PRICING CONVERTIBLE BONDS USING FINITE ELEMENTS

Chris Bardgett

Master of Advanced Studies in Finance
ETH Zürich
and
University of Zürich

Supervisor
Prof. Dr. Ch. Schwab

(December 2007)

Abstract

This paper discusses the pricing of convertible bonds using Finite Element Methods. An overview of the features of this financial product is given. We start by presenting a one factor (stock price) model that allows to consider default. Then, we show that the value of a convertible bond satisfies a Linear Complementarity Problem (LCP), which leads to a parabolic variational inequality (PVI). The Finite Element Methods are applied to solve this PVI. To discretize the variational inequality, we apply a theta scheme in time and a Galerkin method in log-price space. Results are explained and compared to those of a commercial pricing software. Finally, the theoretical framework of the one factor pricing model is extended to include a stochastic interest rate.

Acknowledgments

I would like to thank my supervisor Prof. Christoph Schwab and his assistant Norbert Hilber from the department of Mathematics of the ETH Zürich, for their support, comments and fruitful discussions during the writing of this thesis.

Contents

Part I
Introduction

Convertible bonds (CB) are issued by companies in order to raise capital. Basically, when a company needs money, it can issue debt or equity instruments or a combination of both. A bond (debt) is an instrument paying the holder coupons at particular dates and the principal at maturity. In order to define a convertible bond we first have to explain the concept of an option. When someone buys an option on a company's stock, he purchases the right to buy, for a call, (resp. sell, for a put) a share for a price that is fixed at the beginning of the contract. An option can be European (if you can exercise your right only at maturity) or American (if you can exercise anytime before maturity). A convertible bond is a bond that has additional characteristics: the holder can exchange the bond for a certain number of common stocks during a certain period or at certain dates. It is almost as if there were an option, embedded in the bond, to buy the stock. If converted at time t, the holder of the convertible will not receive the principal or any coupons that would have been delivered at a date later than t. In some sense a convertible enables the holder to benefit from the high performance of the stock (conversion) and to be protected from a drop in the stock value by the bond (the convertible has a higher value than the simple bond with no option to convert). However, it is very important not to forget the credit risk exposure of the convertible. Indeed, some of the companies issuing this type of instrument may have a very low credit quality and therefore issue no vanilla bonds.

The market for convertibles has been growing for the last decade. In 2000, around $60 billion of new convertibles were issued; in 2001 it almost doubled and the number of issuance exploded in 2003 as mentioned in [AFV03].

The aim of this thesis is to price a convertible bond. The CB is a hybrid product since it depends on many different factors: stock price, interest rates, credit risk, all of which makes the pricing of this type of product difficult.

There are two different approaches to pricing convertible bonds. The first that was suggested is a "structural approach": the basic underlying factor is the value of the firm (issuing the stock). It is possible to find an interesting overview of this type of model in [N96] Nyborg (1996). One issue that is encountered in this type of model is that the value of the firm is not a traded asset and the parameters are difficult to estimate.
The other type of modelling is the "reduced-form" approach. The main underlying risk-factor is the share price of the firm issuing the stock. Recently, many authors have developed this modelling approach: Tsiveriotis and Fernandes (1998), Davis and Lischka (1999), [TKN01] Takahashi et al. (2001), [AFV02] Ayache et al. (2002), or also [AB02] Andersen et al. (2002).

In this thesis, we will consider bonds that the holder can convert any time from now to maturity (American style conversion). Moreover, convertible bonds often have many other features and we would like to focus on some of the most common ones: the call and put features. The call feature enables the issuer of the bond to call (buy) it back at a pre-specified price that can depend on time. It is somehow a way to limit the profits of the holder. And the put feature gives the right to the holder to sell the CB back to the issuer for a predefined price (also possibly depending on time) which floors the price of the CB. To summarize, we will consider a CB that can be converted any time with a call and put provision, both of them exercisable anytime from the beginning to maturity.

The topic of this thesis is to price a defaultable convertible bond based on [AFV03] and [AB02], where the stock price and short rate can be stochastic. The first step is to consider one stochastic factor: the stock price. Then, we would like to to extend the framework and include a stochastic interest rate. We have decided to use the Finite Element Methods to price such a product since it provides a complete and rigorous framework to solve partial differential equations (PDEs) and parabolic variational inequalities. Moreover, we believe that the Finite Element Methods (FEM) provide a mathematical framework that is more general than that of the Finite Difference Methods (FDM) since we work on spaces that are less restrictive. Often, FEM allow to achieve more numerical stability.

We will illustrate analogies between the pricing of American options and CBs with an American style conversion to better explain the pricing method for CBs.

Part II
Model Description

In this part we will explain the models that we implement for the pricing of an American option and a CB. We will start, in section one, by working in a one dimensional framework where the stock price is stochastic. Then, in section two, we will consider a stochastic interest rate.

Most of the modelling that we will use here -for the one-dimension framework- is based on the papers of Ayache, Forsyth and Vetzal (see [AFV02], [AFV03]) for the convertible bond. For the American option pricing we refer to [Schw07] and [JLL90].

1 One stochastic factor: stock price

We will describe the model for pricing an American option and a defaultable convertible in the Black Scholes framework. We will start with the pricing system for the American option.

1.1 American option

In the Black and Scholes framework we consider a riskless bond and a share price denoted by S_t paying no dividends. We assume that S_t has the following dynamics under the historical measure $\mathbb{P}$:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \ , \quad S(0) = S_0$$

where μ represents the drift of the stock price, σ its volatility and $(W_t)_{t\geq 0}$ a Brownian motion under $\mathbb{P}$.

Under the martingale measure $\mathbb{Q}$, the dynamics of the price process becomes

$$dS_t = rS_t dt + \sigma S_t d\tilde{W}_t$$

where r is the risk-free rate and $(\tilde{W}_t)_{t\geq 0}$ a Brownian motion under $\mathbb{Q}$.

Let us consider the value $V^{am}(S,t)$ of an American option with maturity T. If $g(S)$ denotes a payoff function, optimal exercising is equivalent to an optimal stopping problem and V^{am} is given by:

$$V_t^{am}(S_t,t) = \sup_{t\leq\tau\leq T} \mathbb{E}[\exp^{-r(\tau-t)} g(S_\tau)|\mathcal{F}_t]$$

where the supremum is taken over all stopping times τ defined on the probability space of S_t.

Denoting the strike price E, the payoff of the call option is $g(S) = \max(S - E, 0)$, and the payoff of a put option is $g(S) = \max(E - S, 0)$.

For g sufficiently regular (see [LL97], chapter 5, section 3 or [Schw07]) and $\sigma \neq 0$, it can be shown that $V^{am}(S,t)$ satisfies the parabolic variational inequality in $\mathbb{R}_+ \times (0,T)$:

$$\begin{cases} \frac{\partial V^{am}}{\partial t} + \underbrace{\frac{\sigma^2}{2} S^2 \frac{\partial^2 V^{am}}{\partial S^2} + rS\frac{\partial V^{am}}{\partial S}}_{A^{BS}V^{am}} - rV^{am} \leq 0 \\ V^{am}(S,t) \geq g(S) \\ (V^{am} - g) \cdot (\frac{\partial V^{am}}{\partial t} + A^{BS}V^{am} - rV^{am}) = 0 \end{cases} \tag{II.1}$$

subject to the terminal condition

$$V^{am}(S,T) = g(S).$$

1.2 The AFV Model for convertible bonds

1.2.1 Idea

The AFV model presented in [AFV02] and [AFV03] is a single factor framework for the valuation of convertible bonds which can be extended to a multi-dimensional framework. The stock price is stochastic and follows the Black-Scholes dynamics until default. Indeed, the authors of [AFV02], [AFV03] include a homogeneous Poisson process for which the time of first jump corresponds to the time of default of the company.

In their paper, Ayache, Forsyth and Vetzal consider precisely what happens when default occurs: at time of default τ_D the stock price process jumps, but not necessarily to 0: $S_{\tau_D^+} = S_{\tau_D^-}(1-\eta)$, where $0 \leq \eta \leq 1$. Then $\eta S_{\tau_D^-}$ represents the loss in stock price at default.

Moreover, at τ_D, the holder of the CB has the option to convert into shares (with value $S_{\tau_D^+}$) or else to recover a certain amount of the CB value: $R_{rec} \cdot X$, where R_{rec} is the recovery rate ($0 \leq R_{rec} \leq 1$) and X could be for instance the face value of the bond.

They consider a CB that can possibly pay coupons and can be exchanged any time for a certain fixed number of shares. The CB has a put and call provision (also American style). The stock may pay a continuous dividend yield.

1.2.2 Model

As in [AFV03], we will consider a stock paying dividends that has a diffusion and a jump term. The dynamics under a risk neutral measure $\mathbb{Q}$ is

$$dS_t^+ = S_t^-(r - q + p\eta)dt + S_t^-\sigma dW_t - S_t^-\eta dN_t \ , \quad S(0) = S_0 \tag{II.2}$$

where r represents the short rate, σ the volatility of the stock and q the continuous dividend yield.

The process $(W_t)_{t\geq 0}$ is a Brownian motion under $\mathbb{Q}$ and $(N_t)_{0\leq t\leq T}$ is a Poisson process with intensity p that models the instant of default. In fact we force $(N_t)_{0\leq t\leq T}$ to remain constant after the first jump. Therefore, the jump models the time of default and no other possible jumps in the stock price. The parameter p is defined as follows:

$$\mathbb{Q}(\tau_D \in (t, t+\Delta t)|\tau_D > t) = \mathbb{E}_{\mathbb{Q}}[\int_t^{t+\Delta t} p(s)\, ds|\ \mathcal{F}_t] \approx p(t)\Delta t.$$

We consider r, q, p and σ constant in this section.

Let us note that before default occurs, the dynamics (II.2) is equivalent to

$$dS_t = (r - q + p\eta)S_t dt + \sigma S_t dW_t$$

It is important to know that in this framework, the market is incomplete due to the jump process, and therefore the uniqueness of a martingale measure $\mathbb{Q}$ is not ensured. We assume that we have chosen a measure (denoted $\mathbb{Q}$ since the beginning) and that it is fixed.

As said before, the CB can be called (bought back) any time until maturity, at a price $B_c(t)$. There is also a put provision which enables the holder to sell the CB any time until maturity for a price $B_p(t)$ (with $B_c > B_p$). And finally, it is possible to convert the CB to a constant number κ of shares any time.

The accrual interest at time t represents the interest associated with the pending coupon. If we denote by $(t_{c_j})_{1\leq j\leq N_c}$ the coupon dates, we can calculate the accrued interest at anytime $AccI(t)$:

$$AccI(t) = K\,\frac{t - t_{c_i}}{t_{c_{i+1}} - t_{c_i}} \ , \quad t_{c_i} \leq t \leq t_{c_{i+1}} \ \ , \ i+1 \leq N_c \ ,$$

where K denotes the value of coupons paid to the CB holder.

As we will see later on, we will need the accrual interest for the calculation of B_c and B_p.

Furthermore, we can represent the coupon payment stream as a function $c(t) = \sum_{p=1}^{N_c} K\delta_{t_{c_p}}(t)$ where δ represents a *Dirac function*. For simplicity, if there are coupons, then they are paid until maturity, in the sense that the last coupon is paid at maturity.

If we denote by $V(S,t)$ the value of the CB, the authors state in their paper [AFV03] that V follows the Linear Complementarity Problem (LCP):

- when $B_c > \kappa S$

$$\begin{pmatrix} \mathcal{M}V - \phi_X(S) - c(t) = 0 \\ V - \max(B_p, \kappa S) \geq 0 \\ V - B_c \leq 0 \end{pmatrix} \vee \begin{pmatrix} \mathcal{M}V - \phi_X(S) - c(t) \geq 0 \\ V - \max(B_p, \kappa S) = 0 \\ V - B_c \leq 0 \end{pmatrix} \vee \begin{pmatrix} \mathcal{M}V - \phi_X(S) - c(t) \leq 0 \\ V - \max(B_p, \kappa S) \geq 0 \\ V - B_c = 0 \end{pmatrix} \tag{II.3}$$

- when $B_c \leq \kappa S$

$$V = \kappa S$$

for S in $[0, \infty)$ and t in $[0, T]$.

In (II.3), $\mathcal{M}$ denotes the operator

$$\mathcal{M} := -\frac{\partial}{\partial t} - [\frac{\sigma^2}{2} S^2 \frac{\partial^2}{\partial S^2} + (r + p\eta - q) S \frac{\partial}{\partial S} - (r + p)], \tag{II.4}$$

and

$$\phi_X(S) := p \max(\kappa S(1 - \eta), R_{rec} X).$$

The notation $(x = 0) \vee (y = 0) \vee (z = 0)$ means that at least one of $(x = 0)$, $(y = 0)$, $(z = 0)$ holds at each point of the solution domain.

The constraints on the value of the CB reflect the fact that

$$\max(B_p(t), \kappa S) \leq V \leq \max(B_c(t), \kappa S)$$

Denoting by F the face value of the bond, the terminal condition on V is $V(S,T) = \max(\kappa S, F + K, B_p(T))$. At maturity, in case the CB has not defaulted, the holder will have a choice between

converting his CB into κ shares, selling the bond for $B_p(T)$ or receiving the $(F+K)$ (nominal + coupon).

In [AFV03], the equation $\mathcal{M}V - p\max(\kappa S(1-\eta), R_{rec}X) - c(t) = 0$ is derived using a no-arbitrage opportunity argument after building a risk free portfolio as it is done in the case of a European option.

As said in [AFV02], the intuition behind (II.3) is that the value of the convertible is given by the solution to $\mathcal{M}V - p\max(\kappa S(1-\eta), R_{rec}X) - c(t) = 0$, subject to the constraints above. When this equation is equal to 0, we are in the continuation region (no constraint is binding), otherwise either the put constraint - middle term - or the call constraint - right term - is binding in (II.3).

Furthermore, let us explain the inequality constraints on V in the following way: on the one hand $V \geq \max(B_p(t), \kappa S)$ has to hold, because at any time the holder has the right to convert (conversion value is κS) or put the bond (for $B_p(t)$).

On the other hand, if the conversion price κS is higher than the call price $B_c(t)$, then the company would call the bond immediately and the holder would decide to convert it (instead of getting $B_c(t)$) to receive $\kappa S \geq B_c$. And in the case the conversion price κS is less than the call price $B_c(t)$, then since the company has the right to call it at the price $B_c(t)$, the value of the CB is necessarily less than $B_c(t)$. Therefore $V \leq B_c(t) = \max(B_c(t), \kappa S)$.

1.2.3 Recovery under the AFV model

In order to be able to solve the system written in the previous section, it is necessary to make an assumption concerning the parameter X. X defines what the holder recovers in case of default.

The assumption made in [AFV03] is that upon default, the holder of the CB can recover $R_{rec}B$ where B is defined as the "pre-default bond component of the convertible". We can then divide the value of the CB into the sum of two components: $V = B + C$ where C is the equity component, and simply $V - B$.

We have now created two artificial instruments that together form the CB. The authors propose the following decomposition for the system of equations and inequalities applying to V:

- when $B_c > \kappa S$

$$\begin{pmatrix} \mathcal{M}B - pR_{rec}B - c(t) = 0 \\ B - B_c \leq 0 \\ B + C - B_p \geq 0 \end{pmatrix} \vee \begin{pmatrix} \mathcal{M}B - pR_{rec}B - c(t) \leq 0 \\ B = B_c \end{pmatrix} \vee \begin{pmatrix} \mathcal{M}B - pR_{rec}B - c(t) \geq 0 \\ B + C = B_p \end{pmatrix}$$

and:

$$\begin{pmatrix} \mathcal{M}C - \phi_B(S) + pR_{rec}B = 0 \\ B + C - \max(B_c, \kappa S) \leq 0 \\ B + C - \kappa S \geq 0 \end{pmatrix} \vee \begin{pmatrix} \mathcal{M}C - \phi_B(S) + pR_{rec}B \leq 0 \\ B + C = \max(B_c, \kappa S) \end{pmatrix} \vee \begin{pmatrix} \mathcal{M}C - \phi_B(S) + pR_{rec}B \geq 0 \\ B + C = \kappa S \end{pmatrix}$$

- when $B_c \leq \kappa S$

$$V = B + C = \kappa S$$

for S in $[0; \infty)$ and t in $[0; T]$. The operator $\mathcal{M}$ has been defined (II.4).

The authors of [AFV03] decompose the constraints on the real product V into new constraints on B and C:

$$\begin{aligned} &B \leq B_c(t) \\ &B \geq B_p(t) - C \\ &C \leq \max(B_c(t), \kappa S) - B \\ &C \geq \kappa S - B \end{aligned}$$

The authors suggest the terminal conditions (no put provision at maturity)

$$\begin{aligned} &B(S, T) = F + K \\ &C(S, T) = \max(\kappa S - (F + K), 0) \end{aligned}$$

1.3 A model with constant recovery

As highlighted in the AFV paper, we can make any assumption concerning the recovery upon default. For instance, in the AFV model, the authors assume the holder recovers a portion of the bond component of the convertible. We choose to consider another model where the holder recovers a fraction of the face value of the bond at default $R_{rec}F$. This is also a standard assumption (see [AFV02]) and this is the model we have decided to implement.

We make the same assumptions as are presented in the previous section. To be clear we recall some assumptions:

- The call feature is American. The price at which the company can call the CB $B_c(t)$ may depend on time. We will consider the case where the call price is a constant plus accrued interest $B_c(t) = B_{cc} + AccI(t)$ where B_{cc} is the (constant) clean price. This means that when the holder sells the CB, he gets the accrued interest which is common in the bond markets.
- The put feature is American. The put price will depend on time. We will choose $B_p(t) = B_{pp} + AccI(t)$ where B_{pp} is the (constant) clean price of B_p.
- The conversion is American style. The holder does not receive accrual interest when he decides to convert into shares. This assumption is made in [AFV02].
- The bond can pay coupons. If coupons are paid, to ease implementation, we make the assumption a coupon is paid at maturity.
- The parameters and assumptions about the stock price dynamics are the same as in part II, section 1.2.2.

Let us denote by B the straight bond with same face value, tenor structure, coupons, as the CB we want to price. We assume that B has a call feature identical to that of the CB. This bond can default and there is no recovery in case of default.

Since the CB has a conversion and put feature embedded, it has necessarily more value than the straight bond B by absence of arbitrage opportunity. Therefore, we can write an additional constraint:

$$V(S,t) \geq B(t) \qquad , \ \forall t \in [0,T], \forall S \in [0,\infty)$$

where V is the value of the CB. We assume $B(t) < B_c(t)$ to ensure the CB is not called immediately.

Therefore we can include an additional term in the lower constraint of V. We use the result of [AFV03] to find

$$\begin{pmatrix} \mathcal{M}V - \phi_F(S) - c(t) = 0 \\ V \geq g_2 \\ V \leq g_1 \end{pmatrix} \vee \begin{pmatrix} \mathcal{M}V - \phi_F(S) - c(t) \geq 0 \\ V = g_2 \\ V \leq g_1 \end{pmatrix} \vee \begin{pmatrix} \mathcal{M}V - \phi_F(S) - c(t) \leq 0 \\ V \geq g_2 \\ V = g_1 \end{pmatrix} \tag{II.5}$$

$$V(S,T) = g_2(S,T) \tag{II.6}$$

for S in $[0,\infty)$ and t in $[0,T]$.

The constraints g_1 and g_2 for the LCP are given by:

$$g_1 = \max(B_c, \kappa S) \tag{II.7}$$
$$g_2 = \max(B_p, B(t), \kappa S) \tag{II.8}$$

At maturity, in case there was no default, the bond holder will receive $\max(B(T), \kappa S, B_p(T))$ since the holder can either choose to receive the terminal value of the callable bond B, put the CB or else convert.

A sketch of the proof that there exists a unique solution to the LCP above is given in Appendix A. It is based on the equivalence between stochastic optimal control and a LCP. This equivalence is based on [BL78].

2 Two stochastic factors: stock price, interest rate

In this section, we include a stochastic interest rate which may be correlated with the stock price. The most common short rate models are the Vasicek, Cox Ingersoll Ross, Ho Lee, Hull White models, etc. (see [BM01]).

2.1 Model specification

We are in the same framework as in the one-factor model except the interest rate is allowed to be stochastic. The dynamics of the factors under a risk neutral measure $\mathbb{Q}$ is then

$$\left.\begin{aligned} dS_t^+ &= S_t^-(r - q + p\eta)dt + S_t^-\sigma dW_t - S_t^-\eta dN_t, \qquad S(0) = S_0 \\ dr_t &= \tilde{\mu}(r,t)dt + \tilde{\sigma}(r,t)d\tilde{W}_t, \qquad r(0) = r_0 \end{aligned}\right\} \tag{II.9}$$

where we denote by $\rho_{(r,S)}$ the correlation between the Wiener processes W_t and $\tilde{W}_t$ where $0 \leq \rho_{(r,S)} \leq 1$. We will assume that $\rho_{(r,S)}$ is constant.

A set of partial differential inequalities can then be derived for the price of a convertible with a stochastic short rate (see for instance for the PDE [AB02] section 6).

A LCP can be derived extending the framework presented in part II. The value V of a CB under the model (II.9) is a solution of the LCP (II.5)-(II.8) where the operator $\mathcal{M}$ in (II.5) is

replaced by the operator $\mathcal{M}_2$:

$$\mathcal{M}_2 = \begin{cases} -\frac{\partial}{\partial t} - [\frac{\sigma^2}{2} S^2 \frac{\partial^2}{\partial S^2} + \frac{\tilde{\sigma}^2}{2} \frac{\partial^2}{\partial r^2} + \rho_{(r,S)} \sigma \tilde{\sigma} S \frac{\partial^2}{\partial S \partial r} + (r + p\eta - q) S \frac{\partial}{\partial S} \\ +\tilde{\mu}(r,t) \frac{\partial}{\partial r} - (r+p)] \end{cases} \tag{II.10}$$

Let us note that the bond price B in constraint (II.8) has to be computed also using the model (II.9) to maintain consistency.

We can see that the LCP is similar to the case where only the stock price is stochastic.

Part III
Finite Element Approach

1 Objective

In many papers, the method used to solve a PDE is the Finite Difference Methods (FDM). We have decided to use the Finite Element Methods (FEM) to solve this pricing problem.

There are several reasons that explain the use of FEM.

First, it is possible to consider grids that are not necessarily rectangular whereas FDM, in the basic form, are restricted to rectangular shapes. Second, since we work with Sobolev spaces to find solutions, we use the concept of weak derivatives and the notion of differentiability is more general than with the FDM. The FEM is therefore more general than the FDM and it often ensures more numerical stability and accuracy.

2 One stochastic factor: stock price

In this section, we will show how to price an American option with FEM in the framework presented in part II, section 1.1. This will enable us to tackle the pricing of CBs in an easier way.

2.1 Modifying the PDEs

In this section, we will introduce changes of variables that are usual in order to formulate and solve our problem more accurately. We consider time to maturity $\tau = T - t$, and change to log-price $x = log(S)(S \neq 0)$ to remove the degeneracy of the operators in (II.1) and (II.4).

2.1.1 American option

The changes of variable described above hold almost everywhere (not when $S = 0$). Let us introduce

$$u(x, \tau) = V^{am}(e^x, T - \tau) \tag{III.1}$$

So, for $x \in \mathbb{R}$ or equivalently for $S \in (0, \infty)$, we have:

$$\frac{\partial V^{am}}{\partial S} = \frac{1}{S}\frac{\partial u}{\partial x} \quad , \quad \frac{\partial^2 V^{am}}{\partial S^2} = \frac{1}{S^2}(\frac{\partial^2 u}{\partial x^2} - \frac{\partial u}{\partial x}).$$

Hence, under the change of variables (III.1), the system (II.1) becomes:

$$\begin{cases} \dfrac{\partial u}{\partial \tau} - A^{BS-log}u + ru \geq 0 & \text{in } (0,T)\times\mathbb{R} \\ u(x,\tau) \geq \psi(x) & \\ (u-\psi)\cdot(\dfrac{\partial u}{\partial \tau} - A^{BS-log}u + ru) = 0 & \text{in } (0,T)\times\mathbb{R} \\ u(x,0) = \psi(x) & \end{cases} \tag{III.2}$$

where the operator A^{BS-log} is given by

$$A^{BS-log} = \frac{\sigma^2}{2}\frac{\partial^2}{\partial x^2} - (\frac{\sigma^2}{2} - r)\frac{\partial}{\partial x},$$

and $\psi(x) = g(e^x)$.

2.1.2 Convertible Bond with constant recovery

Let us denote by u the value of the CB once we have changed variables: $u(x,\tau) = V(S,t)$. Let us moreover write $f(x) := \phi_F(e^x)$. We define the constraints ψ_1 and ψ_2 as follows:

$$\psi_1(x,\tau) = \max(B_c(\tau), \kappa e^x) \tag{III.3}$$
$$\psi_2(x,\tau) = \max(B_p(\tau), \kappa e^x). \tag{III.4}$$

Let us write the change of variables for the coupon payment stream as well as the accrual interest. We define $\tau_{c_p} = T - t_{c_p}$, with $1 \leq c_p \leq N_c$. After ordering in increasing order the $(\tau_{c_p})_{1\leq c_p \leq N_c}$, we have:

$$AccI(0) = 0 \quad ; \quad AccI(\tau) = K \cdot \frac{\tau_{c_{i+1}} - \tau}{\tau_{c_{i+1}} - \tau_{c_i}} \quad \text{for} \quad \tau_{c_i} < \tau \leq \tau_{c_{i+1}}$$

The system (II.5)-(II.8) satisfied by u becomes

$$\begin{pmatrix} \frac{\partial u}{\partial \tau} - \mathcal{A}u - f = 0 \\ u \geq \psi_2 \\ u \leq \psi_1 \end{pmatrix} \vee \begin{pmatrix} \frac{\partial u}{\partial \tau} - \mathcal{A}u - f \geq 0 \\ u = \psi_2 \\ u \leq \psi_1 \end{pmatrix} \vee \begin{pmatrix} \frac{\partial u}{\partial \tau} - \mathcal{A}u - f \leq 0 \\ u \geq \psi_2 \\ u = \psi_1 \end{pmatrix} \tag{III.5}$$

where the operator $\mathcal{A}$ is defined as

$$\mathcal{A} = \frac{1}{2}\sigma^2 \frac{\partial^2}{\partial x^2} + (r + p\eta - q - \frac{1}{2}\sigma^2)\frac{\partial}{\partial x} - (r + p) \tag{III.6}$$

We have decided not to include the coupon stream $c(\tau)$ in the function f. However, at all coupon dates $(\tau_{c_p})_{1 \leq c_p \leq N_c}$ (by a no arbitrage argument) we have to force the solution u of the pricing problem to jump:

$$u(\tau_{c_p}, x) = u(\tau_{c_p}^-, x) + K \qquad (\forall c_p \in [|1 : N_c|] \ , \ x \in \mathbb{R})$$

2.2 Variational formulation on a truncated domain

2.2.1 American option

For computational reasons, we have to truncate the domain where we calculate the price. Indeed we choose x to belong to the set $\Omega_R := (-R, R)$. We then impose homogeneous Dirichlet boundary conditions: $u(\pm R, \tau) = 0, \ \forall \tau \in (0, T)$.

Furthermore, we define the Sobolev space $V := H_0^1(\Omega_R) = \{v, v' \in L^2(\Omega_R)\}$, where v' has to be understood in a weak sense.

The truncated problem to (III.2) reads:

Find $u^R \in L^2(0, T; V)$ such that

$$\frac{\partial u^R}{\partial \tau} - A^{BS-log} u^R + r u^R \geq 0 \qquad \text{in } (0;\mathrm{T}) \times \Omega_\mathrm{R} \tag{III.7}$$

$$u^R(x, \tau) \geq \psi(x) \tag{III.8}$$

$$(u^R - \psi) \cdot (\frac{\partial u^R}{\partial \tau} - A^{BS-log} u^R + r u^R) = 0 \qquad \text{in } (0;\mathrm{T}) \times \Omega_\mathrm{R} \tag{III.9}$$

$$u^R(x, 0) = \psi(x) \tag{III.10}$$

We give the variational formulation of (III.7)-(III.10). To this end, $v \in V$ be a variation function. We define the convex and closed cone (see (III.8))

$$K_\psi := \{v \in V \mid v \geq \psi \ , \text{ for a.e. } x \in \Omega_R\}.$$

Let u^R be the solution of (III.7)-(III.10) and let $v \in K_\psi$. From (III.7) and (III.8), it follows

$$(\frac{\partial u^R}{\partial \tau} + A^{BS-log} u^R)(v - \psi) \geq 0.$$

Integrating this inequality over Ω_R and using (III.9) (we set $\partial_\tau := \frac{\partial}{\partial_\tau}$):

$$\int_{-R}^{R} [(\partial_\tau u^R + A^{BS-log} u^R)(v - \psi)] dx \geq 0$$
$$\int_{-R}^{R} [(\partial_\tau u^R + A^{BS-log} u^R)(u^R - \psi)] dx = 0.$$

Subtracting the equality from the inequality above gives

$$\int_{-R}^{R} [(\partial_\tau u^R + A^{BS-log} u^R)(v - u^R)] dx \geq 0$$

Furthermore, we have (we set $\partial_x := \frac{\partial}{\partial_x}$ and $\partial_{xx} := \frac{\partial}{\partial_{xx}}$):

$$\int_{-R}^{R} A^{BS-log} u^R (v - u^R) dx = \int_{-R}^{R} [-\frac{\sigma^2}{2} \partial_{xx} u^R (v - u^R) + (\frac{\sigma^2}{2} - r) \partial_x u^R (v - u^R) + r u^R (v - u^R)] dx$$

Integration by parts gives

$$\int_{-R}^{R} \partial_{xx} u^R (v - u^R) dx = \underbrace{[\partial_x u^R (v - u^R)]_{-R}^{R}}_{=0 \ since \ v \in V} - \int_{-R}^{R} \partial_x u^R (v' - \partial_x u^R) dx,$$

and therefore we get

$$\int_{-R}^{R} A^{BS-log} u^R (v - u^R) dx = \underbrace{\int_{-R}^{R} [\ \frac{\sigma^2}{2} \partial_x u^R (v' - \partial_x u^R) + (\frac{\sigma^2}{2} - r) \partial_x u^R (v - u^R) + r u^R (v - u^R)\ ] dx}_{a^{BS-log}(u^R, v - u^R)}.$$

Denoting by $(\cdot, \cdot)$ the inner product of V the inequality becomes

$$(\partial_\tau u^R, v - u^R) + a^{BS-log}(u^R, v - u^R) \geq 0.$$

The variational problem is the following. Find $u^R \in L^2(0, T; V) \cap H^1(0, T; V')$, with $V = H_0^1(-R, R)$ and V' denotes the dual space of V, such that $u(\tau, .) \in K_\psi$ a.e. in $(0, T)$ and such that $\forall v \in K_\psi$:

$$(\partial_\tau u^R, v - u^R) + a^{BS-log}(u^R, v - u^R) \geq 0 \tag{III.11}$$

$$u^R(\cdot, 0) = \psi. \tag{III.12}$$

It is not obvious that there exists a unique solution to the problem (III.11) - (III.12). To prove existence and uniqueness, we refer to [Schw07].

The main steps are to show that the bilinear form a is coercive and continuous on $V \times V$. Furthermore, since $f \in L^2(\Omega_R)$ and we are working with a closed and convex set K_ψ, it can be shown that there exists a unique solution $u \in L^2(0, T; V) \cap C^0([0, T]; L^2(\Omega_R))$ to (III.11) - (III.12). Since it is unique, it has to be the solution to the original problem (III.7)-(III.10).

2.2.2 Convertible bond

We are going to truncate the problem for $x \in \mathbb{R}$ to $x \in (-R, R) := \Omega_R$ for computational reasons. We will impose homogeneous Dirichlet boundary conditions for $x = \pm R$. We will follow the same procedure as in part III, section 2.2.1.

Let $v \in L^2(0, T; V)$, where $V = H_0^1(\Omega_R)$. We define $K^R_{\psi_2,\psi_1}$ by:

$$K^R_{\psi_2,\psi_1} = \{v \in L^2(0, T; V) \mid \psi_1(x, \tau) \geq v(x, \tau) \geq \psi_2(x, \tau) \text{ , for a.e. } x \in \Omega_R \text{ , for a.e. } \tau \in [0, T]\}.$$

We choose u^R (approximation of u on $K^R_{\psi_2,\psi_1}$) and v to belong to the same space $K^R_{\psi_2,\psi_1}$.

Let us derive the variational formulation for the LCP (III.5) on the truncated domain $\Omega_R \times (0, T)$. From (III.5), we have

$$(\partial_\tau u^R - \mathcal{A}u^R - f)(v - u^R) \geq 0 \quad \forall\, v \in L^2(0, T; V)$$

Integrating on Ω_R yields

$$(\partial_\tau u^R, v - u^R) - (\mathcal{A}u^R, v - u^R) - (f, v - u^R) \geq 0 \quad \forall\, v \in L^2(0, T; V)$$

where $(g, h) = (g(\cdot, \tau), h(\cdot, \tau)) = \int_{-R}^{R} g(x, \tau)h(x, \tau)dx$.

Furthermore, using integration by parts, we get

$$
\begin{aligned}
(\mathcal{A}u^R, v-u^R) &= \int_{-R}^{R} \frac{1}{2}\sigma^2 \partial_{xx}u^R\ (v-u^R) + (r+p\eta-q-\frac{1}{2}\sigma^2)\partial_x u^R\ (v-u^R) \\
&\quad - (r+p)u^R\,(v-u^R)dx \\
&= \underbrace{[\frac{\sigma^2}{2}\partial_x u^R\ (v-u^R)]_{-R}^{R}}_{=0\ since\ u^R, v\in K_{\psi_2,\psi_1}} - \int_{-R}^{R} \frac{\sigma^2}{2}\partial_x u^R\ (\partial_x v - \partial_x u^R)dx \\
&\quad + \int_{-R}^{R} (r+p\eta-q-\frac{1}{2}\sigma^2)\partial_x u^R\ (v-u^R) - (r+p)u^R\ (v-u^R)dx.
\end{aligned}
$$

We can define the bilinear form a such that $\forall g, h \in K^R_{\psi_2,\psi_1}$

$$
a(g,h) = -(\mathcal{A}g, h) = \int_{-R}^{R} \frac{\sigma^2}{2}\partial_x g\ \partial_x h + (\frac{\sigma^2}{2} + q - r - p\eta)\partial_x g\, h + (r+p)g\, h dx \qquad \text{(III.13)}
$$

The weak formulation of our problem is now :

Find $u^R \in K^R_{\psi_2,\psi_1}$ such that $\partial_\tau u^R \in L^2(0,T;L^2(\Omega_R))$, and such that we have

$$
\begin{aligned}
&(\partial_\tau u^R, v-u^R) + a(u^R, v-u^R) \geq (f(\cdot,\tau), v-u^R) \quad \forall\ v \in K^R_{\psi_2,\psi_1} \\
&u^R(\cdot,0) = \psi_2(\cdot,0).
\end{aligned}
$$

To show that this weak formulation has a unique solution, we refer to [Schw07] and [BL78] (chapter 3, section 2.18). The above problem is basically dealt with in the same way as the problem (part III, section 2.2.1) on American options.

3 Two stochastic factors: stock price & interest rate

In this section, we will do the usual changes of variables for the two-dimension case presented in part II, sections 2.1 , i.e. change to log-price: $x = \log(S)$ and time to maturity $\tau = T - t$.

We will work with the LCP

$$
\begin{pmatrix} \mathcal{M}_2 V - \phi_F(S) = 0 \\ V \geq g_2 \\ V \leq g_1 \end{pmatrix} \vee \begin{pmatrix} \mathcal{M}_2 V - \phi_F(S) \geq 0 \\ V = g_2 \\ V \leq g_1 \end{pmatrix} \vee \begin{pmatrix} \mathcal{M}_2 V - \phi_F(S) \leq 0 \\ V \geq g_2 \\ V = g_1 \end{pmatrix} \qquad \text{(III.14)}
$$

where $\mathcal{M}_2$ has been defined in (II.10).

We have taken out the function $c(t)$. As for the one dimensional case, we will force the solution to jump accordingly at coupon dates.

3.1 Change of variables

We will work with the system (III.14). Let us denote by u, the value of the CB once we have changed variables: $u(x,\tau) = V(S,t)$. Let us moreover define $f(x) = \phi_F(e^x)$. ψ_1 and ψ_2 are the constraints defined by $\psi_1(x,\tau) := \max(B_c(\tau), \kappa e^x)$ and $\psi_2(x,r,\tau) := \max(B_p(\tau), B(\tau), \kappa e^x)$. In the case the bond B is not callable, the expression $B(\tau) = \mathbb{E}_{\mathbb{Q}}[\ e^{-\int_{T-\tau}^{T} r(s)ds}\ ]$ holds, where $\mathbb{Q}$ is a spot martingale measure (see [Scho07]).

Once we have changed variables, we get ($\partial_\tau u = \frac{\partial u}{\partial \tau}$):

$$\begin{pmatrix} \partial_\tau u - \mathcal{A}_2(u) - f(x,\tau) = 0 \\ u \geq \psi_2 \\ u \leq \psi_1 \end{pmatrix} \vee \begin{pmatrix} \partial_\tau u - \mathcal{A}_2(u) - f(x,\tau) \geq 0 \\ u = \psi_2 \\ u \leq \psi_1 \end{pmatrix} \vee \begin{pmatrix} \partial_\tau u - \mathcal{A}_2(u) - f(x,\tau) \leq 0 \\ u \geq \psi_2 \\ u = \psi_1 \end{pmatrix} \tag{III.15}$$

where we define $\mathcal{A}_2$ as the operator:

$$\begin{aligned} \mathcal{A}_2 &:= \frac{\sigma^2}{2}\frac{\partial^2}{\partial x^2} + \frac{\tilde{\sigma}^2}{2}\frac{\partial^2}{\partial r^2} + \rho_{(r,S)}\sigma\tilde{\sigma}\frac{\partial^2}{\partial x \partial r} + (r + p\eta - q - \frac{1}{2}\sigma^2)\frac{\partial}{\partial x} + \tilde{\mu}\frac{\partial}{\partial r} - (r+p) \\ &= \mathcal{A} + \frac{\tilde{\sigma}^2}{2}\frac{\partial^2}{\partial r^2} + \rho_{(r,S)}\sigma\tilde{\sigma}\frac{\partial^2}{\partial x \partial r} + \tilde{\mu}\frac{\partial}{\partial r} \end{aligned}$$

3.2 Variational formulation on a truncated domain

We follow the same procedure as in section 2.2.2. except we are dealing with a truncated domain in $\mathbb{R}^2$.

We are going to truncate the problem for $x \in \mathbb{R}$ to $x \in (-R_1, R_1) := \Omega_x$ and $r \in [R_2, R_3] := \Omega_r$ where R_2 can be for example -10% or 0% and R_3 may be 25% or 50%. We will impose homogeneous Dirichlet boundary conditions on $\partial\Omega_x$ and on $\partial\Omega_r$.

Let $\Omega_R := \Omega_x \times \Omega_r$ ($R := (R_1, R_2, R_3)$) and $V = H_0^1(\Omega_R)$.

Let $v \in L^2(0,T;V)$ and define $K^R_{\psi_2,\psi 1}$ as:

$$K^R_{\psi_2,\psi_1} := \{v \in L^2(0,T;V) \mid \psi_1(x,\tau) \geq v(x,r,\tau) \geq \psi_2(x,r,\tau)\ \text{, for a.e. } x \in \Omega_x, r \in \Omega_r, \tau \in [0,T]\}.$$

Proceeding analogously to section 2.2.2, the parabolic variational inequality to (III.15) reads:

Find $u^R \in L^2(0,T;V)$ such that $\partial_\tau u^R \in L^2(0,T;L^2(\Omega_R))$ and $u^R \in K^R_{\psi_2,\psi_1}$ such that

$$(\partial_\tau u^R, v - u^R) + a_2(u^R, v - u^R) \geq (f, v - u^R) \quad \forall\, v \in K^R_{\psi_2,\psi_1} \tag{III.16}$$

$$u^R(\cdot,\cdot,0) = \psi_2(\cdot,0). \tag{III.17}$$

where the bilinear form $a_2(\cdot,\cdot) : K^R_{\psi_2,\psi_1} \times K^R_{\psi_2,\psi_1} \to \mathbb{R}$ is given by (where $' := \partial_r$)

$$\left.\begin{aligned} a_2(g,h) = \quad & -(\mathcal{A}_2 g, h) = \tfrac{\sigma^2}{2}(\partial_x g, \partial_x h) + \tfrac{1}{2}(\tilde{\sigma}\partial_r g, \tilde{\sigma}\partial_r h) + (\tilde{\sigma}\tilde{\sigma}'\partial_r g, h) \\ & +\rho_{(r,S)}\sigma(\tilde{\sigma}\partial_r g, \partial_x h) - (r\partial_x g, h) + (\tfrac{\sigma^2}{2} + q - p\eta)(\partial_x g, h) - (\tilde{\mu}\partial_r g, h) \\ & +(rg, h) + p(g,h), \end{aligned}\right\} \tag{III.18}$$

and $(\cdot,\cdot)$ denotes the inner product in $L^2(\Omega_R)$.

4 Implementation for pricing convertibles

In this section, we will implement the pricing of convertibles. We will describe the FEM and framework that is used to approximate the price of a CB.

4.1 One stochastic factor: stock price

4.1.1 Discretization in log-price

We will first proceed by discretizing in space (log-price). For $N \geq 1$, we define the sequence $-R = x_0 < x_1 < ... < x_N < x_{N+1} = R$ and we use the Finite Element Space

$$V_N = span\{b_i(x) : i = 1, ..., N\} \subset V = H^1_0(\Omega_R)$$

where the functions b_i are defined by

$$b_i(x) = \begin{cases} \frac{1}{h_i}(x - x_{i-1}) & x \in K_i := (x_{i-1}, x_i), \\ \frac{1}{h_{i+1}}(x_{i+1} - x) & x \in K_{i+1}, \\ 0 & else \end{cases}$$

We will drop the superscript from u^R and simply write u to ease notation. We can approximate the functions u and v by u_N and v_N which are functions in the space V_N.

We can write

$$u_N(x,\tau) = \sum_{i=1}^{N} u_i^N(\tau) b_i(x)$$
$$v_N(x,\tau) = \sum_{i=1}^{N} v_i^N(\tau) b_i(x)$$

where $\underline{u}^N(\tau) := \{u_i^N(\tau)\}_{i=1}^N$ is in $\underline{K}_{\psi_2,\psi_1}(\tau)$ (defined just below).

For discretization, we approximate $K^R_{\psi_2,\psi_1}(\tau)$ by

$$\underline{K}_{\psi_2,\psi_1}(\tau) = \{U_N = \sum_{i=1}^{N} U_i^N(\tau) b_i(x) \ , \ \psi_1(x_i,\tau) \geq U_i^N(\tau) \geq \psi_2(x_i,\tau) \ , \ i = 1, ..., N\}$$

Note that $\underline{K}_{\psi_2,\psi_1} \nsubseteq K^R_{\psi_2,\psi_1}$ since the piecewise linear continuous function consisting of the points $\psi_1(x_i,\tau)$ is above ψ_1 because ψ_1 is convex.

Now, we can discretize our variational inequality in space. We are looking for $\underline{u}^N(\tau) \in \underline{K}_{\psi_2,\psi_1}(\tau)$ for a.e. τ, such that for all $\underline{v}^N(\tau) \in \underline{K}_{\psi_2,\psi_1}(\tau)$

$$(\frac{\partial(\underline{u}^N)^\top}{\partial\tau} \ \underline{b}(x), (\underline{v}^N - \underline{u}^N)\top \ \underline{b}(x)) + a((\underline{u}^N)\top \ \underline{b}(x), (\underline{v}^N - \underline{u}^N)\top \ \underline{b}(x)) \geq (f(.,\tau), (\underline{v}^N - \underline{u}^N)\top \ \underline{b}(x))$$

where $(\underline{u}^N)^\top$ denotes the transpose of vector $\underline{u}^N$ and $\underline{b}(x)$ is the column vector $\{b_i(x)\}_{i=1}^N$.

Let us write some terms in a different way:

- term ($\frac{\partial(\underline{u}^N)^\top}{\partial\tau} \ \underline{b}(x), (\underline{v}^N - \underline{u}^N)^\top \underline{b}(x)$):

$$\begin{aligned}(\frac{\partial(\underline{u}^N)^\top}{\partial\tau} \underline{b}(x), (\underline{v}^N - \underline{u}^N)^\top \underline{b}(x)) &= \frac{\partial(\underline{u}^N)^\top}{\partial\tau} (\underline{b}, \underline{b}^\top (\underline{v}^N - \underline{u}^N)) \\ &= (\dot{\underline{u}}^N)^\top (\underline{b}, \underline{b}^\top) (\underline{v}^N - \underline{u}^N) \\ &= (\underline{v}^N - \underline{u}^N)^\top \underbrace{(\underline{b}, \underline{b}^\top)^\top}_{=(\underline{b},\underline{b}^\top)} \dot{\underline{u}}^N\end{aligned}$$

We find that $\quad (\frac{\partial(\underline{u}^N)^\top}{\partial\tau}\ \underline{b}(x), (\underline{v}^N - \underline{u}^N)^\top\, \underline{b}) = (\underline{v}^N - \underline{u}^N)^\top\, (\underline{b}, \underline{b}^\top)\, \dot{\underline{u}}^N$

- term $a((\underline{u}^N)^\top\, \underline{b}(x), (\underline{v}^N - \underline{u}^N)^\top\, \underline{b}(x))$. Given that a is bilinear, we have:

$$\begin{aligned} a((\underline{u}^N)^\top\, \underline{b}, (\underline{v}^N - \underline{u}^N)^\top\, \underline{b}) &= (\underline{u}^N)^\top\, a(\underline{b}\ ,\ \underline{b}^\top\, (\underline{v}^N - \underline{u}^N)) \\ &= (\underline{u}^N)^\top \cdot a(\underline{b}\ ,\ \underline{b}^\top)\, (\underline{v}^N - \underline{u}^N) \\ &= (\underline{v}^N - \underline{u}^N)^\top\ \underbrace{a(\underline{b}\ ,\ \underline{b}^\top)^\top}_{not\ symmetric}\ \underline{u}^N \end{aligned}$$

So, we can write $\quad a((\underline{u}^N)^\top\, \underline{b}, (\underline{v}^N - \underline{u}^N)^\top\, \underline{b}) = (\underline{v}^N - \underline{u}^N)^\top\, a(\underline{b}\ ,\ \underline{b}^\top)^\top\, \underline{u}^N$

- term $(f(.,\tau), (\underline{v}^N - \underline{u}^N)^\top\, \underline{b}(x))$:

$$\begin{aligned} (f(.,\tau), (\underline{v}^N - \underline{u}^N)^\top\, \underline{b}) &= (f, \underline{b}^\top\, (\underline{v}^N - \underline{u}^N)) \\ &= (f, \underline{b}^\top)\, (\underline{v}^N - \underline{u}^N) \\ &= (\underline{v}^N - \underline{u}^N)^\top\, (\underline{b}, f) \end{aligned}$$

Let us denote by $\mathbf{M}$ the *Mass matrix* , $\mathbf{A}$ the *Stiffness matrix* corresponding to the bilinear form a, and $\underline{f}$ the *load vector* :

$$\begin{aligned} &\mathbf{M}_{i,j} = (b_i, b_j),\ \forall i = 1, ..., N,\ \ j = 1, ..., N \\ &\mathbf{A}_{i,j} = a(b_j, b_i),\ \forall i = 1, ..., N,\ \ j = 1, ..., N \\ &\underline{f}_i\ \ = (b_i, f),\ \forall i = 1, ..., N \end{aligned}$$

Our problem can now be formulated as:

Find $\underline{u}^N \in \underline{K}_{\psi_2,\psi_1}$, such that for all $\underline{v}^N \in \underline{K}_{\psi_2,\psi_1}$

$$(\underline{v}^N - \underline{u}^N)^\top\, [\mathbf{M}\dot{\underline{u}}^N + \mathbf{A}\underline{u}^N - \underline{f}] \geq 0 \qquad \text{for a.e. } \tau \in (0, T) \tag{III.19}$$

4.1.2 Discretization in time

Now, we will discretize in time using a θ-scheme ($0 \leq \theta \leq 1$) for time stepping. Let in $(0, T)$ be a sequence $\{\tau_m\}_{m=0}^{M-1}$ ($M > 1$) of (not necessarily equally sized) time steps be given, and set $\tau_m = \sum_{i=0}^{m} k_i$ so that $\tau_M = T$ ($k_m > 0$ for all m). The discretization of the above inequality in the time and price variable reads:

Find $\{\underline{u}_N^{m+1}\}_{m=0}^{M-1} \in \underline{\mathcal{K}}_{\psi_2,\psi_1}$ such that (we drop the subscript N)

$$(\underline{v}^{m+1} - \underline{u}^{m+1})^\top \left[k_m^{-1}\mathbf{M}(\underline{u}^{m+1} - \underline{u}^m) + \mathbf{A}(\theta\underline{u}^{m+1} + (1-\theta)\underline{u}^m) - (\theta\underline{f}^{m+1} + (1-\theta)\underline{f}^m)\right] \geq 0$$
$$(\psi_1)_i^{m+1} \geq u_i^{m+1} \geq (\psi_2)_i^{m+1}$$
$$u_i^0 = (\psi_2)_i^0$$

where

$$\begin{aligned}
\underline{\mathcal{K}}_{\psi_2,\psi_1} &= \{w \in \mathbb{R}^N \times \mathbb{R}^M, \underline{\psi_2}^m \leq \underline{w}^m \leq \underline{\psi_1}^m, \text{ for } m < M\} \\
f^{m+\theta}(x) &= \theta f(x, t^{m+1}) + (1-\theta) f(x, t^m) \\
u_N^{m+\theta}(x) &= \theta u_N^{m+1}(x) + (1-\theta) u_N^m(x) \\
(\underline{\psi}_p)_i^m &= \psi_p(x_i, t^m) \qquad (\ 1 \leq p \leq 2)
\end{aligned}$$

We define

$$\underline{b}^m := k_m(\theta\underline{f}^{m+1} + (1-\theta)\underline{f}^m) + (\mathbf{M} - k_m(1-\theta)\mathbf{A})\underline{u}^m$$

And we get

$$(\underline{v}^{m+1} - \underline{u}^{m+1})^\top \cdot [\mathbf{M} + k_m\mathbf{A}\theta]\underline{u}^{m+1} \geq (\underline{v}^{m+1} - \underline{u}^{m+1})^\top \underline{b}^m$$
$$(\psi_1)_i^{m+1} \geq u_i^{m+1} \geq (\psi_2)_i^{m+1}$$
$$u_i^0 = (\psi_2)_i^0$$

These LCP are equivalent to

$$(\underline{\psi_1})^{m+1} \geq \underline{u}^{m+1} \geq (\underline{\psi_2})^{m+1} \tag{III.20}$$
$$(\underline{u}^{m+1} - (\underline{\psi_2})^{m+1})^\top [(\mathbf{M} + k_m\mathbf{A}\theta)\underline{u}^{m+1} - \underline{b}^m]((\underline{\psi_1})^{m+1} - \underline{u}^{m+1}) = \underline{0} \tag{III.21}$$
$$u_i^0 = (\psi_2)_i^0 \tag{III.22}$$

where $(\underline{\psi_k})^{m+1}$ is the column vector $\{(\psi_k)_i^{m+1}\}_{i=1}^M$ for $k = 1, 2$.

4.1.3 Calculation of matrices

Let us have a closer look at the matrices $\mathbf{M}$ and $\mathbf{A}$. To simplify calculations (see Appendix B for calculations), we will consider a constant price-space step (h_i is constant equal to $h = \frac{2R}{N+1}$).

First of all, one can observe that the function b_k is non zero only on (x_{k-1}, x_{k+1}) (in particular, we have for instance $(b_i, b_j) = 0,\ \forall |i-j| > 1$). Therefore, we have

$$\mathbf{M}_{i,j} = \mathbf{A}_{i,j} = 0, \forall\ i = 1, ..., N,\ \ j \notin \{i-1, i, i+1\}$$

In order to calculate the matrix $\mathbf{A}$, we will define the matrices $\mathbf{S}$ and $\mathbf{C}$

$$\mathbf{S}_{i,j} = \int_{-R}^{R} b_j'(x) b_i'(x) dx \quad ; \quad \mathbf{C}_{i,j} = \int_{-R}^{R} b_j'(x) b_i(x) dx. \tag{III.23}$$

Then $\mathbf{A}$ can be written as

$$\mathbf{A}_{i,j} = \frac{\sigma^2}{2}\mathbf{S}_{i,j} + (\frac{\sigma^2}{2} + q - r - p\eta)\mathbf{C}_{i,j} + (r+p)\mathbf{M}_{i,j}$$

we find (see Appendix B)

$$\mathbf{M}_{i,i} = \frac{2h}{3} \quad , \quad \mathbf{M}_{i,i+1} = \mathbf{M}_{i,i-1} = \frac{h}{6}$$

$$\begin{aligned}
\mathbf{A}_{i,i} &= \frac{\sigma^2}{h} + (r+p)\frac{2h}{3} \\
\mathbf{A}_{i,i-1} &= \frac{-\sigma^2}{2h} - \frac{1}{2}(\frac{\sigma^2}{2} + q - r - p\eta) + (r+p)\frac{h}{6} \\
\mathbf{A}_{i,i+1} &= \frac{-\sigma^2}{2h} + \frac{1}{2}(\frac{\sigma^2}{2} + q - r - p\eta) + (r+p)\frac{h}{6}
\end{aligned}$$

To calculate the load vector $\underline{f}^m$, we recall that we have $\underline{f}_i^m = (b_i, f^m)$ and $f(x,\tau) = p\ \max(\kappa \exp(x)(1-\eta), R_{rec} \cdot F)$.

$$\begin{aligned}
(f, b_i) = \int_{x_{i-1}}^{x_{i+1}} f(x,\tau) b_i(x) dx &= \frac{1}{h}\int_{x_{i-1}}^{x_i} (p \max(\kappa e^x (1-\eta), R_{rec} \cdot F))(x - x_{i-1}) dx \\
&+ \frac{1}{h}\int_{x_i}^{x_{i+1}} (p \max(\kappa e^x (1-\eta), R_{rec} \cdot F))(x_{i+1} - x) dx]
\end{aligned}$$

The value of this integral depends on the value of $\max(\kappa e^x(1-\eta), R_{rec} \cdot F)$. Check Appendix B for the calculation. We find:

$$\begin{cases} \text{If } \ \log(\frac{R_{rec}F}{\kappa(1-\eta)}) \leq x_0 \\ \qquad \text{then} \\ (f, b_i) = \frac{p\kappa(1-\eta)}{h} \ e^{x_i}(e^{-h} - 2 + e^h) \end{cases}$$

$$\begin{cases} \text{If } \ x_{N+1} \leq \log(\frac{R_{rec}F}{\kappa(1-\eta)}) \\ \qquad \text{then} \\ (f, b_i) = hpR_{rec}F \end{cases}$$

$$\begin{cases} \text{If } \ x_{i-1} < \log(\frac{R_{rec}F}{\kappa(1-\eta)}) < x_i \\ \qquad \text{then} \\ (f, b_i) = \frac{1}{2h}pR_{rec}F(\log(\frac{R_{rec}F}{\kappa(1-\eta)}) - x_{i-1})^2 + \\ \qquad\qquad \frac{1}{h}p\kappa(1-\eta)[\frac{R_{rec}F}{\kappa(1-\eta)}(-\log(\frac{R_{rec}F}{\kappa(1-\eta)}) + 1 + x_i - h) + e^{x_i}(-2 + e^h)] \end{cases}$$

$$\begin{cases} \text{If } \ x_i < \log(\frac{R_{rec}F}{\kappa(1-\eta)}) < x_{i+1} \\ \qquad \text{then} \\ (f, b_i) = hpR_{rec}F - \frac{1}{2h}pR_{rec}F[x_{i+1} - \log(\frac{R_{rec}F}{\kappa(1-\eta)})]^2 \\ \qquad\qquad + \frac{1}{h}p\kappa(1-\eta)[\frac{R_{rec}F}{\kappa(1-\eta)}(\log(\frac{R_{rec}F}{\kappa(1-\eta)}) - x_{i+1}) + e^{x_{i+1}} - \frac{R_{rec}F}{\kappa(1-\eta)}] \end{cases}$$

We can see from the above calculation that if we choose a value of $R < \log(\frac{R_{rec}F}{\kappa(1-\eta)})$, then we do not consider scenarios where conversion is optimal at default. Indeed $R_{rec}F > \kappa(1-\eta)e^x, \ \forall x \in (-R, R)$. We will therefore consider that $R > \log(\frac{R_{rec}F}{\kappa(1-\eta)})$ to allow a broad range of scenarios.

4.2 Two stochastic factors: stock price & interest rate

We discretize in $\Omega_R = (-R_1, R_1) \times (R_2, R_3)$ by a tensor product grid (with constant mesh): let $h_x = \frac{2R_1}{N_x+1}$ and $x_i = -R_1 + ih_x$, $i = 0, ..., N_x + 1$ and let $h_y = \frac{R_3 - R_2}{N_y+1}$ and $y_j = R_2 + jh_y$, $j = 0, ..., N_y + 1$.

Let us denote by $b_i(x) = \max(1-\frac{|x-x_i|}{h_x}, 0)$, $\bar{b}_j(y) = \max(1-\frac{|y-y_j|}{h_y}, 0)$. Then the Finite Elements space $V_{N_x,N_y} \subset V$ is given by $V_{N_x,N_y} = span\{b_i(x)\bar{b}_j(y), 1 \le i \le N_x, 1 \le j \le N_y\}$.
Note that $V_{N_x,N_y} = V_{N_x} \otimes V_{N_y}$ is a tensor product of univariate FE spaces.

For localization and discretization, we will proceed as in section 4.1.2. .

We can write

$$u(x,y,\tau) = \sum_{i=1}^{N_x}\sum_{j=1}^{N_y} u_{ij}(\tau) b_i(x)\bar{b}_j(y)$$
$$v(x,y,\tau) = \sum_{i=1}^{N_x}\sum_{j=1}^{N_y} v_{ij}(\tau) b_i(x)\bar{b}_j(y)$$

The column vector $\underline{u}(\tau) := \{u_{ij}(\tau)\}_{i,j=1}^{N_x,N_y} = (u_{11}(\tau), u_{12}(\tau), ..., u_{1N_y}(\tau), u_{21}(\tau), ..., u_{N_xN_y}(\tau))^T$ is in $\underline{K}_{\psi_2,\psi_1}$ (defined just below).

For discretization, we approximate $K^R_{\psi_2,\psi_1}$ by

$$\underline{K}_{\psi_2,\psi_1} = \{U = \sum_{i=1}^{N_x}\sum_{j=1}^{N_y} U_{ij}(\tau) b_i(x)\bar{b}_j(y)\ ,\ \psi_1(x_i,\tau) \ge U_{ij}(\tau) \ge \psi_2(x_i,y_j,\tau)\ ,$$
$$i = 1,...,N_x,\ j = 1,...,N_y\ \text{, a.e. in } \tau\}$$

Let us notice that $\underline{K}_{\psi_2,\psi_1} \subsetneq K^{R,r_{max}}_{\psi_2,\psi_1}$ since the piecewise linear continuous function consisting of the points $\psi_1(x_i,\tau)$ is above ψ_1 (because ψ_1 is convex).

Since, we have already discretized the one dimensional problem in section 4.1, we will just give the major steps and explain where there are differences.

Let us discretize our variational inequality in space. We are looking for $\underline{u} \in \underline{K}_{\psi_2,\psi_1}$, such that for all $\underline{v} \in \underline{K}_{\psi_2,\psi_1}$

$$(\frac{\partial(\underline{u})^\top}{\partial\tau}\underline{b}, (\underline{v}-\underline{u})^\top\underline{b})_{L^2(\Omega_R)} + a((\underline{u})^\top\underline{b}, (\underline{v}-\underline{u})^\top\underline{b}) \ge (f(.,\tau), (\underline{v}-\underline{u})^\top\underline{b})_{L^2(\Omega_R)}$$

where $(\underline{u})^\top$ (respectively $(\underline{u})^\top$) denotes the transpose of vector $\underline{u}$ (respectively $\underline{v}$) and $\underline{b}$ is the column vector $(b_1\bar{b}_1, b_1\bar{b}_2, ...b_1\bar{b}_{N_y}\ , b_2\bar{b}_1, ..., b_2\bar{b}_{N_y}\ , ..., b_{N_x}\bar{b}_{N_y})^\top$.

Discretizing in time leads to the following system

$$
\begin{aligned}
(\underline{\psi_1})^{m+1} \geq \underline{u}^{m+1} \geq (\underline{\psi_2})^{m+1} \\
(\underline{u}^{m+1} - (\underline{\psi_2})^{m+1})^\top [(\mathbf{M} + k_m \mathbf{A}\theta)\underline{u}^{m+1} - \underline{b}^m]((\underline{\psi_1})^{m+1} - \underline{u}^{m+1}) = \underline{0} \\
u_i^0 = (\psi_2)_i^0
\end{aligned}
$$

where k_m is defined in section 4.1.2.. The matrices $\mathbf{A}$ and $\mathbf{M}$ are defined in the same way as in the section 4.1.2., i.e. we have

$$
\begin{aligned}
\mathbf{M} &= [(\underline{b}, \underline{b}^\top)_{L^2(\Omega_R)}]^\top = \underbrace{(\underline{b}, \underline{b}^\top)}_{\in\ \mathbb{R}^{N_x N_y \times N_x N_y}} \qquad (\mathbf{M} \text{ is symmetric}) \\
\mathbf{A} &= \underbrace{a(\underline{b}, \underline{b}^\top)^\top}_{\in\ \mathbb{R}^{N_x N_y \times N_x N_y}}
\end{aligned}
$$

The matrices $\mathbf{M}$ and $\mathbf{A}$ can be built from tensor products of univariate matrices, e.g. the mass matrix $\mathbf{M} = (\mathbf{M}_{(i,j),(i',j')})$ is of the tensor product form

$$
\begin{aligned}
\mathbf{M}_{(i,j),(i',j')} &= (b_i \bar{b}_j, b_{i'} \bar{b}_{j'})_{L^2(\Omega_R)} \\
&= \int_{\Omega_R} b_i(x) b_{i'}(x) \bar{b}_j(y) \bar{b}_{j'}(y) dy dx \\
&= (b_i, b_{i'})_{L^2(-R_1,R_1)} (b_j, b_{j'})_{L^2(R_2;R_3)} = (\mathbf{M}_x)_{ii'} (\mathbf{M}_y)_{jj'} \\
&= (\mathbf{M}_x \otimes \mathbf{M}_y)_{(i,j),(i',j')}
\end{aligned}
$$

where the matrices $\mathbf{M}_x$ and $\mathbf{M}_y$ are the univariate mass matrices calculated for the one-dimensional problem in section 4.1.2.

This means that we just need to have calculated $\mathbf{M}_x$ and $\mathbf{M}_y$ in order to find $\mathbf{M}$.

We will find a similar result for the matrix $\mathbf{A}$. Let us define the matrices we will use to decompose the matrix $\mathbf{A}$:

$$
\begin{aligned}
\{\mathbf{M}_x\}_{i,i'} &= (b_i, b_{i'})_{L^2(-R,R)} && \text{(III.24)} \\
\{\mathbf{M}_y\}_{j,j'} &= (\bar{b}_j, \bar{b}_{j'})_{L^2(R_2,R_3)} && \text{(III.25)} \\
\{\mathbf{M}_{y,1}\}_{j,j'} &= (y\bar{b}_j, \bar{b}_{j'})_{L^2(R_2,R_3)} && \text{(III.26)} \\
\{\mathbf{C}_x\}_{i,i'} &= (b'_{i'}, b_i)_{L^2(-R,R)} && \text{(III.27)} \\
\{\mathbf{C}_y\}_{j,j'} &= (\bar{b}'_{j'}, \bar{b}_j)_{L^2(R_2,R_3)} && \text{(III.28)} \\
\{\mathbf{C}_{y,1}\}_{j,j'} &= (y\bar{b}'_{j'}, \bar{b}_j)_{L^2(R_2,R_3)} && \text{(III.29)} \\
\{\mathbf{S}_x\}_{i,i'} &= (b'_i, b'_{i'})_{L^2(-R,R)} && \text{(III.30)} \\
\{\mathbf{S}_y\}_{j,j'} &= (\bar{b}'_j, \bar{b}'_{j'})_{L^2(R_2,R_3)} && \text{(III.31)}
\end{aligned}
$$

Assume that the interest rate r follows a Vasicek process, i.e. $\tilde{\mu} = a(b-r)$, $\tilde{\sigma} = \sigma_r$, $a, b, \sigma_r \in \mathbb{R}$ in (II.9).

Then the bilinear form a, we obtain the tensor product matrix representation

$$\begin{aligned}
&\mathbf{M} = \mathbf{M}_x \otimes \mathbf{M}_y \\
&\mathbf{A} = \frac{\sigma^2}{2}\mathbf{S}_x \otimes \mathbf{M}_y - \mathbf{C}_x \otimes \mathbf{M}_{y,1} + (\frac{\sigma^2}{2} + q - p\eta)\mathbf{C}_x \otimes \mathbf{M}_y + \rho_{(r,S)}\sigma\sigma_r \underbrace{\mathbf{C}_x^\top}_{-\mathbf{C}_x} \otimes \mathbf{C}_y + \frac{\sigma_r^2}{2}\mathbf{M}_x \otimes \mathbf{S}_y \\
&\qquad - ab\mathbf{M}_x \otimes \mathbf{C}_y + a\mathbf{M}_x \otimes \mathbf{C}_{y,1} + p\mathbf{M}_x \otimes \mathbf{M}_y + \mathbf{M}_x \otimes \mathbf{M}_{y,1}
\end{aligned}$$

In order to reduce the number of tensor products to calculate, $\mathbf{A}$ can be written as

$$\mathbf{A} = \mathbf{S}_x \otimes \mathbf{A}_{1,y} + \mathbf{C}_x \otimes \mathbf{A}_{2,y} + \mathbf{M}_x \otimes \mathbf{A}_{3,y}$$

with

$$\begin{aligned}
\mathbf{A}_{1,y} &= \frac{\sigma^2}{2}\mathbf{M}_y \\
\mathbf{A}_{2,y} &= -\mathbf{M}_{y,1} + (\frac{\sigma^2}{2} + q - p\eta)\mathbf{M}_y - \rho_{(r,S)}\sigma\sigma_r\mathbf{C}_y \\
\mathbf{A}_{3,y} &= \frac{\sigma_r^2}{2}\mathbf{S}_y - ab\mathbf{C}_y + a\mathbf{C}_{y,1} + p\mathbf{M}_y + \mathbf{M}_{y,1}.
\end{aligned}$$

The components of the matrices above are calculated in the Appendix A and Appendix C.

In order to compute the load vector $\mathbf{F}$, we will denote by $\mathbf{F}_x$ the load vector calculated in the section 4.1.3 and by $\mathbf{F}_y$ the vector

$$\mathbf{F}_y = h_y\,(1, ..., 1)^\top$$

and we have $\mathbf{F} = \mathbf{F}_x \otimes \mathbf{F}_y$.

Part IV
Results

In this part, we will present the results of the implementation using the FEM discussed in part III for the pricing of a CB. We would like to give some intuition of how the price of a CB can be impacted by the different parameters entering in the valuation.

To give us an idea of the accuracy of our method, we have decided to compare our results to a commercial pricing software. It is important to know that it was not possible to improve the accuracy (refine the grid) of the pricing for the commercial tool used. In some rare and extreme cases, this can explain the differences we find.

Furthermore, we would like to get a good idea of the convergence rate of our pricer by testing the speed of convergence numerically (see section 1.2).

Finally, we compute the exercise boundaries for the holder and issuer of a CB, which give us numerically and precisely the continuation region and exercise region (see section 1.3). We also study the impact on the exercise boundaries of a change in a parameter (e.g. the volatility or the interest rate).

1 One stochastic factor

1.1 Accuracy and behavior of the CB price

In this section, we intend to check the accuracy of our pricing algorithm. We will run many tests in order to get a good idea of where our pricing is accurate and where it could be improved.

Let us first show the impact of two parameters in our model: the drop in stock price at default η and the recovery rate when default occurs R_{rec}.

Here is a list of the parameters that we use when testing. We will then test our CB values by changing just one at a time and keeping the other ones fixed to these values.

Parameters:

Parameter	Value	Explanation
T	10	Maturity of CB (years)
F	100	Face Value of Bond
K	0	Absolute coupon value (constant)
$(t_{c_j})_{1\leq j\leq N_c}$	$\emptyset$	Coupons dates in years.
σ	0.2	Volatility (here 20%)
η	1	Proportional loss in stock value at default
R_{rec}	0	Recovery rate of face value at default
κ	1	conversion ratio (here: one CB for one share)
B_c	140	Clean price at which issuer calls the CB
B_p	20	Clean price at which holder puts the CB
r	0.05	Short interest rate (here 5%)
p	0.03	probability of default/year (here 3%)
q	0.06	Continuous dividend yield (here 6%)

Table 1: Parameters used in testing zero-coupons CBs

For computation of prices, we usually chose the following grid parameters:

Parameter	Value	Explanation
$N+1$	1024	$N+2$ is the number of grid points for discretization in log-price
R	6	$x \in [-R, R] \Rightarrow S \in [e^{-R}, e^{R}]$
M	128	$M+1$ is the number of Grid points for time-discretization
θ	0.5	Crank-Nicolson scheme
$jmax$	500	Max number of iterations for the PSOR
tol	10^{-10}	tolerance in the PSOR (Euclidean norm)

Table 2: Parameters for computation of CBs

Figure 1 shows the influence of the parameter η (fractional loss in stock price at default) in the valuation of a CB. We can clearly see that the more the stock price loses value at default the more the price-curve decreases. We have chosen to plot 3 cases: the stock price loses no value ($\eta = 0$), the stock price loses half of its value ($\eta = 0.5$) and if the stock price drops to 0 ($\eta = 1$).

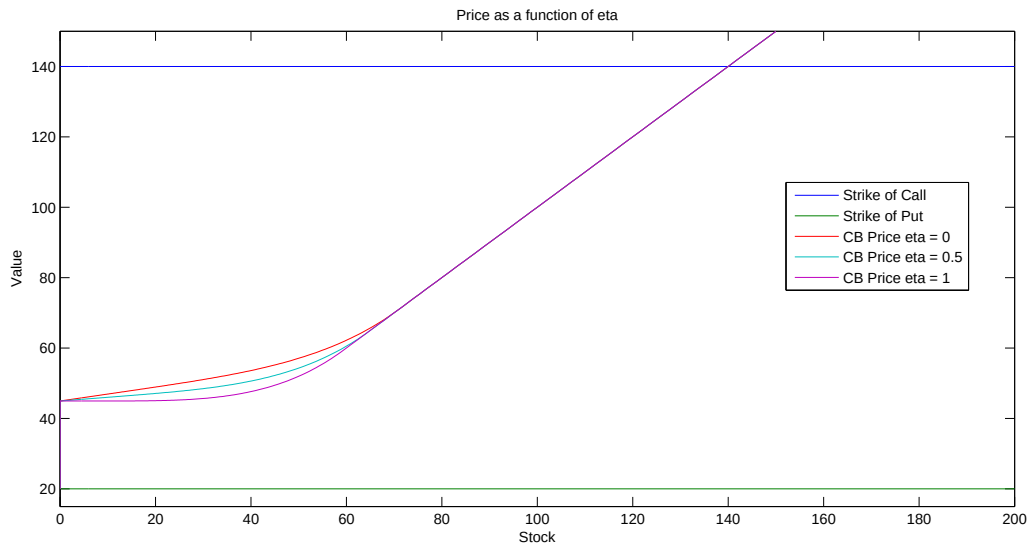

Figure 1: Price of a CB when the jump η in stock price at default varies

It seems intuitive that the more the stock price can fall at default, the less valuable the CB is. This is what we observe on Figure 1.

Moreover, it is normal that the price of a CB grows linearly (for $\kappa S \geq B_c$) as a function of the stock price with exactly the slope κ (conversion ratio). This can be mathematically explained by the constraints on the CB value: $\max(B_p(t), B(t), \kappa S) \leq V \leq \max(B_c(t), \kappa S)$ from (II.5). Indeed, when $\kappa S > B_c(t)$ then, these inequalities become $\kappa S \leq V \leq \kappa S$ (since $B_p(t) < B_c(t)$ and $B_p(t) \leq B_c(t)$). Intuitively, the reason why $V(S,t) = \kappa S$ for $\kappa S \geq B_c$ is that if the issuer did not call the CB when $V(S,t) = B_c(t)$, then it is optimal for the issuer to call it immediately to avoid loosing money (if the stock price were to go up), and obviously the holder would choose to receive κS and not $B_c(t) \leq \kappa S$. Therefore $V(S,t) = \kappa S$ for $\kappa S \geq B_c$.

We have also plotted the value of a CB after changing the percentage R_{rec} of the face value of the bond you recover at default. One would expect that a CB for which we can recover more money at default is more valuable. This is what we observe: $V(S,t,R_{rec}=0) < V(S,t,R_{rec}=50\%) < V(S,t,R_{rec}=100\%)$.

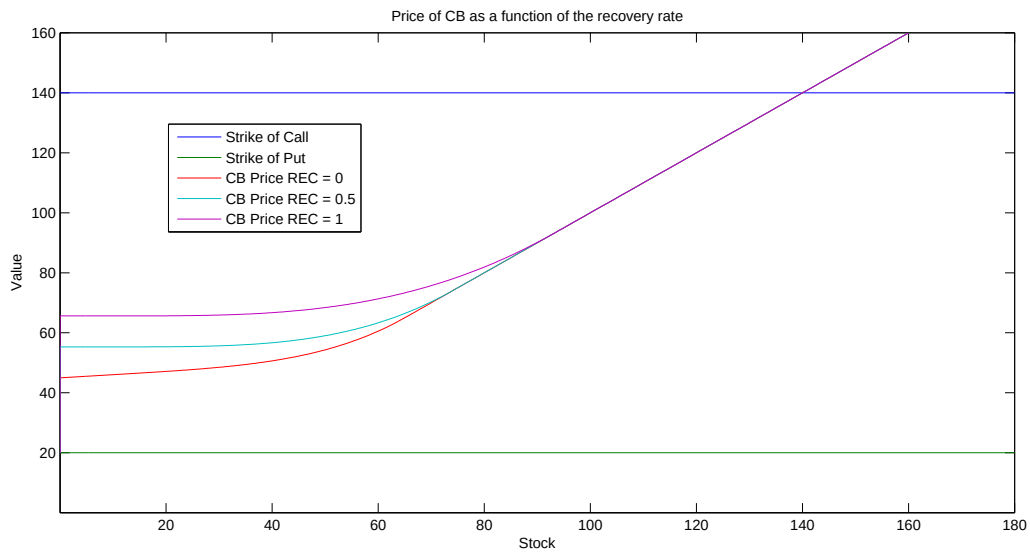

Figure 2: Price of a CB when the recovery at default varies

For other parameters such as interest rate, default probability, dividend rate and volatility, we have plotted the percent error with respect to our benchmark. We define

$$\text{absolute error}(x_i,t^m) = |u_{\text{FEM}}(x_i,t^m) - u_{\text{benchmark}}(x_i,t^m)|$$

$$\text{relative error}(x_i,t^m) = \frac{|u_{\text{FEM}}(x_i,t^m) - u_{\text{benchmark}}(x_i,t^m)|}{u_{\text{benchmark}}(x_i,t^m)}$$

$$\text{percent error}(x_i,t^m) = 100 \times \frac{|u_{\text{FEM}}(x_i,t^m) - u_{\text{benchmark}}(x_i,t^m)|}{u_{\text{benchmark}}(x_i,t^m)}$$

In fact, it is not necessary to test both the interest rate and the default probability because they have the same effect on the pricing. For instance, the price of a CB stays the same (in one dimension) when the sum $(r + p)$ remains constant.

We emphasize that this commercial software may not always provide a very good solution because we could not improve its accuracy (the number of grid points). It will nevertheless give us a good idea of whether or not our CB price is accurate.

We now present some tests we have done. We will not show all of them but focus on the most important ones. Here are some critical parameters we have tested:

- Interest Rate impact

  For this test, we have taken interest rates that go from 0% to 50%. Of course these values are not all realistic, the idea is to test our pricer even in extreme cases. All other characteristics of the CB, or parameters for computation, are those presented in Table 1 and Table 2. Here are some results we get when the interest rate is fixed to 5%

Stock Price	FEM Price	Benchmark Price	Absolute Error	Percent Error
$S = 2.009623$	44.903242	44.903361	1.19e-4	2.65e-4
$S = 4.014968$	44.903243	44.903361	1.19e-4	2.65e-4
$S = 8.810578$	44.903837	44.903983	1.47e-4	3.27e-4

Table 3: $r = 5\%$. the CB behaves like a regular bond

  In Table 3, the CB has exactly the value of a regular bond. The reason is that the stock price is so low that the option to convert (embedded in the CB) has no value. For this stock price range, both errors are very small and probably coming from a day count convention in the software used.

Stock Price	FEM Price	Benchmark Price	Absolute Error	Percent Error
$S = 15.471551$	44.925077	44.925593	5.16e-4	1.15e-3
$S = 19.334225$	44.980021	44.981049	1.03e-3	2.29e-3
$S = 36.002116$	46.580617	46.583925	3.31e-3	7.10e-3
$S = 50.589987$	52.321479	52.313252	8.23e-3	1.57e-2

Table 4: $r = 5\%$. the CB behaves like a hybrid instrument

In Table 4, the value of the CB is more than the bond (Table 3). The stock price has a value large enough (depending on the volatility of the stock) to give value to the option to convert. When the CB behaves like a hybrid instrument, the error is larger than in the other case (see Table 3 and Table 5). This is due to the fact that the algorithm converges with more difficulty.

Stock Price	FEM Price	Benchmark Price	Absolute Error	Percent Error
$S = 58.923874$	58.924709	58.923873	8.35e-3	1.42e-2
$S = 90.945819$	90.945819	90.945818	3.62e-7	3.98e-7
$S = 137.115154$	137.115154	137.115154	1.99e-7	1.45e-7

Table 5: $r = 5\%$. the CB has the conversion value

As we can see in Table 5, when the stock price is higher than a certain value, the CB has exactly the conversion value. In this case, the share price is so high that the bond (which is supposed to provide downside protection to the holder) has no value. The errors are naturally small (here, there are just rounding errors), since the CB has exactly the conversion value.

We should notice that the tolerance parameter in Table 2 for the PSOR is used with the Euclidean norm whereas the errors are computed using the absolute norm. However, for sufficiently small tolerance in the PSOR and smooth solution, it should not matter.

The following graph gives the price of a CB when interest rates $r = 0\%, 5\%, ..., 20\%, 25\%$:

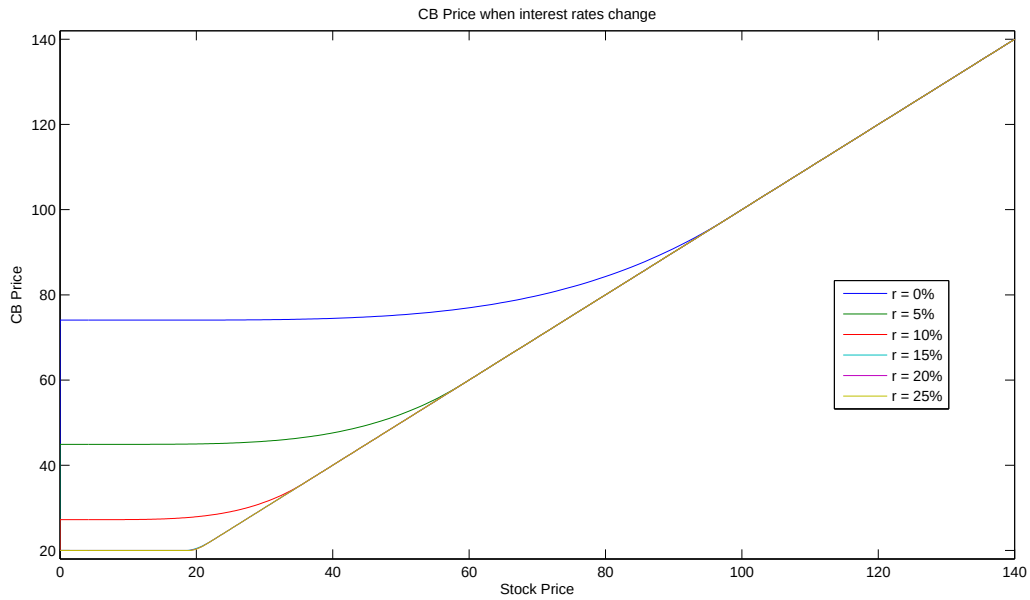

Figure 3: CB Price as a function of Stock Price when the interest rate changes

The price of the CB decreases when interest rates increase, which can be explained mathematically by the fact that we discount the payoff of a convertible with a smaller factor. Moreover, when the interest rate changes, the exercise boundaries are modified. We will see that in more detail in the following sections.

For interest rates higher than 15% the CB prices are very close. However they are not exactly equal and we can see that by zooming close to $S = B_p = 20$:

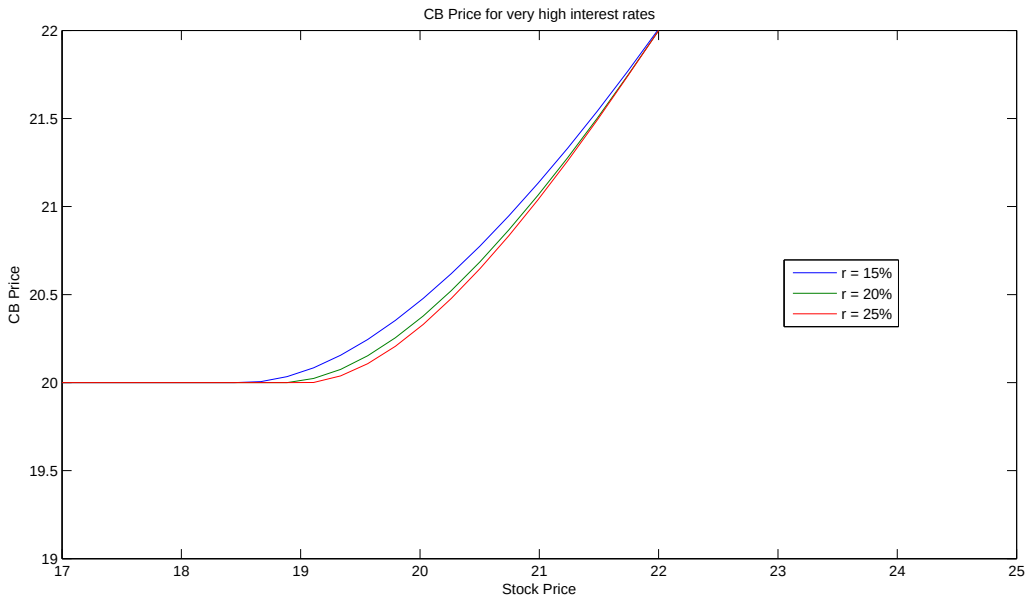

Figure 4: CB Price as a function of Stock Price when the interest rate changes (zoom)

When we focus on the percent error, we observe that for low interest rates (below 15%), the FEM price is really close to our benchmark, i.e. the percent error is less than 10 basis points ($< 0.1\%$). When the interest rate is high (more than 15%), the error is generally bigger when the stock price is close to the lower constraint ψ_2 defined in (III.4) (for a conversion ratio of $\kappa = 1$). Here are plotted the percentage errors as a function of the stock price when the interest rate $r = 0\%, 10\%, 20\%$

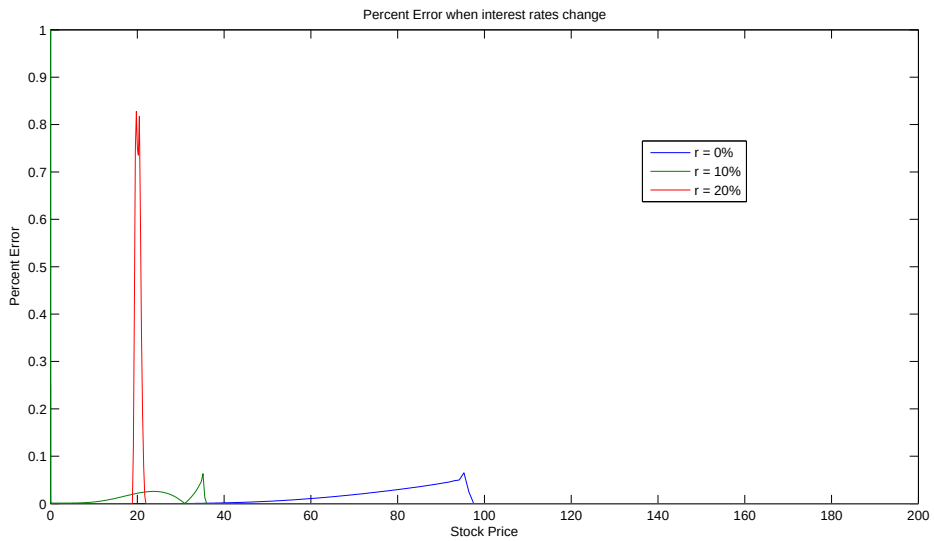

Figure 5: Percent Error as a function of Stock Price when $r = 0\%, 10\%, 20\%$

For some tests, e.g. when $r = 20\%$, we do not force the Psor algorithm to stop after a number $jmax = 500$ of iterations (see Table 2). To check that $jmax = 500$ is not really a limitation, we just remove the constraint in the PSOR algorithm to stop if the maximum number of iterations is reached. The solutions we find for $jmax = 500$ and $jmax = \infty$ are the same up to 10^{-8}.

Here (actually when $r \geq 20\%$), the difference with our benchmark can go up to 0.80% in percent error. It is not due to the slow convergence of the Psor but really to the intrinsic model used. The possible explanation is that the accuracy of our benchmark is not good enough (the accuracy could not be changed for the benchmark price).

- <u>Volatility impact</u>

  For this new set of tests, we consider all the parameters to stay constant, as in Table 1 and Table 2 except the volatility that will vary between 0% and 80%. These values correspond to a frozen or extremely volatile stock price which can give us a broad range of cases.

  Like we have said we have implemented the model presented in part II, section 1.3, but we have used two different lower constraints ψ_2:

$$\psi_2(S,t) = \max(B_p(t), \kappa S) \tag{IV.1}$$

$$\text{and} \qquad \psi_2(S,t) = \max(B_p(t), B(t), \kappa S) \tag{IV.2}$$

  This can have a considerable impact on the accuracy of the solution at the boundaries, and we will see that when we have a look at the values for the CB price.

Here are the prices we obtain when the volatility of the stock increases using the constraint (IV.1):

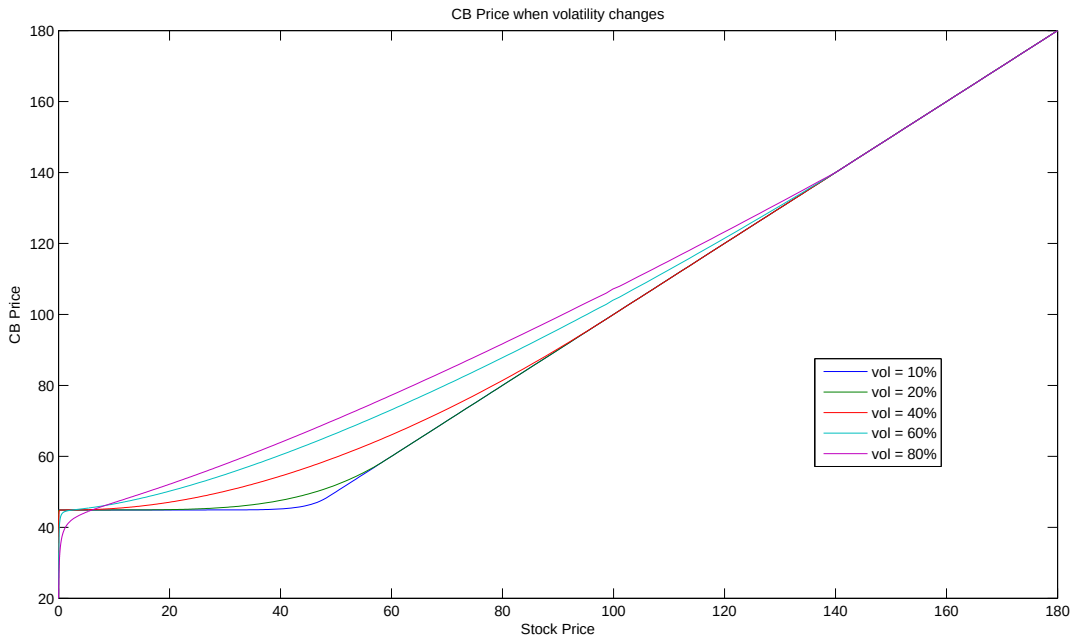

Figure 6: Price of a CB as a function of Stock Price when $\sigma = 10\%, 20\%, 40\%, 60\%, 80\%$

When volatility increases, the value of the CB increases. The same behavior can be observed with European and American options on the stock. The intuitive explanation is that high volatility enables the stock price to fluctuate much more that low volatility. High volatility therefore gives a higher probability that the stock prices reaches a high price, which is what the holder of a CB could hope for.

We can see on Figure 6 that when the stock price is close to 0 and the volatility is very high (typically greater than 60%), the price of the CB can be less than the lower constraint ψ_2. This is not coherent. The reason we observe this phenomenon is that we have for computational reasons forced the solution to have homogeneous Dirichlet boundary conditions (value 0 for $S \mapsto 0$ and $S \mapsto \infty$). In order to have some smoothness in our price the FEM and the Psor produce values that are smaller that ψ_2.

To have a better picture, let us compare these results to those given by our benchmark. We have plotted the relative error for a volatility equal to 0%, 20%, 40% and 60% (Figure 7).

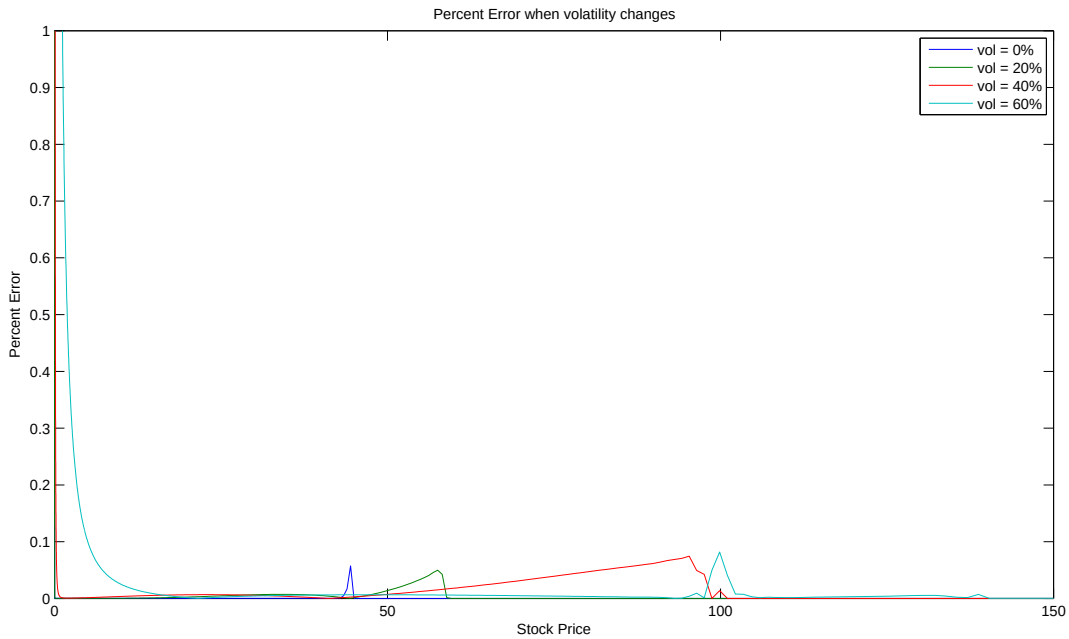

Figure 7: Percent Error as a function of Stock Price when $\sigma = 0\%, 20\%, 40\%, 60\%$

The accuracy is quite good generally since it is less than 10 basis points. We can see thanks to this graph the usefulness of using Excess to Payoff technique or non-Dirichlet boundary conditions (see [Schw07]). Indeed, at the extreme left of our domain (when $S \to e^{-R} \approx 0 \Leftrightarrow x \to -R$), the homogeneous Dirichlet conditions are not a good choice.

We have changed the scale of this graph and plotted the Percent Error as a function of the log-price to focus on the error when the stock price is low (Figure 8).

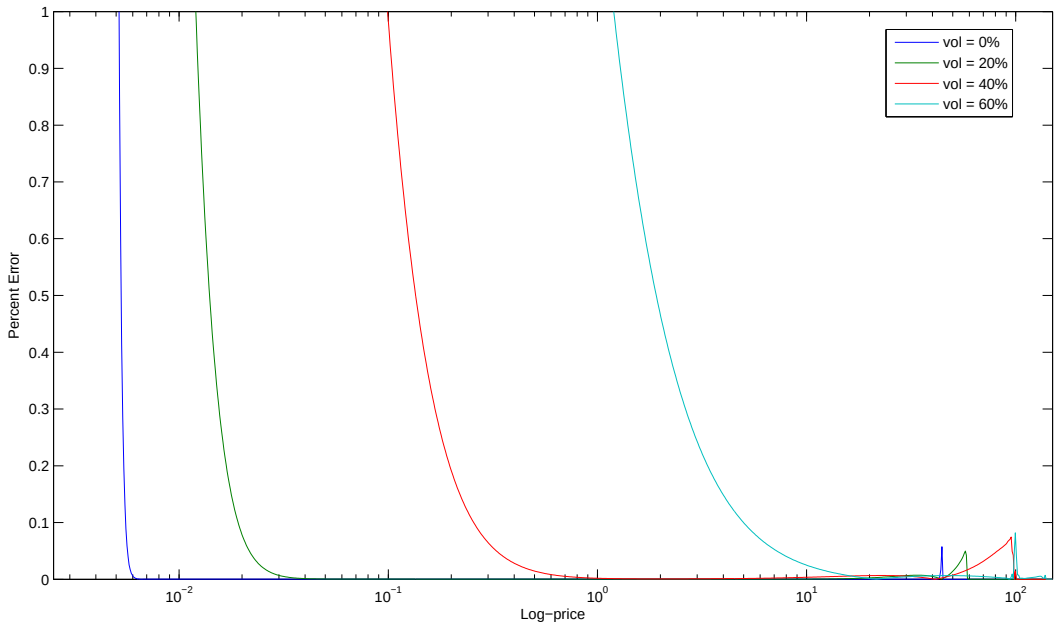

Figure 8: Percent Error as a function of log-Price when $\sigma = 0\%, 20\%, 40\%, 60\%$

There are several ways to improve this issue.

One would be to give a non-zero value to the CB when $S \to 0$. Basically, one could assume the value of the CB when the stock price is zero is just the value of a bond with the same characteristics as the CB (without conversion). This can be explained by the fact that the option to convert has no value. The value of such a bond with no option to convert is ψ_2 (when $\kappa S \to 0$).

Another way is to work with the Excess to payoff (see [Schw07]) and compute the Excess to payoff value instead of that of the CB. Here the Excess to payoff corresponds to the function $u - \psi_2$. However, one would have to derive the corresponding LCP for the Excess to payoff and not (III.5). Then, we would impose homogeneous Dirichlet boundary conditions that are justified by the above paragraph.

The way we have decided to improve the accuracy when the stock price S is "small" is to include the additional constraint that the value of the CB has to be greater than that of the bond. This is the reason why we chose to introduce the value of the bond B (see (IV.2)) whereas it is not done like this for the constraints in the AFV model (see [AFV03] or (IV.1)).

As a result, we can actually decrease tremendously the difference between our price and the benchmark when volatility is very high. The percent error becomes less than 10^{-2} for all volatilities. We did not display the result since there is nothing to see, because the percent error is so low.

We have also tested the impact of other parameters such as dividend yield q, higher/lower call put prices B_c and B_p, no call ($B_c = \infty$) or/and not put ($B_p = -1$). We also looked if there were any noticeable errors when we change maturity, values of coupons, conversion ratio, etc.

The percent error compared to our benchmark remains generally smaller than 0.1 to 0.2 (equivalent to a relative error of 0.1% to 0.2%). The interest rate (or credit spread) and volatility parameters are therefore those that increase the difference between our valuation and our benchmark.

1.2 Convergence rate

We would like to find the speed of convergence of our algorithm. This speed depends on the number of time and space steps.

First of all, since we do not have a closed form solution to check our numerical prices, we can have an idea of the rate of convergence in the following way. Let us denote by u^{k+1} the estimate we get from the PSOR after one iteration with initial value u^k. The PSOR stops when either there has been more than $jmax$ iterations or the amount $||u^{k+1} - u^k|| < tol$, where $||\ ||$ denotes the Euclidean norm.

Let us denote u the exact solution to the LCP (III.5) on the truncated domain. $\tilde{u}_{h,k}$ denotes the solution to the matrix LCP resulting from the discretization of (III.5) into (III.20)-(III.22). The PSOR gives us an approximation $u_{h,k}$ to $\tilde{u}_{h,k}$ up to a certain tolerance tol.

Let us denote u_k^{ref} the solution to (III.20)-(III.22) with a very "low" tolerance (we take $tol_{ref} =$ 1e$-$12), a number of steps in log-price $N_{ref} + 1 = 2^{15}$ and a time step $k = 2^7$ fixed. u_k^{ref} will give us an approximation to $\tilde{u}_{h,k}$.

For $h \geq h_{ref} := \frac{2R}{N_{ref}+1}$, we can find another approximation to $\tilde{u}_{h,k}$ that we denote by $v_{h,k}$ with a tolerance equal to $tol = tol_{ref}$ and the same time step k as for $u_{h,k}^{ref}$.

To find the rate of convergence in the price space, we are looking for a parameter α, such that:

$$||u_k^{ref}(T) - v_{h,k}(T)|| \leq C_1 h^\alpha \qquad ,\forall h \geq h_{ref}.$$

where $v_{h,k}(T)$ is the CB price we get from the PSOR at time to maturity $\tau = T$ or $t = 0$.

Then, we would have

$$\begin{aligned} ||u(T) - v_{h,k}(T)|| &\leq ||u(T) - \tilde{u}_{h,k}(T)|| + ||\tilde{u}_{h,k}(T) - v_{h,k}(T)|| \\ &\leq ||u(T) - \tilde{u}_{h,k}(T)|| + ||\tilde{u}_{h,k}(T) - u_k^{ref}(T)|| + ||u_k^{ref}(T) - v_{h,k}(T)|| \end{aligned}$$

So, in order to have $||u(T) - v_{h,k}(T)|| \leq C_1 h^\alpha$, we need $||u_k^{ref}(T) - v_{h,k}(T)|| \leq C_1 h^\alpha$, which is what we will measure numerically.

A log-log plot of $||u_k^{ref}(T) - v_{h,k}(T)||_{L^\infty}$ as a function of h gives us an approximation of α (Figure 9).

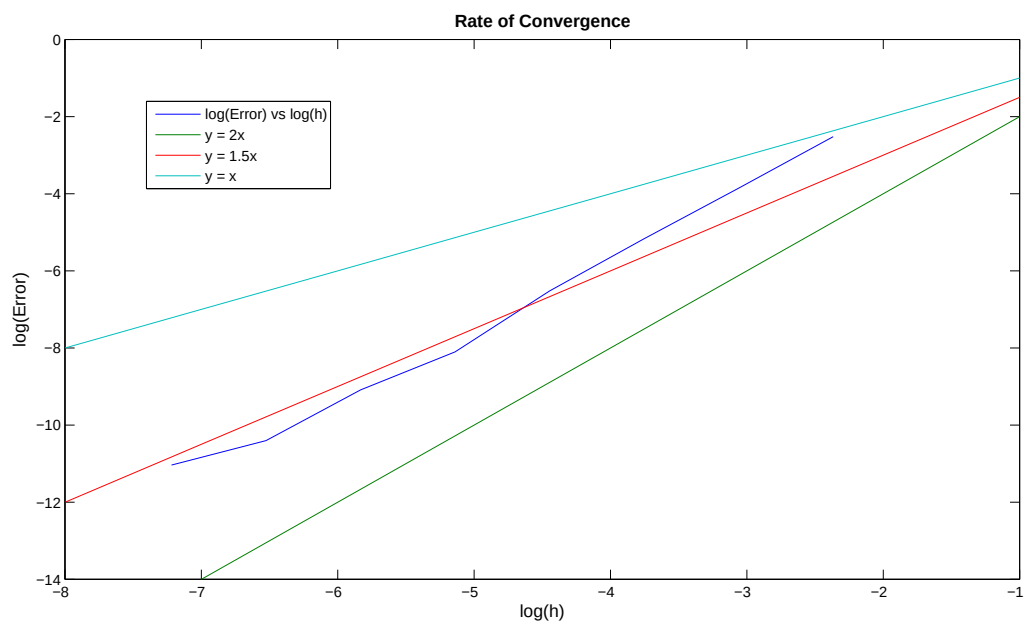

Figure 9: $\log(||u_k^{ref}(T) - v_{h,k}(T)||_{L^\infty})$ as a function of $log(h)$(blue curve)

The parameter α can be found by evaluating the slope of the function plotted on Figure 9. When, h is not "very small" ($N+1 < 2^{12}$), then we can see and calculate that the convergence rate α is close to 2.

When h becomes "very small" ($N+1 \geq 2^{12}$), then the rate of convergence α is roughly 1.5. We can not fully rely on Figure 9 to find the convergence rate because if we have $||u_k^{ref}(T) - v_{h,k}(T)|| \leq C_1 h^\alpha$, then we would need $tol \leq C_1 h^\alpha$ to observe the real rate of convergence. For computational power restrictions, we did not consider $N_{ref}+1 > 2^{15}$ or a tolerance $tol < 10^{-12}$. Therefore, the convergence rate α is between 1.5 and 2.

The rate of convergence α in price space does not depend on the tolerance required in the PSOR algorithm. α only depends on characteristics of the Finite Element space the solution belongs to, which is independent of the PSOR. Furthermore, it has been shown that α does not depend on k.

We can do the same analysis for the rate of convergence in time. We fix $N+1 = 2^8$, and we consider $tol = 10^{-12}$ and $M_{ref} = 2^{16}$. Then we take $M \leq M_{ref}$ or equivalently $k = \frac{T}{M} \geq k_{ref} := \frac{T}{M_{ref}}$, we can compute a reference $u_h^{ref}(T)$, and approximations $v_{h,k}(T)$ to this reference by changing $k \geq k_{ref}$.

We are searching for a parameter $\beta > 0$ such that

$$||u_h^{ref}(T) - v_{h,k}(T)|| \leq C_2 k^\beta$$

We have plotted $log(||u_h^{ref}(T) - v_{h,k}(T)||_{L^\infty})$ as a function of $log(k)$ (Figure 10):

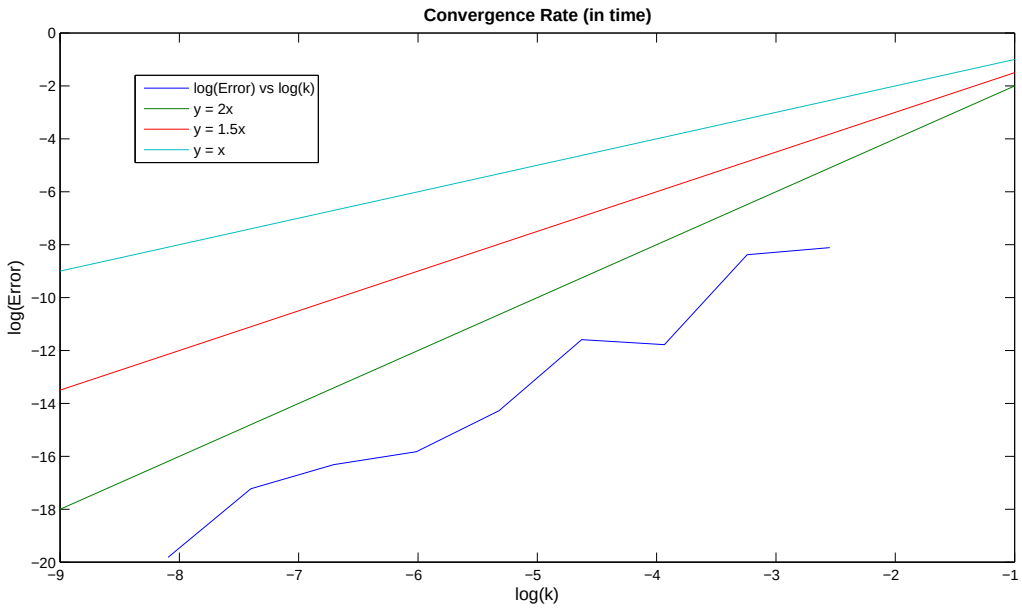

Figure 10: $\log(||u_h^{ref}(T) - v_{h,k}(T)||_{L^\infty})$ as a function of $log(k)$ (blue curve)

We clearly do not observe a straight line (blue curve). However, if we do a regression analysis we find a β approximately equal to 2.5. The real value of β is probably around 2, which is the maximum we can obtain.

Let us add that the parameter β does not depend on the tolerance used in the PSOR algorithm or N.

1.3 Exercise boundaries

In this section, we will consider the case where there are no coupons paid, so that the price of the CB and the constraints applied to the value of the CB are continuous.

The price of a CB is subject to two constraints (see the LCP (II.5) and constraints (III.3)-(III.4)).

Let us consider the options the holder of the convertible has: he has the option to convert his bond into shares or sell the CB. The option to convert at time t is the reason why $V(S,t) \geq \kappa S$ (see part II, section 1.2.2). We can actually find for each moment t, the set of values of the stock price for which we have $V(S,t) = \kappa S$. The first moment this equality holds, the holder decides to exercise his right to convert. Let us consider the following set:

$$\mathcal{E}_{conv} := \{(\inf\{S, V(S,t) = \kappa S\}, t), \qquad , \forall t \in [0,T]\} \tag{IV.3}$$

This corresponds to the conversion exercise boundary. It is the values (S,t) for which it is optimal to the CB holder to convert.

The set $\inf\{S, V(S,t) = \kappa S\}$ is not empty and has a unique element $\forall t \in [0,T]$ (see [BL78]). Therefore $\mathcal{E}_{conv}$ is well defined.

The holder can also sell the CB, either at the price $B_p(t)$ at time t or at least at the price of the straight bond $B(t)$ defined in part II, section 1.3. This is the reason we have $V(S,t) \geq \max(B(t), B_p(t))$. For every moment t, we can define the set

$$\mathcal{E}_{sell} := \{(\sup\{S, V(S,t) = \max(B(t), B_p(t))\}, t), \qquad , \forall t \in [0,T]\} \tag{IV.4}$$

This set is once again well defined (see [BL78]) and corresponds to the selling exercise boundary. At maturity, if we have $V(S,T) = B(T)$, the CB is not sold, the holder simply receives the redemption payment.

These are the exercise boundaries of the holder, i.e. the values of (S,t) for which the holder needs to take action. The issuer has the right to call the CB. This is the reason why we have $V(S,t) \leq B_c(t)$ when $\kappa S < B_c(t)$. If $\kappa S \geq B_c(t)$, the issuer should clearly exercise his call option to avoid loosing money (if the stock price increases). Therefore we can define the set

$$\mathcal{E}_{call} := \{(\inf\{S, V(S,t) = \max(B_c(t), \kappa S)\}, t), \;\; \forall t \in [0,T]\} \tag{IV.5}$$

This set is well defined and corresponds to the values of (S,t) for which the issuer should call the CB.

1.3.1 Computation of the exercise boundaries

We can compute the exercise boundaries presented above. Of course the more precise the price and time steps are, the more the exercise boundary will be computed accurately. It is also possible to make an adaptive Finite Element grid that is more accurate around these boundaries to have a better approximation.

Here are the exercise boundaries, we find with a constant time step with $N + 1 = 2^{10}$, $M = 2^7$ and $tol = 10^{-12}$. The other parameters used are those of Table (1) and Table (2).

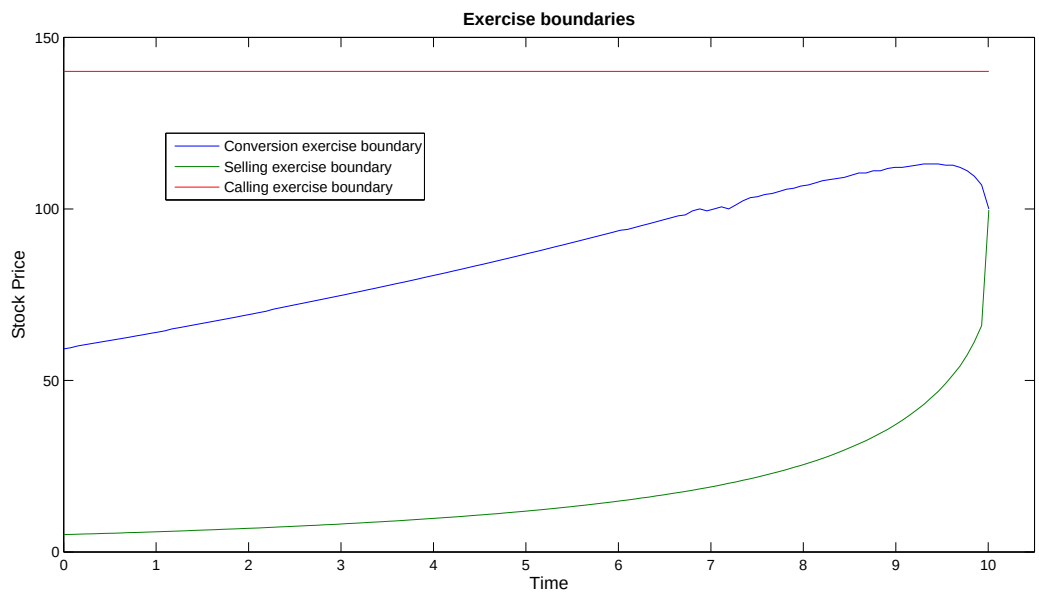

Figure 11: exercise boundaries with the parameters presented in Table (1) and Table (2), except $tol = 10^{-12}$

First of all, it is clear that the moment the issuer should call the bond is the moment the value of the CB is equal to the call price. This seems intuitive since, if the issuer called when $V < B_c$ then the issuer would loose the amount $B_c - V$ and give it for free to the holder. And if the issuer calls the CB when $V > B_c$, since the CB price does not have upward jumps, the issuer could have called it when $V = B_c$. Therefore, calling the CB at the moment its price has the call value is optimal.

The selling exercise boundary is the same as the exercise boundary of a put option. The holder of a CB has actually got a put option with strike $\max(B_p, B)$ on the value of the CB.

The conversion exercise boundary has the shape of a call option exercise boundary. The reason is that one can view a CB as a callable and puttable bond ($\max(B_p, B)$) plus an option to convert into κ shares with a strike equal to this special bond $\max(B_p, B)$. Since the strike increases with time, the value of the CB is less (if S remains constant) with time, and therefore, for the CB value to reach κS, one needs a higher S. This is why the exercise boundary increases in the beginning. Of course, the exercise boundary cannot increase indefinitely, because at maturity the CB has a known value.

1.3.2 Impact of a change in parameters in the exercise boundaries

It is interesting to know how the exercise boundaries can move in relation to key parameters that enter in the pricing of a CB. We have decided to focus on the impact of a change in interest rate (equivalently credit spread) and volatility.

In Table 12, we can see that the volatility parameter has a big impact on the location of the exercise boundaries. Indeed, the higher the volatility, the higher the conversion boundary (resp. the lower the selling boundary). When the stock price volatility is high, it gives an incentive to the holder of the CB to keep the CB as long as he can, because the possibility of the stock price being favorable is high. Therefore, the moment the holder would decide to convert/sell is postponed.

The exercise boundary of the issuer of the CB does not change at all. In fact, it is always optimal for him to call the CB as soon as the stock price reaches the call value $B(t)$.

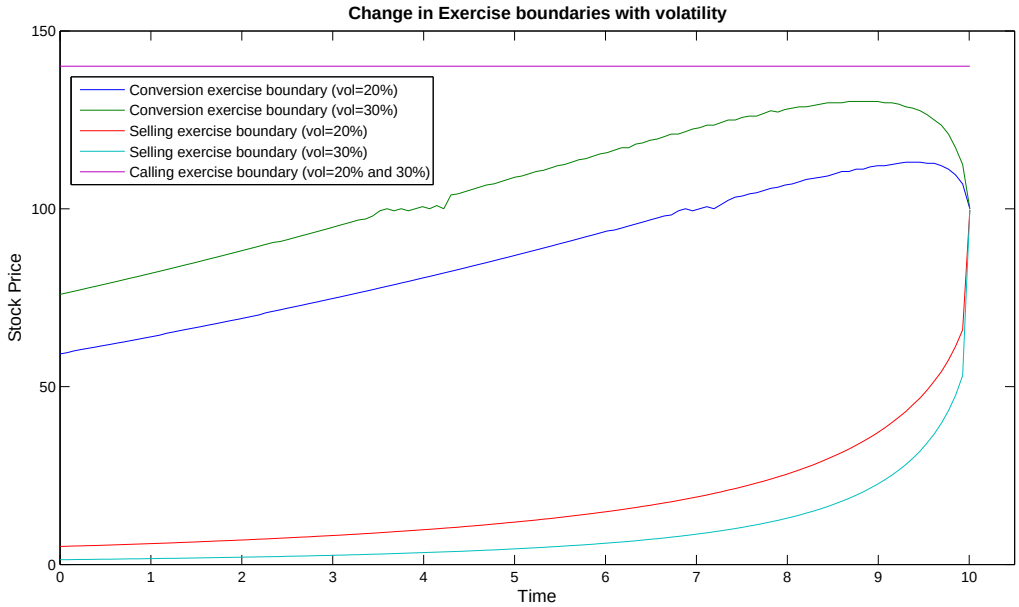

Figure 12: Impact of a change in volatility on the exercise boundaries

We can see some instability in the conversion exercise boundary on Figure 12 around $\kappa S = F$. This can be dealt with by increasing the number of grid points and iterations in the PSOR.

Figure 13 indicates how the exercise boundaries should change when the interest rate (or credit spread) increases.

If the interest rate increases and the stock price remains constant, the value of the CB will decrease (since we discount at a higher rate). Therefore, we will have $V < \kappa S$. The exercise boundary is then lower when interest rates increase.

The same reasoning applies to the selling exercise boundary. If r increases, then $\max(B_p, B)$ decreases, and so for V to attain $\max(B_p, B)$, the stock price is smaller. This means that the selling boundary is lower, which is what we observe.

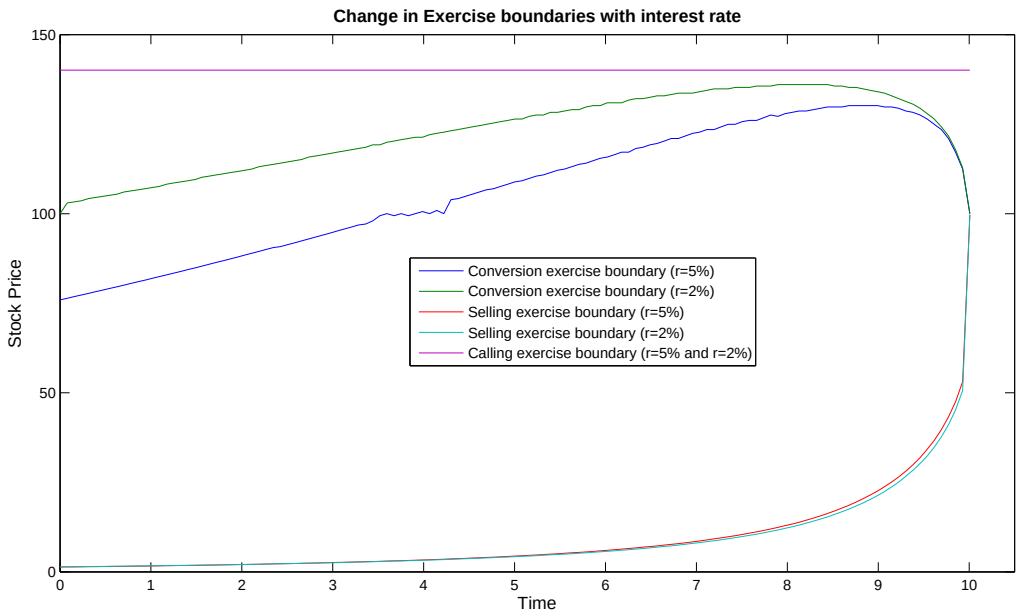

Figure 13: Impact of a change in interest rate on the exercise boundaries

Part V
Conclusion

In this paper, we have presented a model, based on [AFV03] and [AB02], that is used to price a CB. We started with a one factor model where the stock price is stochastic. For this model, we have derived the Linear Complementarity Problem and parabolic variational inequality that the price of a CB has to satisfy. We also have presented a model where both the stock price and the interest rate are stochastic for a general short rate process.

Furthermore, since there is no closed-form solution, the PVI has been localized and discretized to find a numerical solution. We have discretized in time by a theta scheme and in log-price using Finite Elements. Even though, FEM are used commonly in Physics, they are not yet popular in Finance. This thesis is innovative in the sense we provide a detailed analysis on how to apply FEM to price a CB with complex features for a two factor model.

We have compared the FEM prices to those provided by a software. The results are generally very good, since the absolute and relative difference do not generally exceed $2e10^{-3}$. We also provide a numerical estimate of the rate of convergence of our algorithm: 1.5 in space and 2 in time, which is very good.

To have a better understanding of the product, we have also computed the exercise boundaries which correspond to the stock value for which the holder or issuer needs to take action (convert, sell, call).

Finally, to continue this work, one could extend the model to a three factor model where for instance the credit spread is stochastic or perhaps introduce a FX factor (see [Y01]).

References

[AB02] Andersen L., Buffum D. (2002). *Calibration and Implementation of Convertible Bond Models.* SSRN.

[AFV02] Ayache E., Forsyth P.A., Vetzal K.R. (2002). Next generation models for convertible bonds with credit risk. *Wilmott Magazine.*

[AFV03] Ayache E., Forsyth P.A., Vetzal K.R. (2003). The valuation of convertible bonds with credit risk. *Journal of Derivatives.*

[BL78] Bensoussan A., Lions J.L. (1978). *Applications des inéquations variationnelles en contrôle stochastique.* Dunod.

[BM01] Brigo D., Mercurio F. (2001). *Interest Rate Models - Theory and Practice.* Springer, Second Edition.

[DS99] Duffie D., Singleton K. (1999). Modeling Term Structures of Defaultable Bonds. *The Review of Financial Studies.*

[Hul03] Hull J. (2003). *Options, Futures and other Derivatives.* 5th edition, Prentice Hall.

[JLL90] Jaillet P., Lamberton D., Lapeyre B. (1990). Variational Inequalities and the pricing of American options. *Acta Applicandae Mathematicae.*

[L07] Li D. (2007). *On convergence of Numerical Methods for Pricing Convertible Bonds.* M.Sc. Thesis, Department of Computer Science, University of Toronto.

[LL97] Lamberton D., Lapeyre B. (1997). *Introduction au Calcul Stochastique Appliqué à la Finance.* Ellipses.

[N96] Nyborg K.G. (1996). The use and Pricing of convertible bonds. *Applied Mathmatical Finance.*

[Scho07] Schönbucher P. (2007). *Term Structure and Credit Risk Models.* Lecture Notes. ETH Zürich.

[Schw07] Schwab C. (2007). *Computational Methods for Quantitative Finance.* Lecture Notes. ETH Zürich.

[TKN01] Takahashi A., Kobayashi T., Nakagawa N. (2001). *Pricing Convertible Bonds with Default Risk: A Duffie-Singleton Approach.* EconPapers.

[WHD95] Wilmott P., Howison S., Dewynne J. (1995). *The Mathematics of Financial Derivatives, a Student Introduction.* Cambridge University Press.

[Y01] Yigitbasioglu A. B. (2001). *ISMA Centre.* Pricing Convertible Bonds with Interest Rate, Equity, Credit and FX Risk.

Appendix

A Derivation of the LPC satisfied by the CB price

In this section, we provide an idea of how to derive the LCP (III.5). The LCP (III.5) are obtained after changing variables in LCP (II.5). Even though we implement in our work a model for convertibles with constant recovery, we can actually derive the LCP for a more general recovery $R_{rec}X$, where X is a random variable (LCP (II.3)).

The change of variables $x = log(S)$ for $S \in (0, \infty)$ leads to the following dynamics (using Itô's formula)

$$dx_t = (r - q - \frac{\sigma^2}{2} + p\eta)dt + \sigma dW_t$$

To have a clearer idea of how (III.5) can be derived, we have to come back to the expression of the payoff of a convertible. In fact the problem of pricing a CB is exactly the problem that arises in stochastic games and optimal exercising. On the one side the holder wants to maximize his payoff and on the other the issuer wants to minimize it.

To ease calculations and notation, we consider the case where there are no coupons paid. Then the dirty price (= clean price + accrual interest) is equal to the clean price for all quantities (B_c, B_p, B and the CB value u). Furthermore, B_c and B_p are constant and $B \in C^\infty(0, T)$.

We recall that function ψ_1 and ψ_2 have been defined by $\psi_1 = \max(B_c, \kappa e^x)$ and $\psi_2 = \max(B_p, B(t), \kappa e^x)$.

The function B is the value of a callable bond (with constant interest rates) with face value F, so using the notations of part II, section 1.2.2, we get

$$B(t) = \underbrace{\max(F, B_c)}_{\text{constant}} e^{-(r+p)(T-t)}$$

Let us denote by τ_2 (respectively τ_1) the first moment the holder decides to convert his CB into shares, put to the CB or the moment the CB has the same value as B (respectively the

moment the issuer decides to call the bond). Basically, τ_1 (resp. τ_2) correspond to the moment the value of the CB u is equal to ψ_2 (resp. ψ_1). The stopping times τ_1 (resp. τ_2) belong to the set of stopping times with values in $([t,T] \cup \{\infty\})$ (resp. $[t,T]$).

We will write equivalently S or e^x. To ease notation, we choose to use S.

We can write the payoff -if no default occurs- of a convertible depending on the stopping times τ_1 and τ_2 as a function h

$$h_{\tau_1,\tau_2}(S, t = \tau_1 \wedge \tau_2 \wedge T) = \begin{cases} \psi_1(S_{\tau_1}, \tau_1) \cdot \ 1_{(\tau_1<\tau_2)\cap(\tau_1\wedge\tau_2\leq T)} \\ \\ + \ \psi_2(S_{\tau_2}, \tau_2) \cdot \ 1_{(\tau_2\leq\tau_1)\cap(\tau_1\wedge\tau_2\leq T)} \end{cases}$$

Nevertheless, if no option is exercised either by the holder or the issuer before maturity and the CB has not defaulted ($\tau_D > T$), we can write the payoff at maturity :

$$h(S,T) = \max(B(T), B_p, \kappa S) = \psi_2(S,T)$$

This explains why $\tau_2 \in [0,T]$ a.s.

Let us define the *Reward Function* . In the general case where the recovery is X (LCPs (II.3)), we have

$$\begin{aligned} J_{S,t}(\tau_1,\tau_2) = \mathbb{E}_{\mathbb{Q}}[\ & (e^{-r(\tau_1-t)}\psi_1(S_{\tau_1},\tau_1)\cdot 1_{(\tau_1<\tau_2)\cap(\tau_1\wedge\tau_2\leq T)} \\ & + e^{-r(\tau_2-t)}\psi_2(S_{\tau_2},\tau_2)\cdot 1_{(\tau_2\leq\tau_1)\cap(\tau_1\wedge\tau_2\leq T)}) \cdot 1_{\tau_D>\tau_1\wedge\tau_2\wedge T} \\ & + e^{-r(\tau_D-t)} \max(\kappa S_{\tau_D}(1-\eta), R_{rec}X(\tau_D)) \cdot 1_{\tau_D\leq\tau_1\wedge\tau_2\wedge T} | \ \mathcal{F}_t] \end{aligned}$$

where $\mathcal{F}_t$ denotes the filtration generated by the process $(x_s)_{0\leq s\leq t}$ and by the Poisson process $(N_s)_{0\leq s\leq t}$ defined in section 1.2.2.

The first part (multiplied by $1_{\tau_D>\tau_1\wedge\tau_2\wedge T}$) can be viewed as pricing a contingent claim that yields the random payoff of a CB if default does not occur. It can be shown (see [Scho07]) that

$$\begin{aligned} J_{S,t}(\tau_1,\tau_2) = \ & \underbrace{1_{\tau_D>t}}_{\text{deterministic}} \cdot \mathbb{E}_{\mathbb{Q}}[\ e^{-(r+p)(\tau_1-t)}\psi_1(S_{\tau_1},\tau_1)\cdot 1_{(\tau_1\leq\tau_2)\cap(\tau_1\wedge\tau_2\leq T)} \\ & \qquad + e^{-(r+p)(\tau_2-t)}\psi_2(S_{\tau_2},\tau_2)\cdot 1_{(\tau_2\leq\tau_1)\cap(\tau_1\wedge\tau_2\leq T)} | \ \mathcal{F}_t] \\ & + \mathbb{E}_{\mathbb{Q}}[\ e^{-r(\tau_D-t)} \quad \max(\kappa S_{\tau_D}(1-\eta), R_{rec}X(\tau_D)) \cdot 1_{\tau_D\leq\tau_1\wedge\tau_2\wedge T} | \ \mathcal{F}_t] \end{aligned}$$

since the stopping time $\tau_1 \wedge \tau_2 \wedge T$ is independent of τ_D.

To deal with the term $\mathbb{E}_{\mathbb{Q}}[e^{-r(\tau_D-t)}\max(\kappa S_{\tau_D}(1-\eta), R_{rec}X(\tau_D))\cdot 1_{\tau_D\le\tau_1\wedge\tau_2\wedge T}]$, we refer to [Scho07] to calculate

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}}[e^{-r(\tau_D-t)}\max(\kappa S_{\tau_D}(1-\eta), R_{rec}X(\tau_D))\cdot 1_{\tau_D\le\tau_1\wedge\tau_2\wedge T}]\\
&\quad=\mathbb{E}_{\mathbb{Q}}[\int_t^{\tau_1\wedge\tau_2\wedge T} e^{-r(s-t)}\max(\kappa S_s(1-\eta), R_{rec}X(s))\cdot 1_{\{N(s^-)=0\}}dN(s)|\ \mathcal{F}_t]
\end{aligned}$$

using the compensator of $N(t)$

$$=\mathbb{E}_{\mathbb{Q}}[\int_t^{\tau_1\wedge\tau_2\wedge T} e^{-r(s-t)}\max(\kappa S_s(1-\eta), R_{rec}X(s))\cdot 1_{\{N(s^-)=0\}}p(s)ds|\ \mathcal{F}_t]$$

assuming sufficient regularity to interchange expectation and integration

$$\begin{aligned}
&=\int_t^{\tau_1\wedge\tau_2\wedge T}\mathbb{E}_{\mathbb{Q}}[e^{-r(s-t)}p\max(\kappa S_s(1-\eta), R_{rec}X(s))\cdot 1_{\{N(s^-)=0\}}|\ \mathcal{F}_t]ds\\
&=\int_t^{\tau_1\wedge\tau_2\wedge T}\underbrace{\mathbb{E}_{\mathbb{Q}}[e^{-r(s-t)}p\max(\kappa S_s(1-\eta), R_{rec}X(s))\cdot 1_{\{\tau_D>s\}}|\ \mathcal{F}_t]}_{=1_{\{\tau_D>t\}}\mathbb{E}_{\mathbb{Q}}[e^{-(r+p)(s-t)}p\max(\kappa S_s(1-\eta),R_{rec}X(s))\ |\ \mathcal{F}_t]}ds\\
&=1_{\{\tau_D>t\}}\int_t^{\tau_1\wedge\tau_2\wedge T}\mathbb{E}_{\mathbb{Q}}[e^{-(r+p)(s-t)}p\max(\kappa S_s(1-\eta), R_{rec}X(s))\ |\ \mathcal{F}_t]ds\\
&=1_{\{\tau_D>t\}}\mathbb{E}_{\mathbb{Q}}[\int_t^{\tau_1\wedge\tau_2\wedge T} e^{-(r+p)(s-t)}p\max(\kappa S_s(1-\eta), R_{rec}X(s))ds\ |\ \mathcal{F}_t].
\end{aligned}$$

Therefore, we can write our Reward Function as

$$\begin{aligned}
J_{S,t}(\tau_1,\tau_2)=1_{\tau_D>t}\cdot\mathbb{E}_{\mathbb{Q}}[\ & e^{-(r+p)(\tau_1-t)}\psi_1(S_{\tau_1},\tau_1)\cdot 1_{(\tau_1\le\tau_2)\cap(\tau_1\wedge\tau_2\le T)}\\
&+e^{-(r+p)(\tau_2-t)}\psi_2(S_{\tau_2},\tau_2)\cdot 1_{(\tau_2\le\tau_1)\cap(\tau_1\wedge\tau_2\le T)}\\
&+\int_t^{\tau_1\wedge\tau_2\wedge T} e^{-(r+p)(s-t)}p\max(\kappa S_s(1-\eta), R_{rec}X(s))ds|\ \mathcal{F}_t].
\end{aligned}$$

Given that default has not occurred at time t, we have

$$\begin{aligned}
J_{S,t}(\tau_1,\tau_2)=\mathbb{E}_{\mathbb{Q}}[\ & e^{-(r+p)(\tau_1-t)}\psi_1(S_{\tau_1},\tau_1)\cdot 1_{(\tau_1\le\tau_2)\cap(\tau_1\wedge\tau_2\le T)}\\
&+e^{-(r+p)(\tau_2-t)}\psi_2(S_{\tau_2},\tau_2)\cdot 1_{(\tau_2\le\tau_1)\cap(\tau_1\wedge\tau_2\le T)}\\
&+\int_t^{\tau_1\wedge\tau_2\wedge T} e^{-(r+p)(s-t)}p\max(\kappa S_s(1-\eta), R_{rec}X(s))ds|\ \mathcal{F}_t].
\end{aligned}$$

The issuer will try to minimize $J_{S,t}(\tau_1,\tau_2)$ by choosing τ_1 optimally and the holder will choose τ_2 to maximize $J_{S,t}(\tau_1,\tau_2)$.

The price of a convertible at time t can be written as

$$V^R(S,t) = \inf_{t\le\tau_1\le\infty} \sup_{t\le\tau_2\le T} J_{S,t}(\tau_1,\tau_2) = \sup_{t\le\tau_2\le T} \inf_{t\le\tau_1\le\infty} J_{S,t}(\tau_1,\tau_2) \tag{V.1}$$

$$u^R(x,t) = \inf_{t\le\tau_1\le\infty} \sup_{t\le\tau_2\le T} J_{x,t}(\tau_1,\tau_2) = \sup_{t\le\tau_2\le T} \inf_{t\le\tau_1\le\infty} J_{x,t}(\tau_1,\tau_2) \tag{V.2}$$

where $J_{x,t} = J_{S=e^x,t}$ and u^R (resp. V^R) denotes the solution to the LCP (III.5) on a truncated domain as explained below (resp. $V^R = u^R(T-t, e^x)$) (see part III, section 2.2.2).

It can be shown under certain conditions that the system of inequalities (III.5) are a PDI formulation of this stochastic control problem (V.2) (see [BL78], chapter 3, part 5.2.1.).

In order to operate in the framework of [BL78], we need to check several assumptions.

First of all, since we are not interested in the value of the convertible bond on for $x \in \mathbb{R}$, let us truncate the problem and work on the open and bounded domain $\Omega_R = (-R, R)$ where $R > 0$. We define $Q = \Omega_R \times (0,T)$. We give homogeneous Dirichlet boundary conditions at $x = \pm R$ to the CB price u as well as the constraints ψ_1 and ψ_2.

Furthermore, in order to use the result in [BL78], we need to consider $\sigma > 0$. We can show that $f \in L^2(Q)$ and $\psi_i \in W^{1,1,2}(Q),\ \forall i \in \{1,2\}$.

Let us assume that the interest rate and credit spread are 0. This enables ψ_2 not to depend on time.

Note that in [BL78] we need $\psi_i \in W^{2,1,2}(Q)$, however using that $r, p = 0$, we can follow [JLL90] (section 3) and find a sequence of functions $\{(\psi_i)_n \in W^{1,1,2}(Q),\ \forall n\}$ that converges uniformly in x to ψ_i and therefore reduce the assumption $W^{2,1,2}(Q)$ to $W^{1,1,2}(Q)$.

Following these assumptions, the solution to the LCP (III.5) on the domain Q exists and is unique on Q. For the general case, we have not derived the LCP and assume there exists a unique solution.

B Calculation of matrices in one dimension

We calculate here the matrices that enter in the implementation of the one dimensional pricing problem for a convertible with constant recovery.

To ease calculations, we will consider a constant price-space step (h_i is constant equal to h). According to the section 4.1.3, we just need to calculate the terms $\mathbf{M}_{i,j}$ and $\mathbf{A}_{i,j}$ for $|i-j| \leq 1$.

Let us calculate the terms $\mathbf{M}_{i,i}$, $\mathbf{M}_{i,i+1}$ ($\mathbf{M}_{i,i+1} = \mathbf{M}_{i,i-1}$, since $\mathbf{M}$ is symmetric and $h_i = h$):

$$\begin{aligned}
\mathbf{M}_{i,i} &= (b_i, b_i) = \int_{-R}^{R} b_i(x)^2 dx = 2\int_{x_{i-1}}^{x_i} b_i(x)^2 dx = \frac{2}{h^2}\int_{x_{i-1}}^{x_i} (x - x_{i-1})^2 dx \\
&= \frac{2}{3h^2}[(x - x_{i-1})^3]_{x_{i-1}}^{x_i} = \frac{2h}{3} \\
\mathbf{M}_{i,i+1} &= (b_i, b_{i+1}) = \int_{-R}^{R} b_i(x) b_{i+1}(x) dx = \int_{x_i}^{x_{i+1}} b_i(x) b_{i+1} dx = \frac{1}{h^2}\int_{x_i}^{x_{i+1}} (x_{i+1} - x)(x - x_i) dx \\
&= \frac{1}{h^2}\int_{x_i}^{x_{i+1}} (x_{i+1} - x_i + x_i - x)(x - x_i) dx = \frac{1}{h^2}(h\int_{x_i}^{x_{i+1}} (x - x_i) dx - \int_{x_i}^{x_{i+1}} (x - x_i)^2 dx) \\
&= \frac{1}{h^2}(\frac{h^3}{2} - \frac{h^3}{3}) = \frac{h}{6}
\end{aligned}$$

To calculate $\mathbf{A}_{i,i}$, $\mathbf{A}_{i,i+1}$ and $\mathbf{A}_{i+1,i}$, we use

$$\mathbf{A}_{i,j} = \frac{\sigma^2}{2}\mathbf{S}_{i,j} + (\frac{\sigma^2}{2} + q - r - p\eta)\mathbf{C}_{i,j} + (r + p)\mathbf{M}_{i,j}$$

where $\mathbf{S}$ and $\mathbf{C}$ are defined in (III.23) and $\mathbf{M}$ is the Mass matrix from above. Thus it is sufficient to calculate

$$
\begin{aligned}
\mathbf{S}_{i,i} &= \int_{-R}^{R} b_i'(x)^2 dx = \frac{1}{h^2}\int_{x_{i-1}}^{x_i} dx = \frac{2}{h} \\
\mathbf{S}_{i,i+1} &= \int_{-R}^{R} b_{i+1}'(x) b_i'(x) dx = \int_{x_i}^{x_{i+1}} -dx = -\frac{1}{h} \\
\mathbf{S}_{i,i-1} &= \int_{-R}^{R} b_{i-1}'(x) b_i'(x) dx = \frac{1}{h^2}\int_{x_{i-1}}^{x_i} -dx = -\frac{1}{h} \\
\mathbf{C}_{i,i} &= \int_{-R}^{R} b_i'(x) b_i dx \\
&= \frac{1}{h^2}[\int_{x_{i-1}}^{x_i} (x - x_{i-1}) dx - \int_{x_i}^{x_{i+1}} (x_{i+1} - x) dx] \\
&= \frac{1}{h^2}(\frac{h^2}{2} - \frac{h^2}{2}) = 0 \\
\mathbf{C}_{i,i+1} &= \int_{-R}^{R} b_{i+1}'(x) b_i dx = \frac{1}{h^2}\int_{x_i}^{x_{i+1}} (x_{i+1} - x) dx = \frac{1}{2} \\
\mathbf{C}_{i,i-1} &= \int_{-R}^{R} b_{i-1}'(x) \cdot b_i dx = -\frac{1}{h^2}\int_{x_{i-1}}^{x_i} (x - x_{i-1}) dx = -\frac{1}{2}
\end{aligned}
$$

So we find

$$
\begin{aligned}
\mathbf{A}_{i,i} &= \frac{\sigma^2}{h} + (r+p)\frac{2h}{3} \\
\mathbf{A}_{i,i-1} &= \frac{-\sigma^2}{2h} - \frac{1}{2}(\frac{\sigma^2}{2} + q - r - p\eta) + (r+p)\frac{h}{6} \\
\mathbf{A}_{i,i+1} &= \frac{-\sigma^2}{2h} + \frac{1}{2}(\frac{\sigma^2}{2} + q - r - p\eta) + (r+p)\frac{h}{6}
\end{aligned}
$$

C Calculation of Load vector in one dimension

We recall that we have $\underline{f_i^m} = (b_i, f^m)$ and $f(x,\tau) = p \cdot \max(\kappa \exp(x)(1-\eta), R_{rec}\, F)$.

$$(f, b_i) = \int_{x_{i-1}}^{x_{i+1}} f(x,\tau) b_i(x) dx = \frac{1}{h} \int_{x_{i-1}}^{x_i} (p \max(\kappa e^x (1-\eta), R_{rec}\, F))(x - x_{i-1}) dx$$
$$+ \frac{1}{h} \int_{x_i}^{x_{i+1}} (p \max(\kappa e^x (1-\eta), R_{rec}\, F))(x_{i+1} - x) dx].$$

The value of this integral depends on the value of $\omega(x) := \max(\kappa e^x (1-\eta), R_{rec} \cdot F)$. To implement this integral in an efficient way, we have to check if this function even changes value on the chosen x interval $(-R, R)$.

We assume $R_{rec} \neq 0$, $\eta \neq 1$ and $\kappa \neq 0$ in the following and distinguish three cases

(i) $\log(\frac{R_{rec}F}{\kappa(1-\eta)}) \leq x_0 = -R \quad \Rightarrow \quad \omega(x) = \kappa e^x (1-\eta), \quad \forall x \in (-R; R).$

(ii) $R = x_{N+1} \leq \log(\frac{R_{rec}F}{\kappa(1-\eta)}) \quad \Rightarrow \quad \omega(x) = R_{rec}\, F, \quad \forall x \in (-R; R).$

(iii) $\exists j \in \{1, ..., N+1\} s.t. \quad x_{j-1} < \log(\frac{R_{rec}F}{\kappa(1-\eta)}) < x_j$
$\Rightarrow \quad \omega(x) = \kappa e^x (1-\eta) 1_{x > \log(\frac{R_{rec}F}{\kappa(1-\eta)})} + R_{rec} F 1_{x \leq \log(\frac{R_{rec}F}{\kappa(1-\eta)})}$

Let us calculate the load vector in these cases. Using integration by parts and assuming the price step is constant $h_i = h$, we find:

(i)

$$(f, b_i) = \frac{p\kappa(1-\eta)}{h} \int_{x_{i-1}}^{x_i} e^x (x - x_{i-1}) dx + \frac{p\kappa(1-\eta)}{h} \int_{x_i}^{x_{i+1}} e^x (x_{i+1} - x_i) dx$$
$$= \frac{p\kappa(1-\eta)}{h} e^{x_i} (e^{-h} - 2 + e^h),$$

(ii)

$$(f, b_i) = \frac{1}{h} [\int_{x_{i-1}}^{x_i} (p R_{rec} \cdot F)(x - x_{i-1}) dx + \int_{x_i}^{x_{i+1}} (p R_{rec} F)(x_{i+1} - x) dx]$$
$$= h p R_{rec} F$$

(iii) for some $j \in \{1, ..., N+1\}$:
If $x_{i-1} < \log(\frac{R_{rec}F}{\kappa(1-\eta)}) < x_i$ (for calculating (f, b_i)):

$$
\begin{aligned}
&\frac{1}{h}\int_{x_{i-1}}^{x_i} (p\max(\kappa e^x(1-\eta), R_{rec}\,F))(x - x_{i-1})dx \\
&= \frac{1}{h}[\int_{x_{i-1}}^{\log(\frac{R_{rec}F}{\kappa(1-\eta)})} (pR_{rec}F)(x - x_{i-1})dx \\
&\quad + \int_{\log(\frac{R_{rec}F}{\kappa(1-\eta)})}^{x_i} (p\kappa e^x(1-\eta))(x - x_{i-1})dx] \\
&= \frac{1}{2h}(pR_{rec}F)(\log(\frac{R_{rec}F}{\kappa(1-\eta)}) - x_{i-1})^2 \\
&\quad + \frac{1}{h}[p\kappa(1-\eta)\int_{\log(\frac{R_{rec}F}{\kappa(1-\eta)})}^{x_i} e^x(x - x_{i-1})dx] \\
&= \frac{1}{2h}(pR_{rec}F)(\log(\frac{R_{rec}F}{\kappa(1-\eta)}) - x_{i-1})^2 \\
&\quad + \frac{1}{h}p\kappa(1-\eta)([xe^x]_{\log(\frac{R_{rec}F}{\kappa(1-\eta)})}^{x_i} - [e^x]_{\log(\frac{R_{rec}F}{\kappa(1-\eta)})}^{x_i} - x_{i-1}[e^x]_{\log(\frac{R_{rec}F}{\kappa(1-\eta)})}^{x_i}) \\
&= \frac{1}{2h}pR_{rec}F(\log(\frac{R_{rec}F}{\kappa(1-\eta)}) - x_{i-1})^2 \\
&\quad + \frac{1}{h}p\kappa(1-\eta)[x_ie^{x_i} - \frac{R_{rec}F}{\kappa(1-\eta)}\log(\frac{R_{rec}F}{\kappa(1-\eta)}) - e^{x_i} + \frac{R_{rec}F}{\kappa(1-\eta)} - x_{i-1}e^{x_i} + x_{i-1}\frac{R_{rec}F}{\kappa(1-\eta)}]
\end{aligned}
$$

Using the calculations we have done before, we have

$$
\frac{1}{h}\int_{x_i}^{x_{i+1}} (p\max(\kappa e^x(1-\eta), R_{rec}\,F))(x_{i+1} - x)dx = \frac{p\kappa(1-\eta)}{h}e^{x_i}[-h + e^h - 1]
$$

Therefore, we find:

$$
\Rightarrow \left|
\begin{array}{l}
\text{If } \; x_{i-1} < \log(\frac{R_{rec}F}{\kappa(1-\eta)}) < x_i \\
\qquad \text{then} \\
(f, b_i) = \frac{1}{2h}pR_{rec}F(\log(\frac{R_{rec}F}{\kappa(1-\eta)}) - x_{i-1})^2 + \\
\qquad\qquad \frac{1}{h}p\kappa(1-\eta)[\frac{R_{rec}F}{\kappa(1-\eta)}(-\log(\frac{R_{rec}F}{\kappa(1-\eta)}) + 1 + x_i - h) + e^{x_i}(-2 + e^h)]
\end{array}
\right.
$$

If $x_i < \log(\frac{R_{rec}F}{\kappa(1-\eta)}) < x_{i+1}$ (for calculating (f, b_i)):

$$\frac{1}{h}\int_{x_{i-1}}^{x_i}(p\max(\kappa e^x(1-\eta), R_{rec}\cdot F))(x - x_{i-1})dx = \frac{h}{2}(pR_{rec}F)$$

And

$$\begin{aligned}
&\frac{1}{h}\int_{x_i}^{x_{i+1}}(p\max(\kappa e^x(1-\eta), R_{rec}\cdot F))(x_{i+1} - x)dx \\
&= \frac{1}{h}[\int_{x_i}^{\log(\frac{R_{rec}F}{\kappa(1-\eta)})}(pR_{rec}F)(x_{i+1} - x)dx + \int_{\log(\frac{R_{rec}F}{\kappa(1-\eta)})}^{x_{i+1}}(p\kappa e^x(1-\eta))(x_{i+1} - x)dx] \\
&= \frac{1}{2h}pR_{rec}F(h^2 - [x_{i+1} - \log(\frac{R_{rec}F}{\kappa(1-\eta)})]^2) \\
&\quad + \frac{1}{h}p\kappa(1-\eta)([e^x(x_{i+1} - x)]_{\log(\frac{R_{rec}F}{\kappa(1-\eta)})}^{x_{i+1}} + [e^x]_{\log(\frac{R_{rec}F}{\kappa(1-\eta)})}^{x_{i+1}}) \\
&= \frac{1}{2h}pR_{rec}F(h^2 - [x_{i+1} - \log(\frac{R_{rec}F}{\kappa(1-\eta)})]^2) \\
&\quad + \frac{1}{h}p\kappa(1-\eta)[\frac{R_{rec}F}{\kappa(1-\eta)}(\log(\frac{R_{rec}F}{\kappa(1-\eta)}) - x_{i+1}) + e^{x_{i+1}} - \frac{R_{rec}F}{\kappa(1-\eta)}]
\end{aligned}$$

Therefore, we find:

$$\begin{aligned}
(f, b_i) = hpR_{rec}F &- \frac{1}{2h}pR_{rec}F[x_{i+1} - \log(\frac{R_{rec}F}{\kappa(1-\eta)})]^2 \\
&+ \frac{1}{h}p\kappa(1-\eta)[\frac{R_{rec}F}{\kappa(1-\eta)}(\log(\frac{R_{rec}F}{\kappa(1-\eta)}) - x_{i+1}) + e^{x_{i+1}} - \frac{R_{rec}F}{\kappa(1-\eta)}]
\end{aligned}$$

$$\Rightarrow \left|\begin{array}{l}
\text{If } \ x_i < \log(\frac{R_{rec}F}{\kappa(1-\eta)}) < x_{i+1} \\
\qquad \text{then} \\
(f, b_i) = hpR_{rec}F - \frac{1}{2h}pR_{rec}F[x_{i+1} - \log(\frac{R_{rec}F}{\kappa(1-\eta)})]^2 \\
\qquad\qquad + \frac{1}{h}p\kappa(1-\eta)[\frac{R_{rec}F}{\kappa(1-\eta)}(\log(\frac{R_{rec}F}{\kappa(1-\eta)}) - x_{i+1}) + e^{x_{i+1}} - \frac{R_{rec}F}{\kappa(1-\eta)}]
\end{array}\right.$$

D Calculation of matrices in two dimension

We calculate here the matrices that enter in the implementation of the two dimensional pricing problem for a convertible with constant recovery.

To ease calculations, we consider a constant log-price step h_x and a constant interest rate step h_y.

We need to calculate the terms of the following matrices: $\mathbf{M}_x, \mathbf{M}_y, \mathbf{M}_{y,1}, \mathbf{C}_x, \mathbf{C}_y, \mathbf{C}_{y,1}, \mathbf{S}_x, \mathbf{S}_y$ (equations (III.24) - (III.31)).

The calculation of the matrices $\mathbf{M}_x, \mathbf{M}_y, \mathbf{C}_x, \mathbf{C}_y, \mathbf{S}_x, \mathbf{S}_y$ is done in Appendix B (just set $h = h_x$ for the matrices $\mathbf{M}_x, \mathbf{C}_x$ and $\mathbf{S}_x$, as well as $h = h_y$ for the matrices $\mathbf{M}_y, \mathbf{C}_y$ and $\mathbf{S}_y$). For the calculation of $\mathbf{M}_{y,1}$ and $\mathbf{C}_{y,1}$, we have

- Matrix $M_{y,1}$:

  By definition, we have $(M_{y,1})_{i,j} = \int_{R_2}^{R_3} y\bar{b}_i(y)\bar{b}_j(y)dy$

$$
\begin{aligned}
(M_{y,1})_{i,i} &= \int_{y_{i-1}}^{y_i} y(\frac{y-y_{i-1}}{h_y})^2 dy + \int_{y_i}^{y_{i+1}} y(\frac{y_{i+1}-y}{h_y})^2 dy \\
&= \frac{1}{h_y^2}\int_{y_{i-1}}^{y_i}(y-y_{i-1})^3 dy + \frac{1}{h_y^2}y_{i-1}\int_{y_{i-1}}^{y_i}(y-y_{i-1})^2 dy - \frac{1}{h_y^2}\int_{y_i}^{y_{i+1}}(y_{i+1}-y)^3 dy \\
&\quad + \frac{1}{h_y^2}y_{i+1}\int_{y_i}^{y_{i+1}}(y_{i+1}-y)^2 dy = \frac{h_y^2}{4} + \frac{h_y}{3}y_{i-1} - \frac{h_y^2}{4} + \frac{h_y}{3}y_{i+1} = \frac{2}{3}h_y y_i
\end{aligned}
$$

$$
\begin{aligned}
(M_{y,1})_{i+1,i} &= \int_{y_i}^{y_{i+1}} y(\frac{y_{i+1}-y}{h_y})(\frac{y-y_i}{h_y})dy = \frac{1}{h_y^2}\int_{y_i}^{y_{i+1}}(y-y_i+y_i)(y_{i+1}-y_i+y_i-y)(y-y_i)dy \\
&= \frac{1}{h_y^2}\int_{y_i}^{y_{i+1}} h_y(y-y_i)^2 - (y-y_i)^3 + y_i h_y(y-y_i) - y_i(y-y_i)^2 dy \\
&= \frac{1}{h_y^2}(\frac{h_y^4}{3} - \frac{h_y^4}{4} + y_i\frac{h_y^3}{2} - y_i\frac{h_y^3}{3}) = \frac{h_y^2}{12} + y_i\frac{h_y}{6}.
\end{aligned}
$$

  We get from the above calculation:

  $(M_{y,1})_{i-1,i} = \int_{y_{i-1}}^{y_i} y(\frac{y-y_{i-1}}{h_y})(\frac{y_i-y}{h_y})dy = \frac{h_y^2}{12} + y_{i-1}\frac{h_y}{6}$.

- Matrix $C_{y,1}$:

  By definition, we have $(C_{y,1})_{i,j} = \int_{R_2}^{R_3} y\bar{b}'_j(y)\bar{b}_i(y)dy$

$$
\begin{aligned}
(C_{y,1})_{i,i} &= \int_{y_{i-1}}^{y_i} y(\frac{y-y_{i-1}}{h_y})\frac{1}{h_y}dy - \int_{y_i}^{y_{i+1}} y(\frac{y_{i+1}-y}{h_y})\frac{1}{h_y}dy \\
&= \frac{1}{h_y^2}\int_{y_{i-1}}^{y_i}(y-y_{i-1})^2dy + \frac{1}{h_y^2}y_{i-1}\int_{y_{i-1}}^{y_i}(y-y_{i-1})dy + \frac{1}{h_y^2}\int_{y_i}^{y_{i+1}}(y_{i+1}-y)^2dy \\
&\quad - \frac{1}{h_y^2}y_{i+1}\int_{y_i}^{y_{i+1}}(y_{i+1}-y)dy = \frac{h_y}{3} + \frac{1}{2}y_{i-1} + \frac{h_y}{3} - \frac{1}{2}y_{i+1} = -\frac{1}{3}h_y.
\end{aligned}
$$

$$
\begin{aligned}
(C_{y,1})_{i+1,i} &= \int_{y_i}^{y_{i+1}} y(\frac{y_{i+1}-y}{h_y})\frac{1}{h_y}dy = \frac{-1}{h_y^2}\int_{y_i}^{y_{i+1}}(y-y_{i+1}+y_{i+1})(y-y_{i+1})dy \\
&= \frac{-1}{h_y^2}\int_{y_i}^{y_{i+1}}(y-y_{i+1})^2 + y_{i+1}(y-y_{i+1})dy = \frac{-1}{h_y^2}(\frac{h_y^3}{3} - y_{i+1}\frac{h^2}{2}) = -\frac{h_y}{3} + \frac{1}{2}y_{i+1}.
\end{aligned}
$$

$$
\begin{aligned}
(C_{y,1})_{i-1,i} &= \int_{y_{i-1}}^{y_i} y\frac{1}{h_y}(\frac{y_i-y}{h_y})dy = -\frac{1}{h_y^2}\int_{y_{i-1}}^{y_i}(y-y_i+y_i)(y-y_i)dy \\
&= -\frac{1}{h_y^2}(\int_{y_{i-1}}^{y_i}(y-y_i)^2 + y_i(y-y_i)dy) = -\frac{h_y}{3} + \frac{1}{2}y_i.
\end{aligned}
$$