!}\right)^n \left[F_{d+2n} \left(\frac{a_i+1 - \varphi_{t_{m+1}}}{\eta}\right) - F_{d+2n} \left(\frac{a_i - \varphi_{t_{m+1}}}{\eta}\right)\right]$$

where

$$Q_n(a_{i+1},a_i) = -2 \left(a_{i+1} f_{d+2n} \left(\frac{a_i+1}{\eta}\right) - a_{i} f_{d+2n} \left(\frac{a_i}{\eta}\right)\right) + (d + 2n) \times \left(F_{d+2n} \left(\frac{a_i+1}{\eta}\right) - F_{d+2n} \left(\frac{a_i}{\eta}\right)\right)$$

with $h = \sqrt{\kappa^2 + 2\sigma^2}$ and $d = \frac{\kappa}{\sqrt{\sigma}}$ and $F_{d+2n}$ and $f_{d+2n}$ are respectively the cumulative and density functions of an khi-2 with $d + 2n$ degrees of freedom model.

This implies

$$B_{k,i}(t_m,t_{m+1}) = P(t_m,t_{m+1},a_k) \times \sum_{n=0}^{\infty} e^{-\frac{(\lambda h/2)^n}{n!}} \left(\frac{\lambda h/2}{n!}\right)^n \left(F_{d+2n} \left(\frac{a_i+1 - \varphi_{t_{m+1}}}{\eta}\right) - F_{d+2n} \left(\frac{a_i - \varphi_{t_{m+1}}}{\eta}\right)\right)$$

24