Master

Mathematical Finance

Master's Final Work

Dissertation

Forward Backward Stochastic Differential Equations and Pricing in Emission Markets

André Monteiro Bento

Supervisor

João Guerra

October 2022

Abstract

This thesis is concerned with the pricing of emission allowance certificates in carbon markets. We study a stochastic model based on an FBSDE system. The forward processes are the Demand for electricity, the fuel prices and the total greenhouse gases emissions. The backward process is the price of the certificate. We present the construction of the model that mimics the market mechanics and give the theoretical results for existence and uniqueness of solution. We then provide a numerical algorithm for the system and test it with different functions and scenarios. We end with the inclusion of a potential policy that interferes with the market, in particular, with the Demand process.

Keywords: Forward Backward SDE, Numerical Methods, Monte-Carlo Method, Environmental Finance, Emission Markets

Acknowledgements

I would like to thank my parents for their support and guidance throughout my entire academic career. I am also grateful to my friends and to my teachers over these years.

Regarding this thesis, I want to give a special thank you to my mom, who had to deal with my stubbornness every time I asked for her help. I would also like to thank my supervisor, Professor João Guerra, for his support and availability during this time.

Contents

1 Introduction

The technological advancements that humanity has achieved in the last century have propelled it to grow, both in number and quality of life. The developments in medicine made the life expectancy of the average person increase and, as a result, the world population has more than doubled over the last sixty years. In parallel, the need for technology on a daily basis implies an increase in the need for electricity. The combination of these factors led to a surge in the demand for electric power. This, in turn, caused a soar in the emission of greenhouse gases (GHGs).

Approaching the turn of the millennium, the Kyoto protocol was a treaty that envisioned the sustainability of the environment through the reduction of the emissions of such gases. In order to comply with this goal, the European Union created the EU Emissions Trading System (EU ETS). This is a system that lays on a *cap-and-trade* principle. This means that a limit is imposed on the emissions of the GHGs with the intent of reducing the limit over time. The "trade" comes in the form of Emission Allowances. These are certificates that represent the right to emit GHGs. In this thesis, we study the pricing mechanism and behaviour of the Emission Allowances, which can be seen as derivatives in that which is called the Carbon Market.

In our study, we follow the development of a model for the carbon market that consists in a Forward Backward Stochastic Differential Equation System. We then apply a numerical scheme and consider small modifications to the model. These modifications are based on recent political policies implemented in Spain, which affect the Demand process.

This thesis is structured in the following way: in the second chapter, we address the existing theory of Forward-Backward Stochastic Differential Equations and associated numerical techniques and schemes. We also give a brief overview on the work that has been developed for the Carbon Emissions Market. In chapter three, we describe the Carbon Emissions Market and the EU ETS. We then proceed to construct the equations that model the market. Finally, we provide a result of existence and uniqueness to the FBSDE system that arises. In the fourth chapter, we describe the numerical scheme and respective algorithm that was used in the numerical simulations. We also present the results we obtained using the algorithm on the equations proposed by the works we followed, as well as with modifications we suggest. Lastly, we summarize all the results, both theoretical and numerical.

2 Literature Review

2.1 Forward-Backward Stochastic Differential Equations

A Forward-Backward Stochastic Differential Equation (FBSDE) is a system of the type

$$\begin{cases} X_t = \zeta + \displaystyle\int_{t_0}^{t} \mu(s, X_s, Y_s, Z_s)\ ds + \int_{t_0}^{t} \sigma(s, X_s, Y_s, Z_s)\ dW_s \\ \\ Y_t = \xi + \displaystyle\int_{t}^{T} f(s, X_s, Y_s, Z_s)\ ds - \int_{t}^{T} g(s, Y_s, Z_s)\ dW_s, \end{cases}$$

where μ, σ, f and g are deterministic functions, (W_t) represents the standard Brownian Motion and ζ and ξ are random variables, the initial and terminal condition respectively. (X_t) is the forward component of the system, (Y_t) is the backward component and (Z_t) is a process that is necessary to guarantee the adaptedness of the solution, in particular of (Y_t).

A Backward Stochastic Differential Equation (BSDE) is a particular case where there is no forward process (X_t). BSDEs were first introduced in 1973 by Bismut [4], who treated the linear case in the context of optimal stochastic control. The general case was then studied by Pardoux and Peng [20] in 1990. The transition to FBSDEs arises in the so called Markovian set-up, which has the terminal condition as a function of the forward component: $\xi = \phi(X_T)$. In this case, it is also standard to have $g(t, y, z) = z$.

Peng [21] was also the first to relate FBSDEs with Partial Differential Equations (PDEs). This relation (through a Feynman-Kac type formula) has been widely used in the field as a way to prove the existence of solution using classic PDE results and techniques. It is also important because the two processes end up being related by

$$Y_t = u(t, X_t),$$

where u is called a decoupling field (see [19]) and solves the PDE associated to the FBSDE.

The earliest result of existence and uniqueness of solution to an FBSDE was given in 1993 by Antonelli [1], in the small-time duration framework. Later, Ma et al. [17] and Delarue [12] proved the existence and uniqueness in more general cases, for an

arbitrary time duration and when the diffusion coefficient (σ) is non-degenerate (i.e. $\sigma \neq 0$ for all of its components). In [17], the authors also present a method to obtain the solution explicitly.

These equations arise very naturally in Finance (see [16] for example). The most well-known example to be given is the European Option pricing problem. Generally, this problem is modelled with Itō diffusions for the price of the underlying and the pricing function of the option. The initial value of the underlying is known but, for the option, we can only define the terminal value as a function of the underlying's terminal value. Thus, the diffusion modelling the underlying is going forward in time, while the one modelling the option is going backward. This is an example of a decoupled FBSDE, since the price of the underlying is independent of the value of the option, and the functions μ_X and σ would not depend on Y_s and Z_s. Another example is given in [10] regarding a regime switching hypothesis in the Black-Karasinski model for the short rate.

The model considered in this thesis has two main features that come as obstacles regarding it's well-posedness and the existence and uniqueness of solution. The first, is that the diffusion coefficient on the forward equation is degenerate. We consider a four-dimensional forward component in which the last process does not depend on the Brownian Motion. The second, and most important, is that the terminal condition on the backward component is not a Lipschitz function. As in the European Option pricing problem, the backward terminal condition is a function of the terminal value of the forward diffusion. However, due to the nature of the emissions certificates, the terminal function becomes of the type $\phi(e) = \Pi\mathcal{X}_{[\Lambda,+\infty[}(e)$, which has a discontinuity at Λ. Both of these problems were treated by Carmona and Delarue in [5], where the authors used a mollifying argument to prove existence. However, the result can only be guaranteed for a relaxed terminal condition. We also mention another type of discontinuity which has been studied in [10]. In this paper, the authors consider possibly multiple, but finite, discontinuities on the drift coefficient of the forward equation (μ) with respect to the backward variable Y_s.

2.2 Numerical methods for FBSDEs

The PDE Approach

As we have mentioned in the previous section, any FBSDE problem has a corre-

spondent PDE formulation. The literature regarding numerical methods for PDEs is extensive and dates back to Euler. Thus, one of the methods used to solve FBSDEs numerically is to consider their PDE representation and then use already known techniques. Interestingly, Delarue and Menozzi [13] use the probabilistic representation to solve the PDE problem.

One of the disadvantages of dealing with the PDE representation is that it might require stricter assumptions on smoothness and other such regularity conditions. One of the advantages is that the numerical solution is more accurate and faster to achieve. However, as the number of dimensions increases, it becomes harder, if not impossible, to obtain a solution. In these cases, we need to go with the probabilistic approach: The Monte-Carlo Method.

The Monte-Carlo Approach

The Monte-Carlo Method consists in simulating a process multiple times and then taking the average of the results. For example, if we want to determine the value of a derivative with payoff function $\phi(S_T)$. We can simulate the trajectory of the underlying price say, 10000 times: $(\hat{S}_t^i),\ 1 \leq i \leq 10000$. We then consider that the expected value of the derivative's payoff is the average of the simulated trajectories evaluated with ϕ:

$$\mathbb{E}\left[\phi(S_T)\right] = \frac{1}{10000} \sum_{i=1}^{10000} \phi\left(\hat{S}_T^i\right).$$

This is essentially relying on brute force computer power (and on the Law of Large Numbers) to bypass the theoretical difficulties (or impossibilities) that one might encounter. As we have stated, higher dimensional PDE problems become harder to treat numerically, and thus, we turn to the Monte-Carlo solutions.

Regarding BSDEs, the most common problem is to estimate the process (Z_t). In [18] the authors propose a method for the case

$$Y_t = \xi + \int_t^T f(s, Y_s)\ ds - \int_t^T Z_s\ dW_s,$$

where the Brownian Motion is approximated by a symmetric random walk. In [24], Zhang develops a method that aims to balance both the regularity conditions and the convergence rate. The case considered allows for f to depend on (Z_t) as well but has $\xi = \phi(X)$, where X is a diffusion process (he also considers a decoupled FBSDE). We

mention also the compilation of methods for BSDEs surveyed in [11].

We now turn to the Markovian case of the FBSDEs.

In a decoupled system

$$\begin{cases} X_t = \xi + \displaystyle\int_{t_0}^{t} \mu(s, X_s)\ ds + \int_{t_0}^{t} \sigma(s, X_s)\ dW_s \\ \\ Y_t = \phi(X_T) + \displaystyle\int_{t}^{T} f(s, X_s, Y_s, Z_s)\ ds - \int_{t}^{T} Z_s\ dW_s, \end{cases}$$

the forward component is independent of the backward component. Hence, we could simply simulate (X_t) first and then work backwards in time, since $\phi(X_T)$ would be known as well as the values of (X_t) for $\mu_Y(s, X_s, Y_s, Z_s)$. This would allow a very direct (and easy) application of the Monte-Carlo Method, and the computational obstacles would compare to those of BSDEs. However, our problem is not decoupled.

With a coupled system like

$$\begin{cases} X_t = \xi + \displaystyle\int_{t_0}^{t} \mu(s, X_s, Y_s)\ ds + \int_{t_0}^{t} \sigma(s, X_s, Y_s)\ dW_s \\ \\ Y_t = \phi(X_T) + \displaystyle\int_{t}^{T} f(s, X_s, Y_s, Z_s)\ ds - \int_{t}^{T} Z_s\ dW_s, \end{cases}$$

a typical discretization would be

$$\begin{cases} X_0 = x_0 \\ X_{i+1} = X_i + \mu\left(t_i, X_i, Y_i\right) \Delta t_{i+1} + \sigma\left(t_i, X_i, Y_i\right) \Delta W_{i+1} \\ Y_n = \phi(X_n) \\ Z_i = \dfrac{1}{\Delta t_{i+1}} \mathbb{E}\left[Y_{i+1} \Delta W_{i+1} \middle| \mathcal{F}_{t_i} \right] \\ Y_i = \mathbb{E}\left[Y_{i+1} + f\left(t_i, X_i, Y_{i+1}, Z_i\right) \Delta t_{i+1} \middle| \mathcal{F}_{t_i} \right]. \end{cases}$$

However, because (X_t) is discretized forwardly in time and (Y_t) backwardly, the scheme is fully implicit and becomes difficult to implement. The method developed by Bender and Zhang in [2] is based in a Picard iteration type scheme, where the values of the coupled process used in the forward equation come from the previous iteration. Thus, in each iteration, we can simulate the forward process first and then work backwards in time to reach Y_0, like in a decoupled system.

Note that the process (Z_t) is never coupled in the forward SDE. This is essential to the convergence of the method (see [2]). We describe the method fully in chapter 4.

2.3 Carbon Emissions Market

The creation of the EU ETS in 2005 and, in particular, the entrance in its second phase, led to the growing interest in modelling the new market mechanisms. For example, Frikha and Lemaire ([14]) model Gas and Electricity spot prices through the empirical correlations between the two.

In Carmona et. al ([8]), the authors analyse several complex alternative schemes to model the carbon markets. They consider multiple products causing emissions in the production process, analyse the costs and profits of the firms and the consequences of the emissions cap in a probabilistic scenario. They also consider tax-based schemes.

The inclusion of FBSDEs in the models came shortly after. In [7, 6], Carmona, Coulon and Schwarz introduce a structural model (similar to the one we studied), for pricing electricity and related derivatives. Such derivatives include the emission certificates and spread options relating the price of electric power when produced either with natural gas or coal. Howison and Schwarz ([22]) propose a related model, without the inclusion of fuel prices, and focus on the PDE representation of the problem. They also include the case of a multi period scenario. We mention as well Schwarz's Ph.D. thesis, ([23]) which is entirely dedicated to this topic.

Chassagneux et. al ([9]) follow the previously mentioned works and apply a version of the already mentioned structural models to the UK Energy Market. In our study, we implemented their model and analyse potential modifications (see chapter 4).

3 A stochastic model for the Carbon Emissions Market

3.1 Description of a Carbon Emissions Market

We start by describing the interactions between the electricity and emission markets. We give Portugal as the example.

In Portugal, the electricity market is open to suppliers, and these can make their offers freely. Consumers may also change their electricity supplier as they please. An electricity supplier may not be an electricity producer. The supplier buys electricity from the producer and then sells it to the consumer. We focus on the producers. The most commonly used method to produce electricity is the burning of fossil fuels. This method causes the release of CO_2 and other greenhouse gases. In view of the Kyoto protocol, to reduce the emissions of CO_2, the European Union (EU) created the EU Emissions Trading System (EU ETS).

The EU ETS is a market mechanism based on a *cap-and-trade* scheme. This means that the EU sets a limit on total CO_2 emissions by allocating a limited number of Emission Allowance Certificates per year. One allowance certificate gives the producers the right to emit one tonne of CO_2 equivalent greenhouse gases (we henceforth refer only to CO_2). At the end of the year (the compliance period), electricity producers must present enough certificates to fully cover their emissions throughout the period. If they have certificates to spare, they can sell them to other producers or save them for the following year. In case of excess emissions (non-compliance), they must pay a fine.

The limit on the total number of allowance certificates ensures that they have a monetary value. The trading of the certificates between producers (or other emitting entities) is called the emissions market (or carbon market). We note that in section 1 of [9], the authors argue that, under perfect information, imposing a tax on carbon emissions is in fact equivalent to the carbon market. In this section, we present a structural model for a simplified version of the electricity and emission markets which aims to determine the fair price of an allowance certificate.

In this thesis, we consider the following structure for the electricity market:

- The market is open to both suppliers and consumers and, for the sake of sim-

plicity, we consider that producers are also suppliers;

- There exists a market administrator who is in charge of matching supply and demand. This implies that it is not the consumers who choose their suppliers, but instead it considers that each consumer would choose their optimal supplier;
- The electricity suppliers submit their bids to the market administrator: a bid consists in a pair of quantity (of electricity) and price (asked for that amount). Note that a supplier can submit several bids, for example, they might be willing to supply the first 100 megawatts for 50€ but the next 100 megawatts for 60€;
- The market administrator is tasked with organizing the supply bids in merit order, that is, he must match demand to the lowest possible current supply bid. This results in the so-called bid stack;
- There exists a capacity limit for electricity production. Additionally, if demand were to surpass supply, the maximum limit would be supplied. So, we also assume that demand is bounded in the same way.

Also, for the sake of simplicity, we ignore the existence of alternative electricity production methods, both nuclear and environmentally friendly.

Regarding the emissions market, we have the following assumptions:

- All CO_2 emissions which are subjected to the cap are a consequence of the production of one good only, this being electricity;
- There exists a market regulator who imposes a limit on CO_2 emissions. This limit is imposed through emission allowances which are allocated at the beginning of the period (we ignore whether they are auctioned, sold or simply distributed);
- Emission allowances are tradable assets.

Additionally, we choose to work within a single compliance period, so that allowances lose its transition value over multiple periods.

3.2 Construction of the Model

We follow [9, 22] concerning the construction of the model for the emissions market. The purpose of the model is to determine the current price of an emissions allowance when we are given the current levels of demand for electricity, the price of the used fuels and the total cumulative emissions of CO_2 for the period.

We start by addressing the framework regarding our probability space. As opposed to starting to model the market under an empirical probability measure, say $\mathbb{P}$, we choose to work directly under the risk neutral measure $\mathbb{Q}$. This will be useful to define the equation governing the dynamics of the price of an allowance certificate. Thus, we have the following assumption:

Assumption 1. *There exists a probability measure, denoted by $\mathbb{Q}$, under which the discounted price of any tradable asset is a martingale.*

As stated, we choose to work solely with one compliance period, represented by the interval $[t_0, T]$. We consider a probability space $(\Omega, \mathcal{F}, \mathbb{Q})$ along with a natural filtration $(\mathcal{F}_t)_{t_0 \leq t \leq T}$ generated by a standard $\mathbb{Q}$-Brownian Motion $(W_t)_{t_0 \leq t \leq T}$ and augmented by the $\mathbb{Q}$-null sets (we consider (W_t) to be $n+1$ dimensional, with W_t^0 representing its first component and W_t^n the other n).

We denote by (D_t) and (ξ_t) the processes representing the demand and supply of electricity, measured in megawatts (henceforth, we also use the letter ξ to represent electricity production). We assume that there exists a capacity limit for electricity production and, by the argument above, there is a positive constant ξ_{max} such that

$$0 \leq D_t, \xi_t \leq \xi_{max}, \ t_0 \leq t \leq T.$$

As long as $D_t < \xi_{max}$, there is no need to produce excess electricity. So, as in [22], we consider that (D_t) and (ξ_t) are related through a Walrasian equilibrium imposed by the market administrator, implying

$$D_t = \xi_t, \ t \ a.s..$$

To characterize (D_t), we first assume that the demand for electricity is perfectly inelastic, that is, it does not depend on the price of electricity. This is a reasonable assumption, given that there are no simple substitute goods. Also, in [3], the authors found that the elasticity levels between demand and price are low. In the scope of the model, we admit the following Itō dynamics

$$dD_t = \mu_D(t, D_t)dt + \sigma_D(t, D_t)dW_t^0.$$

We denote by $(S_t) = (S_t^1, \ldots, S_t^n)$, the process representing the prices of the fuels consumed in electricity production. Similarly to D_t, we assume that S_t follows the Itō

dynamics

$$dS_t = \mu_S(t, S_t)dt + \sigma_S(t, S_t)dW_t^n.$$

We denote by (E_t) the process representing the total emissions of CO_2 up to time t, measured in tonnes of CO_2. Since the production level is bounded, we also have

$$0 \leq E_t \leq E_{max},\ t_0 \leq t \leq T,\ E_{max} > 0.$$

Also, we have the limit on total emissions, $E_{cap} > 0$, for the duration $[t_0, T]$ imposed by the market regulator, satisfying

$$0 \leq E_{cap} \leq E_{max}.$$

We construct (E_t) in the next sections.

Finally, we denote by (A_t) the process representing the price of an allowance certificate. The main characteristic about (A_t), is that its initial value is unknown. However, its terminal value can be deduced in the following way:

> *At time T, any excess emissions by a firm, say K, will be penalized at the amount* Π *per tonne. If another firm holds unused allowances, for positive* ε*, any price of* $\Pi - \varepsilon$ *will be favourable for firm K. Hence, under a no arbitrage hypothesis, we make* ε *tend to 0, and the fair price of an allowance should be* Π *at time T. In case there are no excess emissions, there will be no demand for the certificates, meaning that their value is 0.*

In other words, any firm that emits more than it can cover through allowances, should suffer the full penalty. So, we have

$$A_T = \begin{cases} 0 & if\ 0 \leq E_T < E_{cap} \\ \Pi & if\ E_{cap} \leq E_T \leq E_{max}. \end{cases}$$

Due to assumption (1), we know that the process $e^{-r(t-t_0)}A_t$ is a martingale (r stands for the risk free rate). By the martingale representation theorem, we know that there exists an adapted process (Z_t) such that

$$d[e^{-r(t-t_0)}A_t] = Z_t dW_t,$$

which, by the Itō formula, gives

$$dA_t = rA_t dt + e^{-r(t-t_0)} Z_t dW_t.$$

3.2.1 The bid and emissions stacks

To characterize the process (E_t), we need to model the bid and emissions stacks. We consider the superscript BAU (standing for $business - as - usual$) to represent a function which is not influenced by the cap on CO_2 emissions, that is, the version of a function before imposing restrictions on emissions.

As described in section 3.1, the market administrator arranges the production bids in increasing order (or merit order). This means that, for each level, the required energy is supplied at the cheapest possible price. So, the BAU bid stack is given by the function

$$b^{BAU}(\xi, s) : [0, \xi_{max}] \times \mathbb{R}^+ \to \mathbb{R}_0^+$$

where we have $b^{BAU} \in \mathcal{C}^1$ and $\dfrac{\partial b^{BAU}}{\partial \xi} > 0$. Note that, in theory, the bid stack should be an increasing step function. However, we assume that there are enough bids to approximate the real bid stack by a smooth function. An example bid stack function could be

$$b^{BAU}(\xi) = \underline{b} + \left(\frac{\overline{b} - \underline{b}}{\xi_{max}^{\theta_1}} \right) \xi^{\theta_1}, \quad \theta_1 > 2$$

as used in the test model of [22] (this paper does not consider fuel prices). Economically, we can interpret b^{BAU} as follows:

For the given level of electricity supply ξ and fuel prices $s = (s^1, \ldots, s^n)$,
$b^{BAU}(\xi, s)$ is the cost of production of the next unit of electricity.

Similarly, we can define the marginal emissions stack through the function

$$e : [0, \xi_{max}] \to \mathbb{R}^+,$$

with $e \in \mathcal{C}^1$. The test model of [22] used

$$e(\xi) = \underline{e} + \left(\frac{\overline{e} - \underline{e}}{\xi_{max}^{\theta_2}} \right) \xi^{\theta_2}, \quad 0 \leq \theta_2 < 1.$$

Analogously, we can read:

For the given level of electricity supply ξ, $e(\xi)$ represents the emissions necessary to produce the next unit of electricity.

Following this definition, the market emissions rate function is naturally defined as

$$\mu_E^{BAU}(D) := \kappa \int_0^D e(x)\ dx,\ 0 \leq D \leq \xi_{max},$$

that is:

For the given level of electricity demand D, $\mu_E^{BAU}(D)$ represents the total emissions necessary to produce $\xi = D$ units of electricity.

The parameter κ serves as a scale parameter.

3.2.2 The bid stack with capped emissions

We now consider the consequences of imposing a cap on total emissions. The main obstacle is to find a suitable integration set for the equivalent definition of μ_E^{BAU}. We start with the following definitions.

Definition 3.1. *For each fuel type $i = 1, \ldots, n$, the function*

$$b_i^{BAU}(\xi, s_i) : [0, \xi_{max}^i] \times \mathbb{R} \to \mathbb{R}_0^+$$

represents the bid stack if there were only the fuel of type i available. ξ_{max}^i represents the maximum amount of electricity that can be produced using fuel i. The function $s_i \mapsto b_i^{BAU}(\cdot, s_i)$ should be strictly increasing.

Definition 3.2. *For each fuel type $i = 1, \ldots, n$, the function*

$$e_i(\xi) : [0, \xi_{max}^i] \to \mathbb{R}^+$$

represents the marginal emissions stack when the fuel of type i is used to produce electricity.

For the sequel, we assume that e_i is non decreasing.

As we have mentioned before, a carbon market is equivalent to a carbon tax. This means that we can treat the emission allowances as an added production cost. By the definitions of b_i^{BAU} and e_i, it is straightforward to define the bid stack for each fuel as

$$b_i(\xi, s, A) = b_i^{BAU}(\xi, s) + Ae_i(\xi).$$

This way, the excess cost of production is proportional to emissions. One could reason that the natural bid stack would be the sum of all b_i's. However, this would leave us without a proper integration set for μ_E. Hence, we need the following arguments.

Note that each $b_i(\xi, \cdot, \cdot)$ is strictly increasing, so we can consider its generalized right-continuous inverse function

$$\begin{aligned} b_i^{-1}(p, s_i, A) \quad &: \quad \mathbb{R} \to [0, \xi_{max}^i] \\ &\qquad p \mapsto \xi_{max}^i \wedge \inf\{\xi \in [0, \xi_{max}^i] : b_i(\xi, s_i, A) > p\}. \end{aligned}$$

We also denote its minimum by

$$\underline{b_i} := b_i(0, s_i, A).$$

We can say that $b_i^{-1}(p, s_i, A)$ represents the total amount of electricity produced in the market using fuel type i (for a given electricity price p and allowance price A). Then, the total amount of electricity produced is given by

$$b^{-1}(p, s, A) = \sum_i b_i^{-1}(p, s_i, A).$$

By the monotonicity of the individual b_i's, we can also invert $p \mapsto b^{-1}(p, s, A)$. So, we define the market bid stack by

$$b(\xi, s, A) := \min_i \underline{b_i} \vee \sup \left\{ p \in \mathbb{R}^+ : \sum_i b_i^{-1}(p, s_i, A) < \xi \right\}.$$

Definition 3.3. *For each fuel type $i = 1, \ldots, n$, electricity price p and allowance price A, we define the set of potential active generators as*

$$G_i(p, s_i, A) := \{\xi \in [0, \xi_{max}^i] : b_i(\xi, s_i, A) \leq p\}, \ 0 \leq p, s_i, A < \infty.$$

That is, for given electricity, fuel and allowance prices (p, s_i, A), $G_i(p, s_i, A)$ represents the set of production levels that are still producible (using fuel i) without incurring in losses.

We have, finally, the definition for the market emissions rate when the market

regulator imposes a cap on emissions through emission allowances:

$$\mu_E(D, s, A) := \kappa \sum_i \int_{G_i} e_i(x)\ dx,\ 0 \leq A, s_i < \infty,\ 0 \leq D \leq \xi_{max}$$

where $G_i := G_i(b(D, s, A), s, A)$.

Finally, (E_t) represents the cumulative emissions over time, so we simply have

$$E_t = \int_0^t \mu_E(D_s, S_s, A_s)ds.$$

3.3 Existence and Uniqueness of Solution

We want to show that the model

$$\begin{cases} dD_t = \mu_D(t, D_t)dt + \sigma_D(t, D_t)dW_t^0 \\ dS_t = \mu_S(t, S_t)dt + \sigma_S(t, S_t)dW_t^n \\ dE_t = \mu_E(D_t, S_t, A_t)dt \\ dA_t = rA_t dt + e^{-r(t-t_0)}Z_t dW_t, \quad t_0 \leq t \leq T \\ (D_{t_0}, S_{t_0}, E_{t_0}) = (d, s, e) \\ A_T = \Pi \mathcal{X}_{[E_{cap}, +\infty[}(E_T) \end{cases} \tag{1}$$

admits a unique solution ($\mathcal{X}$ stands for the indicator function). However, we will have to settle for a relaxed terminal condition, which can be seen as an approximation.

In order to simplify the notation, because D_t and S_t are both non-degenerate forward components,we define:

$$X_t := \begin{bmatrix} D_t \\ S_t \end{bmatrix}, \quad \mu_X(t, X_t) := \begin{bmatrix} \mu_D(t, D_t) \\ \mu_S(t, S_t) \end{bmatrix}, \quad \sigma_X(t, X_t) := \begin{bmatrix} \sigma_D(t, D_t) & \mathbf{0} \\ \mathbf{0} & \sigma_S(t, S_t) \end{bmatrix}.$$

Following [5], under the appropriate conditions, a system of the type

$$\begin{cases} dX_t = \mu_X(X_t)dt + \sigma_X(X_t)dW_t \\ dE_t = \mu_E(X_t, A_t)dt \\ dA_t = Z_t dW_t, \quad t_0 \leq t \leq T \end{cases}$$

admits a unique solution for any initial condition $(x, e) \in \mathbb{R}^n \times \mathbb{R}$ and terminal

condition

$$\Pi\mathcal{X}_{]E_{cap},+\infty[}(E_T) \leq A_T \leq \Pi\mathcal{X}_{[E_{cap},+\infty[}(E_T).$$

Since the process A_t can be directly obtained from $Y_t = e^{-r(t-t_0)}A_t$, we can use the latter to solve an equivalent system by defining the following function:

$$f(x,y) := \mu_E\left(x, e^{r(t-t_0)}y\right),\ t_0 \leq t \leq T,$$

which leads to the simplified system:

$$\begin{cases} dX_t = \mu_X(t,X_t)dt + \sigma_X(t,X_t)dW_t, \\ dE_t = f(X_t,Y_t)dt \\ dY_t = Z_t dW_t, \quad t_0 \leq t \leq T. \end{cases} \tag{2}$$

So, we need to show that the result of existence and uniqueness still holds when the coefficients b and σ also depend on time. Hence, we work with the following assumption:

Assumption 2. *$\mu_X : [t_0,T] \times \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$, $\sigma_X : [t_0,T] \times \mathbb{R}^{n+1} \to \mathbb{R}^{(n+1)\times(n+1)}$, $f : \mathbb{R}^{n+1} \times \mathbb{R} \to \mathbb{R}$ and there exist three constants, $L \geq 1$, $l_1, l_2 > 0$, $1/L \leq l_1 \leq l_2 \leq L$, such that*

(i) μ_X and σ_X are L-Lipschitz continuous and of at most L-linear growth in the spatial variable for almost every $t \in [t_0,T]$, that is,

$$|\mu_X(t,x_1) - \mu_X(t,x_2)| + |\sigma_X(t,x_1) - \sigma_X(t,x_2)| \leq L|x_1 - x_2|,\ x_1,x_2 \in \mathbb{R}^{n+1},\ t\ a.s.,$$

$$|\mu_X(t,x)| + |\sigma_X(t,x)| \leq L(1+|x|),\ x \in \mathbb{R}^{n+1},\ t\ a.s..$$

ii) for any $y \in \mathbb{R}$, $x \mapsto f(x,y)$ is L-Lipschitz continuous and satisfies

$$|f(x,y)| \leq L(1+|x|+|y|),\ x \in \mathbb{R}^{n+1}.$$

iii) For any $x \in \mathbb{R}^{n+1}$, $y \mapsto f(x,y)$ is strictly decreasing and satisfies

$$l_1|y_1 - y_2|^2 \leq (y_1 - y_2)[f(x,y_1) - f(x,y_2)] \leq l_2|y_1 - y_2|^2,\ y_1,y_2 \in \mathbb{R}.$$

Note that $y \mapsto f(x,y)$ is decreasing if $A \mapsto \mu_E(X,A)$ is decreasing, which is our

case.

We follow section 2 of [5]. We aim to show that the function $(t_0, x, e) \mapsto Y_{t_0}^{\varepsilon,t_0,x,e}$ still verifies the same PDE and a-priori estimates as in propositions 2.4, 2.6 and 2.8 of [5] (see Appendix). Hence, we start by considering a non-decreasing Lipschitz terminal condition ϕ bounded by 0 and 1. We also consider the non-degenerate mollified equation system:

$$\begin{cases} dX_t^\varepsilon = \mu_X(t, X_t^\varepsilon)dt + \sigma_X(t, X_t^\varepsilon)dW_t + \varepsilon d\tilde{W}_t \\ dE_t^\varepsilon = f(X_t^\varepsilon, Y_t^\varepsilon)dt + \varepsilon dB_t \\ dY_t^\varepsilon = Z_t^\varepsilon dW_t + \tilde{Z}_t^\varepsilon d\tilde{W}_t + \Upsilon_t^\varepsilon dB_t, \end{cases} \tag{3}$$

where we work in an appropriate augmented probability space. Here, $(\tilde{W}_t)_{t_0 \le t \le T}$ and $(B_t)_{t_0 \le t \le T}$ are two additional independent Brownian Motions of dimensions $n+1$ and 1 respectively and $0 < \varepsilon < 1$. Also, $(Z_t^\varepsilon, \tilde{Z}_t^\varepsilon, \Upsilon_t^\varepsilon)_{t_0 \le t \le T}$ are the processes arising from the martingale representation of $(Y_t^\varepsilon)_{t_0 \le t \le T}$ with respect to $(W_t, \tilde{W}_t, B_t)$.

Section 2 in [12] allows us to suppose that the functions μ_D, σ_D, f and ϕ are $\mathcal{C}^\infty$ with bounded derivatives, as then we can construct sequences of functions that converge to the real coefficients. By [12], for any initial condition (t_0, x, e) and terminal condition $Y_T = \phi(E_T)$, the system (3) admits a unique solution which we denote by

$$(X_t^{\varepsilon,t_0,x,e}, E_t^{\varepsilon,t_0,x,e}, Y_t^{\varepsilon,t_0,x,e}, \widehat{Z}_t^{\varepsilon,t_0,x,e})_{t_0 \le t \le T},$$

with $\widehat{Z}_t^{\varepsilon,t_0,x,e} = (Z_t^{\varepsilon,t_0,x,e}, \tilde{Z}_t^{\varepsilon,t_0,x,e}, \Upsilon_t^{\varepsilon,t_0,x,e})_{t_0 \le t \le T}$. We then define the function

$$\theta^\varepsilon(t_0, x, e) := Y_{t_0}^{\varepsilon,t_0,x,e}. \tag{4}$$

With $\mathcal{L}_x(t) := \langle \mu_X(t,x), \frac{\partial}{\partial x} \rangle + \frac{1}{2}\mathrm{Tr}\left((\sigma_X \sigma_X^*) \circ (t,x) \frac{\partial^2}{\partial x^2}\right)$, we have the following properties:

Proposition 3.4. *The function θ^ε takes values in $[0,1]$ only and is of class $\mathcal{C}^{1,2,2}$ on $[0,T] \times \mathbb{R}^{n+1} \times \mathbb{R}$ with bounded derivatives. Moreover, it satisfies the following PDE:*

$$\left[\frac{\partial \theta^\varepsilon}{\partial t} + \mathcal{L}_x(t)\theta^\varepsilon + \frac{\varepsilon^2}{2}\left(\Delta_{xx}\theta^\varepsilon + \frac{\partial^2 \theta^\varepsilon}{\partial e^2}\right)\right] \circ (t,x,e) + f(x, \theta^\varepsilon(t,x,e)) \frac{\partial \theta^\varepsilon}{\partial e}(t,x,e) = 0 \tag{5}$$

with terminal condition $\theta^\varepsilon(T,x,e) = \phi(e)$ for all $x \in \mathbb{R}^{n+1}$.

Proof. θ^ε being of class $\mathcal{C}^{1,2,2}$ with bounded derivatives and solving the PDE are immediate from Lemma 2.1 in [12] by noticing that the Brownian Motion coefficient is given by the $(n+2)\mathsf{x}(2n+3)$ matrix

$$\Sigma := \begin{bmatrix} \sigma_X & \varepsilon I_{n+1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \varepsilon \end{bmatrix}$$

and thus, the matrix $a := \Sigma\Sigma^*$ is obtained by augmenting $\sigma_X\sigma_X^*$ with a column and a row of zeros and adding $\varepsilon^2 I_{n+2}$. (I_p standing for the identity matrix of order p).

□

We now present the desired result.

Theorem 3.5. *Under Assumption* (2)*, for any initial condition* $(x, e) \in \mathbb{R}^{n+1}$ x $\mathbb{R}$*, problem* (2) *admits a unique solution* $(X_t, E_t, Y_t, Z_t)_{t_0 \le t \le T}$ *satisfying the terminal condition*

$$\mathbb{Q}\{e^{-r(T-t_0)}\Pi\mathcal{X}_{]E_{cap},+\infty[}(E_T) \le Y_T \le e^{-r(T-t_0)}\Pi\mathcal{X}_{[E_{cap},+\infty[}(E_T)\} = 1$$

and the integrability condition

$$\mathbb{E}\left[\sup_{t_0 \le t \le T}\left\{|X_t|^2 + |E_t|^2 + |Y_t|^2\right\} + \int_0^T |Z_t|^2 dt\right] < +\infty.$$

Consequently, problem (1) *admits a unique solution when the terminal condition is relaxed to*

$$\mathbb{Q}\{\Pi\mathcal{X}_{]E_{cap},+\infty[}(E_T) \le A_T \le \Pi\mathcal{X}_{[E_{cap},+\infty[}(E_T)\} = 1.$$

Additionally, there exists a continuous function $\theta = \lim_{\varepsilon\to 0}\theta^\varepsilon$ *which verifies*

$$A_t = \theta(t, D_t, S_t, E_t).$$

Proof. The proof is in all aspects similar to the proofs in section 2 of [5]. Assumption 2.i) ensures that the processes remain well-defined and that no problems of integrability arise. Additionally, proposition 4.1 in [16] confirms that all processes remain differentiable as needed. Thus, θ^ε verifies the same a-priori estimates and the result holds true. □

4 Numerical simulation of the model

In this section we present the numerical scheme and algorithm proposed in [2]. We then discuss our results obtained with its application to a test model suggested by [9] and to the model described in the previous chapter.

4.1 Numerical Scheme

Consider the FBSDE

$$\begin{cases} X_t = x_0 + \int_{t_0}^{t} \mu_X(s, X_s, Y_s)\ ds + \int_{t_0}^{t} \sigma(s, X_s, Y_s)\ dW_s \\ \\ Y_t = \phi(X_T) + \int_{t}^{T} \mu_Y(s, X_s, Y_s, Z_s)\ ds - \int_{t}^{T} Z_s\ dW_s \end{cases} \tag{6}$$

with the following discretization scheme

$$\begin{cases} X_0 = x_0 \\ X_{i+1} = X_i + \mu\left(t_i, X_i, Y_i\right)\Delta t_{i+1} + \sigma\left(t_i, X_i, Y_i\right)\Delta W_{i+1} \\ Y_n = \phi(X_n) \\ Z_i = \dfrac{1}{\Delta t_{i+1}}\mathbb{E}\left[Y_{i+1}\Delta W_{i+1} \middle| \mathcal{F}_{t_i}\right] \\ Y_i = \mathbb{E}\left[Y_{i+1} + f\left(t_i, X_i, Y_{i+1}, Z_i\right)\Delta t_{i+1} \middle| \mathcal{F}_{t_i}\right]. \end{cases} \tag{7}$$

The problem of applying (7) to (6) is that the scheme is fully implicit. Note that Y_i depends on Y_{i+1} which depends on X_{i+1} which in turn depends on Y_i. This turns (7) computationally unfeasible given the dimensionality of the problem we are modelling. We thus need to find a way to decouple the system.

The idea is the following: in many cases, the method used to prove existence and uniqueness of solution to an equation is by using a fixed point theorem on a contraction. In such cases, an iteration scheme would converge to the said fixed point. The method developed in [2] decouples the FBSDE using the iteration argument.

We start with an initial trajectory of the backward process (Y_t) which stands as an initial guess of the solution. Then, we use this trajectory to simulate the forward process (X_t). The system is now decoupled, and we can use the values of (X_t) to compute the new trajectory of (Y_t). We then repeat the same process until the scheme converges.

The discretization scheme proposed in [2] is

$$\begin{cases} X_0^m = x_0 \\ X_{i+1}^{\Pi,m} = X_i^{\Pi,m} + \mu_X\left(t_i, X_i^{\Pi,m}, u_i^{\Pi,m-1}(X_i^{\Pi,m})\right)\Delta t_{i+1} \\ \qquad\qquad\qquad + \sigma\left(t_i, X_i^{\Pi,m}, u_i^{\Pi,m-1}(X_i^{\Pi,m})\right)\Delta W_{i+1} \\ Y_n^{\Pi,m} = \phi(X_n^{\Pi,m}) \\ \\ Z_i^{\Pi,m} = \dfrac{1}{\Delta t_{i+1}}\mathbb{E}\left[Y_{i+1}^{\Pi,m}\Delta W_{i+1}\middle|\mathcal{F}_{t_i}\right] \\ \\ Y_i^{\Pi,m}\mathbb{E}\left[Y_{i+1}^{\Pi,m} + \mu_Y\left(t_i, X_i^{\Pi,m}, Y_{i+1}^{\Pi,m}, Z_i^{\Pi,m}\right)\Delta t_{i+1}\middle|\mathcal{F}_{t_i}\right] \\ \\ u_i^{\Pi,m} = Y_i^{\Pi,m} \end{cases}$$

where Π stands for the time grid. The main idea is to use a Picard iteration method to reach an approximate solution $(X^{\Pi,M}, Y^{\Pi,M}, Z^{\Pi,M})$. $(X^{\Pi,m})$ is calculated using $(Y^{\Pi,m-1}, Z^{\Pi,m-1})$ and $(Y^{\Pi,m}, Z^{\Pi,m})$ using $(X^{\Pi,m})$. The usual assumptions stand: μ_X, σ, μ_Y, ϕ are Lipschitz continuous functions of linear growth. Also, it is assumed that the drift and diffusion coefficients are uniformly Hölder-$\frac{1}{2}$ continuous w.r.t. t.

We use the following notation regarding the processes of the model:

$$X_t = (D_t, S_t, E_t).$$

We also note that the proposed method admits the degeneracy of the σ coefficient included in our model.

Remark 4.1. *The discretization for $Z_i^{\Pi,m}$ is not immediate, but comes from the following idea:*

$$\begin{aligned} Y_{t_i} &\approx Y_{t_{i+1}} + \mu_Y\left(t_i, X_{t_i}, Y_{t_i}, Z_{t_i}\right)\Delta t_{i+1} - Z_{t_i}\Delta W_{i+1} \\ &\Leftrightarrow \\ Z_{t_i}\Delta W_{i+1} &\approx Y_{t_{i+1}} - Y_{t_i} + \mu_Y\left(t_i, X_{t_i}, Y_{t_i}, Z_{t_i}\right)\Delta t_{i+1}. \end{aligned}$$

Multiplying both sides by ΔW_{i+1} on the right and applying the conditional expectation

on $\mathcal{F}_{t_i}$ results in

$$Z_{t_i}\Delta t_{i+1} \approx \mathbb{E}\big[Y_{t_{i+1}}\Delta W_{t_{i+1}}|\mathcal{F}_{t_i}\big],$$

given that the rest is $\mathcal{F}_{t_i}$ measurable and thus, the expectation is simply on the Brownian increment which is null.

4.2 Numerical Algorithm

We now present the algorithm to employ the scheme described above. In [2, 15], the authors' suggestion is to calculate the conditional expectations using a least squares regression estimator, for which we need to simulate Λ different paths for each iteration. The algorithm is the following:

- Fix Λ and the number of time steps n and define a time grid $\Pi = \{t_0 < t_1 < \cdots < t_n = T\}$ (we henceforth omit the dependence of the simulation regarding n, Π and Λ).
- Set $u_i^0(x) \equiv 0$ for all $0 \leq i \leq n$ and choose x_0.
- Simulate Λ independent Brownian Motion trajectories indexed by (W_t^λ), $1 \leq \lambda \leq \Lambda$.
- For $m = 1, \ldots, M$ do:
  1. Simulate Λ trajectories of X^m as follows:

$$\begin{aligned} X_0^{m,\lambda} &= x_0 \\ X_{i+1}^{m,\lambda} &= X_i^{m,\lambda} + \mu_X\left(t_i, X_i^{m,\lambda}, u_i^{m-1}(X_i^{m,\lambda})\right)\Delta t_{i+1} \\ &\quad + \sigma_X\left(t_i, X_i^{m,\lambda}, u_i^{m-1}(X_i^{m,\lambda})\right)\Delta W_{i+1}^\lambda \end{aligned}$$

  2. Choose a set of Lipschitz continuous basis functions

$$\mathcal{B}_i^K = \{p_{k,i}(x),\ 1 \leq k \leq K\}$$

     such that

$$\Big\{p_{k,i}(X_i^{m,\lambda}),\ 1 \leq k \leq K\Big\}$$

     forms a subset of $L^2(\Omega)$.

  3. Fix $u_n^m(\cdot) = \phi(\cdot)$.

4. For $i = n-1, \ldots, 1$ do:

$$
\begin{aligned}
&Y_{i+1}^{m,\lambda} = u_{i+1}^m(X_{i+1}^{m,\lambda}), \quad 1 \le \lambda \le \Lambda \\
&v_i^m(x) = \arg\inf \left\{ \frac{1}{\Lambda} \sum_{\lambda=1}^{\Lambda} \left| \frac{1}{\Delta t_{i+1}} Y_{i+1}^{m,\lambda} \Delta W_{i+1}^{\lambda} - V(X_{i+1}^{m,\lambda}) \right|^2 : \right. \\
&\qquad\qquad \left. V \in \operatorname{span}\left(\mathcal{B}_i^K\right) \right\} \\
&Z_i^{m,\lambda} = v_i^m(X_{i+1}^{m,\lambda}) \\
&u_i^m(x) = \arg\inf \left\{ \frac{1}{\Lambda} \sum_{\lambda=1}^{\Lambda} \left| Y_{i+1}^{m,\lambda} + \mu_Y \left(t_i, X_i^{\Pi,m}, Y_{i+1}^{\Pi,m}, Z_i^{\Pi,m}\right) \Delta t_{i+1} \right. \right. \\
&\qquad\qquad \left. \left. - U(X_i^{m,\lambda}) \right|^2 : U \in \operatorname{span}\left(\mathcal{B}_i^K\right) \right\}
\end{aligned}
$$

5. Set

$$
\begin{aligned}
Y_1^{m,\lambda} &= u_1^m(X_1^{m,\lambda}) \\
Z_0^{m,\lambda} &= \frac{1}{\Lambda \Delta t_1} \sum_{\bar{\lambda}=1}^{\Lambda} Y_1^{m,\bar{\lambda}} \Delta W_1^{\bar{\lambda}} \\
Y_0^{m,\lambda} &= \frac{1}{\Lambda} \sum_{\bar{\lambda}=1}^{\Lambda} \left[Y_1^{m,\bar{\lambda}} + \mu_Y \left(t_0, x_0, Y_1^{m,\bar{\lambda}}, Z_0^{m,\bar{\lambda}}\right) \Delta t_1 \right]
\end{aligned}
$$

Remark 4.2. *Note that with our model, steps 4 and 5 are reduced to*
For $i = n-1, \ldots, 0$ do:

$$
\begin{aligned}
&Y_{i+1}^{m,\lambda} = u_{i+1}^m(X_{i+1}^{m,\lambda}), \quad 1 \le \lambda \le \Lambda \\
&u_i^m(x) = \arg\inf \left\{ \frac{1}{\Lambda} \sum_{\lambda=1}^{\Lambda} \left| Y_{i+1}^{m,\lambda} - U(X_i^{m,\lambda}) \right|^2 : U \in \operatorname{span}\left(\mathcal{B}_i^K\right) \right\} \\
&Y_0^{m,\lambda} = \frac{1}{\Lambda} \sum_{\bar{\lambda}=1}^{\Lambda} Y_1^{m,\bar{\lambda}}
\end{aligned}
$$

since $\mu_Y \equiv 0$, which leads to $Z_i^{m,\lambda}$ becoming unnecessary in u_i^m.

As in [2, 9], for the numerical simulations, we used

$$\mathcal{B}_i^K = \mathcal{B}^D = \{1\} \cup \{x_d, 1 \le d \le D\} \cup \{(-R) \vee x_d x_q \wedge R, 1 \le d \le q \le D\}$$

where X takes values in $\mathbb{R}^D$.

4.3 Results

In this section we present the results of our numerical experimentations of the algorithm described above. We started by testing the algorithm with a mock system proposed in [9] of a simple singular FBSDE. We then tested the model of chapter 3 with the dynamics and parameters estimated in [9]. The algorithm was implemented using Python.

4.3.1 Simple singular FBSDE

Consider the following FBSDE

$$\begin{cases} dX_t = dW_t \\ dE_t = (X_t - Y_t)dt \\ dY_t = Z_t dW_t \\ (X_0, E_0) = (0, 0) \\ Y_1 = \mathcal{X}_{[0,+\infty[}(E_1). \end{cases}$$

Our implementations had the parameters $\Lambda = 10000$, $R = 10$ and 101 equidistant time-steps.

A direct application of the algorithm proved unsuccessful, as it did not converge (on 20 iterations). This was expected, given that the terminal function is not Lipschitz continuous. We thus considered the mollified terminal condition $Y_1 = \phi_\varepsilon(E_1)$, where

$$\phi_\varepsilon(x) = \begin{cases} 0 & \text{if } \ x \leq -\varepsilon \\ \frac{x+\varepsilon}{\varepsilon} & \text{if } \ -\varepsilon < x < 0 \\ 1 & \text{if } \ x \geq 0. \end{cases}$$

The idea was to start at some value for which the algorithm converged and then make $\varepsilon \to 0$. However, it did not converge for any of the following values: $\varepsilon = 5$, $\varepsilon = 1$, $\varepsilon = 0.5$ and $\varepsilon = 0.1$ (in 50 iterations). Following the methodology of [5] which recurred to system (3), we included a non-degenerate "small" noise term to dE_t. Hence, we

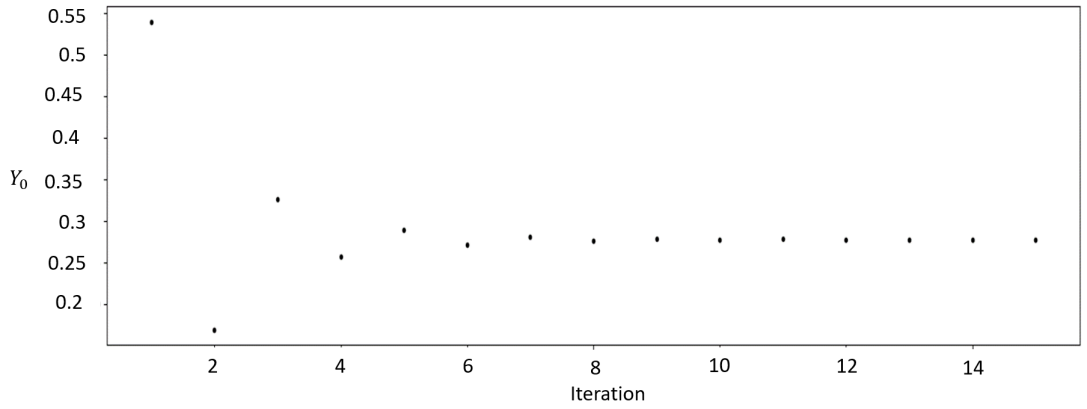

Figure 1: Y_0 for system (8) with parameters $\varepsilon = 0.1$ and $\delta = 0.01$.

have

$$\begin{cases} dX_t = dW_t \\ dE_t = (X_t - Y_t)dt + \delta dB_t \\ dY_t = Z_t dW_t + \tilde{Z}_t dB_t \\ (X_0, E_0) = (0, 0) \\ Y_1 = \phi_\varepsilon(E_1). \end{cases} \tag{8}$$

With $\varepsilon = 0.1$ and $\delta = 0.01$ the algorithm converged in less than 10 iterations (fig. 1). For $\delta = 0.005$ the result was the same but for $\delta = 0.001$ the algorithm no longer converged. Hence, we fixed $\delta = 0.005$ and tested lower values for ε. The result stayed unchanged for ε as low as 0.01. For $\varepsilon = 0.005$ the algorithm needed more iterations to converge (approximately 25) but the value of the solution dropped around 5%.

We thus conclude that, with a Lipschitz terminal condition, the algorithm fails with a degenerate diffusion coefficient. Also, the non-degenerate mollification cannot be arbitrarily small.

4.3.2 Simulation with estimated functions and parameters

Our approach to simulate the model developed in chapter 3 was to use the functions estimated in [9] for the UK Energy Market. We omit the process of estimation of the parameters which can be found in the book. The authors suggest the following

dynamics for the Demand and fuel price processes:

$$\begin{aligned} D_t &= \kappa^D \exp\left\{h^D(t) + X_t^D\right\} \\ S_t^C &= \kappa^C \exp\left\{h^C(t) + X_t^C\right\} \\ S_t^G &= \kappa^G \exp\left\{h^G(t) + X_t^G\right\} \end{aligned}$$

where the h^is represent the seasonality terms, (X_t^i) are diffusion processes and the κ^is are the scale parameters. The seasonality functions were given by

$$\begin{aligned} h^D(t) &= a^D + b^D t + c_1^D \cos\left(\frac{2\pi t}{63/252}\right) + d_1^D \sin\left(\frac{2\pi t}{63/252}\right) + \\ &\quad + c_2^D \cos(2\pi t) + d_2^D \sin(2\pi t) \end{aligned}$$

$$h^C(t) = a^C + b_1^C t + b_2^C t^2 + c^C \cos(2\pi t) + d^C \sin(2\pi t)$$

$$\begin{aligned} h^G(t) &= a^G + b_1^G t + b_2^G t^2 + b_3^G t^3 + c_1^G \cos\left(\frac{2\pi t}{63/252}\right) + \\ &\quad + d_1^G \sin\left(\frac{2\pi t}{63/252}\right) + c_2^G \cos(2\pi t) + d_2^G \sin(2\pi t). \end{aligned}$$

Thus, h^D and h^G present annual and quarterly tendencies while h^C only has annual components.

Regarding the diffusion processes, (X_t^D) and (X_t^C) were estimated as Ornstein-Uhlenbeck processes while (X_t^G) was assumed to have non-constant volatility. The dynamics chosen were:

$$\begin{aligned} dX_t^D &= -\lambda^D X_t^D dt + \sigma^D dW_t^D \\ dX_t^C &= -\lambda^C X_t^C dt + \sigma^C dW_t^C \\ dX_t^G &= -\lambda^G (X_t^G - \mu^G) dt + \sqrt{v^G(X_t^G)} dW_t^G \end{aligned}$$

with

$$v^G(x) = \frac{2\lambda^G \sqrt{\delta^2 + x^2}\left(x - \dfrac{\beta\delta}{\sqrt{\alpha^2 - \beta^2}}\right)}{\alpha x - \beta\sqrt{\delta^2 + x^2}}.$$

The numerical values estimated for the parameters are given in tables 1 and 2.

Parameter	Estimate	Parameter	Estimate	Parameter	Estimate
κ^D	24000	κ^C	0.1228	κ^G	0.3412
a^D	3.32	a^C	4.16	a^G	3.88
b^D	-0.0867	b_1^C	-0.159	b_1^G	0.698
		b_2^C	0.0103	b_2^G	-0.419
				b_3^G	0.0587
c_1^D	-0.0119			c_1^G	-0.00613
d_1^D	-0.0212			d_1^G	0.0124
c_2^D	0.145	c^C	0.031	c_2^G	0.0817
d_2^D	0.0301	d^C	-0.00615	d_2^G	0.1
λ^D	80.3	λ^C	10.1	λ^G	8.08
σ^D	1.32	σ^C	0.16	μ^G	0

Table 1: Estimated parameters for the Demand and fuel price processes.

α	β	δ
14.7	0.524	0.126

Table 2: Estimated parameters for the function v^G.

Regarding the construction of the bid stack, we have

$$\begin{aligned} i \in \{c, g\} \qquad & x \in [0, \xi^i_{max}] \\ e_i(x) &= \alpha_i \eta_i \\ b_i^{BAU}(x) &= \gamma_i + \delta_i \eta_i x \end{aligned}$$

with the parameters of table 3 and $\xi^C_{max} = 3.997\text{x}10^5$, $\xi^G_{max} = 5.9013\text{x}10^5$.

α_C	η_C	α_G	η_G	γ_C	δ_C	γ_G	δ_G
0.411	2.04	0.973	2.63	0	6.05x 10^{-6}	0.178	3.73x 10^{-6}

Table 3: Estimated parameters for the bid stack functions.

Following the discussion in section 3.2.2, we have:

$$b_i(x, a, s_i) = a\alpha_i\eta_i + s_i(\gamma_i + \delta_i\eta_i x)$$

$$b_i^{-1}(p, a, s_i) = \xi_{max}^i \wedge \left(\frac{p - a\alpha_i\eta_i - s_i\gamma_i}{s_i\delta_i\eta_i}\right) \vee 0,$$

for $p \in [b_i(0, a, s_i), b_i(\xi_{max}^i, a, s_i)]$.

$b(x, a, s)$ is the inverse of $p \mapsto \sum_i b_i^{-1}(p, a, s_i)$, and the computations of μ_E are reduced to

$$\mu_E(x, a, s) = 252 \sum_i \int_0^{b^{-1}(b(x,a,s),a,s_i)} e(x)dx = 252 \sum_i \left[\alpha_i\eta_i b_i^{-1}(b(x, a, s), a, s_i)\right].$$

$\kappa = 252$ represents the number of trading days in one year. For the remaining parameters we considered $\Lambda = 5000$, $T = 8$ and $\Delta t_i \equiv 0.02$.

Following the same idea as in system (8), we considered the terminal condition:

$$\phi_\varepsilon(x) = \begin{cases} 0 & \text{if} \quad x \leq E_{cap} - \varepsilon \\ \Pi \ \dfrac{x - E_{cap} + \varepsilon}{\varepsilon} & \text{if} \quad E_{cap} - \varepsilon < x < E_{cap} \\ \Pi & \text{if} \quad x \geq E_{cap}. \end{cases}$$

We started with $\varepsilon = E_{cap}/10$ and the non-degenerate volatility $\sigma_E = 10^7$ as in [9]. To determine E_{cap} we started with the suggested value $E_{cap} = 2.74\text{x}10^9$. This led to a scenario where the trajectories of (E_t) never reached the limit and, therefore, Y_0 would remain always at 0. We then decided to try the approach in [9] with

$$E_{cap} = E_{avg} = T\mu_E(D_0, S_0^C, S_0^G, 0),$$

which is the same as assuming a constant Demand and fuel prices and a null certificate price. However, the opposite happened, and all the trajectories surpassed E_{cap} making $Y_0 = \Pi$. So, we decided to make E_{avg} the real average of the terminal values of the trajectories of the first iteration:

$$E_{avg} = \frac{1}{\Lambda} \sum_{\lambda=1}^{\Lambda} E_T^{0,\lambda}. \tag{9}$$

This also represents the average in a scenario where there is no regulation, since the

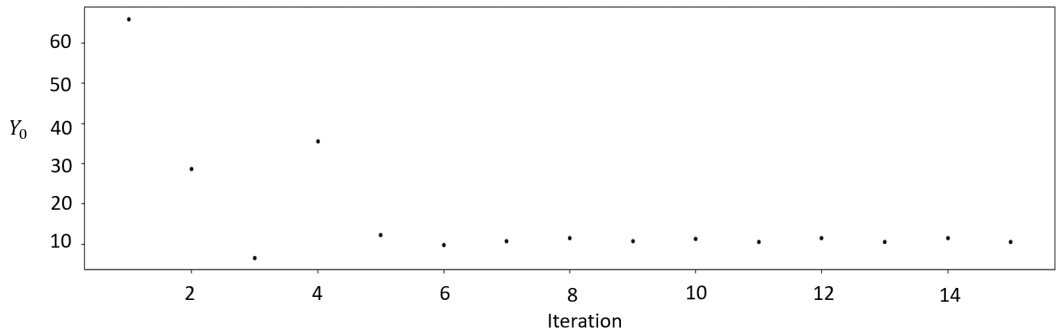

Figure 2: Y_0 with E_{cap} given by (9), $\varepsilon = E_{cap}/10,\ \sigma_E = 10^7$.

trajectories in the first iteration are calculated with $Y_t \equiv 0$. Fig. 2 shows the results obtained at 15 iterations.

In the same reasoning as in the mock system, we then tested different values for σ_E and ε. However, we could not find any other value for either parameter resulting in the algorithm's convergence. We note that we did not have the computer power to test the algorithm in more ideal conditions, for example, we could increase the number of trajectories per iteration as well as the total number of iterations.

4.3.3 Regime changes and linear reduction Demand

As part of the efforts to fight climate change, the EU has set several targets for 2030. One of those targets is to reduce the total energy consumption. To achieve this, several countries have already started to test some options in order to do so. For example, starting in the summer of 2022, Spain decreed that all shops must turn their window lights off after 10 p.m.. To incorporate these actions in our simulations, we decided to compare two different kinds of reduction methods in the Demand process.

The first method we tested was the regime changes. This is essentially adding a sudden shock to the process at a specified time. The second method was the gradual decay. We also tried a mixture of the two, where the gradual decay would start at a specified time (not at the beginning). We sticked to 15 iterations.

Regime Changes

We decided to include the shock in the Demand process at the halfway mark, that is, in the beginning of the fourth simulated year. So, we had:

$$D_t^{shock} = \begin{cases} D_t & \text{if} \quad t < 4 \\ pD_t & \text{if} \quad t \geq 4, \end{cases}$$

where $1-p$ represents the percentage of the shock (a 10% reduction means $p = 0.9$). The numerical results are presented in table 4 and fig. 3.

p	1	0.9	0.8	0.7	0.6	0.5
Y_0^{shock}	10.36	12.61	14.32	15.56	16.18	16.51

Table 4: Y_0^{shock} for different magnitudes of regime change at $t = 4$.

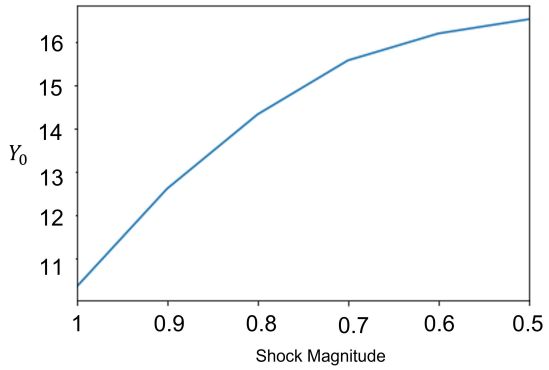

Figure 3: Y_0^{shock} for different magnitudes of regime change at $t = 4$.

We also tested a double shock approach, where the shocks were of similar magnitude at times $t = 4$ and $t = 6$:

$$D_t^{double\_shock} = \begin{cases} D_t & \text{if} \quad t < 4 \\ qD_t & \text{if} \quad 4 \leq t < 6 \\ q^2 D_t & \text{if} \quad t \geq 6. \end{cases}$$

The results are shown in table 5.

q	1	0.9	0.8	0.7
$Y_0^{double\_shock}$	10.36	12.93	14.26	15.32

Table 5: $Y_0^{double\_shock}$ for different magnitudes of regime change at $t = 4$ and $t = 6$.

We can see that for $p = q = 0.9$, $Y_0^{double\_shock}$ is around 2.5% greater than Y_0^{shock}. Also, for $p = q = 0.8$ and $p = q = 0.7$, we have $Y_0^{shock} > Y_0^{double\_shock}$ which is unexpected given the increasing tendency shown in the graph of fig. 3. However, the more logical comparisons should be, for example, $p = 0.8 \wedge q = 0.9$ or $p = 0.5 \wedge q = 0.7$, since, in both cases, D_8^{shock} and $D_8^{double\_shock}$ would have a similar multiplier when compared with D_8. In both of these cases, the impact on Y_0 of the single shock is more accentuated.

These observations lead us to suspect that the Demand process has greater influence on Y_0 when t is closer to 0. This hypothesis will be reinforced with the results obtained with the gradual decay method.

Gradual Decay

In order to simulate a gradual decay, we had the following:

$$D_t^{decay} = \left(1 - \overline{p}\,\frac{t}{8}\right) D_t,$$

where $\overline{p}$ represents the target percentage reduction at $t = 8$. The results are presented in table 6 and fig. 4.

$\overline{p}$	0	0.1	0.2	0.3	0.4	0.5
Y_0^{decay}	10.36	12.83	15.91	21.50	31.14	42.42

Table 6: Y_0^{decay} for different values of the decay.

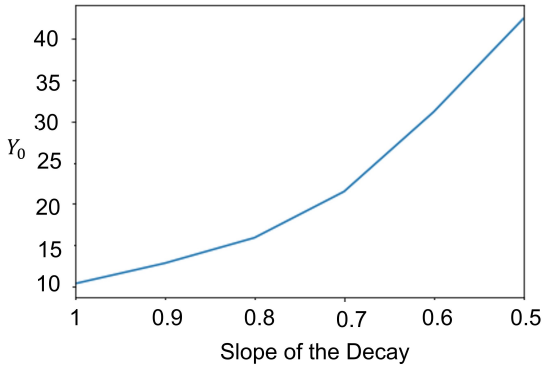

Figure 4: Y_0^{decay} for different values of the decay.

We can see that the values of Y_0^{decay} are very superior to those of Y_0^{shock}. The graph even becomes convex. These results support the hypothesis previously mentioned. This method modifies D_t from the initial time forward, but the final percentage change ends up being the same as with the shock method. However, the values of Y_0^{decay} escalate much higher than Y_0^{shock}. Thus, D_t impacts Y_0 more severely earlier in time.

We also tested a scenario where the decay would start at $t = 4$, but there was no significant change to the values when compared with the original Demand process. For example, with a 50% reduction, we had $Y_0^{delayed\_decay} = 12.85$, which is very close to Y_0^{shock} with $p = 0.9$.

5 Conclusion

The objective of this thesis was to describe a pricing mechanism in the Carbon Emissions Market through a model that has been developed over the past 15 years. The model aims to determine the price, at the initial time, of an Emissions Certificate Allowance, which can be seen as a derivative in this market. It features a Demand and a multi-dimensional fuel price processes, the CO_2 emissions process and the price process of the derivative. The main idiosyncrasy of the model is that it is constructed with a Forward-Backward SDE. This means that some processes are defined forwardly in time, while the allowance process is defined backwardly, having a terminal instead of initial condition.

We then presented a numerical scheme based on a Picard iteration method in order to approximate the solution of the equation system. The high complexity of the scheme was due to the fact that the FBSDE was coupled, that is, the forward processes depended on the backward process. Despite the complexity, we could still describe the scheme through a well-defined algorithm. However, given that the allowance process had to be simulated backwardly, at each time step, we had to approximate a conditional expected value. The approach we followed consisted in calculating the conditional expectation via a Least Squares Method. We then tested the algorithm with a mock equation and with a model adapted to the UK Energy market.

Finally, we tested the inclusion of potential environmental measures in the model. We sought to understand how the initial price of the Allowance variates when the Demand process suffers structural changes. Our results showed that Demand has greater impact on the initial Allowance price when time is closer to 0. That is, a small variation in the Demand process at an earlier time can be more impactful than a greater variation later in time.

The results presented in this thesis still leave room for research. On the theoretical framework, it would be desirable to develop a numerical method that considers the singularity of the terminal condition. In the practical framework, it would be ideal to apply an algorithm that could sustain the degeneracy of the forward component, which we were unable to do. Another possible improvement would be to calibrate the model parameters to different countries and to more recent data. A good try would be to consider the whole Iberian Peninsula instead of just a single country like Portugal.

References

[1] Antonelli, F. (1993) "Backward-forward stochastic differential equations", *Ann. Appl. Probab.* **3(3)**:777-793.

[2] Bender, C., Zhang, J. (2008) "Time discretization and Markovian iteration for coupled FBSDEs", *The Annals of Applied Probability* **18(1)**:143-177.

[3] Bernstein, M. A., Griffin, J. M. (2006) "Regional differences in the price-elasticity of demand for energy", *National Renewable Energy Laboratory.* Available on-line in http://www.osti.gov/bridge.

[4] Bismut, J.-M. (1973) "Conjugate Convex Functions in Optimal Stochastic Control", *Journal of Mathematical Analysis and Applications* **44(2)**:384-404.

[5] Carmona, R., Delarue, F. (2013) "Singular FBSDEs and scalar conservation laws driven by diffusion processes", *Probab. Theory Relat. Fields* **157**:333-388.

[6] René Carmona , Michael Coulon & Daniel Schwarz (2012) The valuation of clean spread options: linking electricity, emissions and fuels, Quantitative Finance, 12:12, 1951-1965, DOI: 10.1080/14697688.2012.750733.

[7] Carmona, R., Coulon, M., Schwarz, D. (2013) "Electricity price modeling and asset valuation: a multi-fuel structural approach", *Math Finan Econ* **7**:167-202.

[8] Carmona, R., Fehr, M., Hinz, J., Porchet, A. (2010) "Market Design for Emission Trading Schemes", *SIAM Review* **52(3)**:403-452.

[9] Chassagneux, J.-F., Chotai, H., Muûls, M. (2017) *A Forward-Backward SDEs Approach to Pricing in Carbon Markets*, SpringerBriefs in Mathematics of Planet Earth, Springer.

[10] Jianfu Chen, Jin Ma & Hong Yin (2018) Forward-backward SDEs with discontinuous coefficients, Stochastic Analysis and Applications, 36:2, 274-294294, DOI: 10.1080/07362994.2017.1399799.

[11] Chessari, J., Kawai, R., Shinozaki, Y., Yamada, T. (2021) "Numerical Methods for Backward Stochastic Differential Equations: A Survey", https://doi.org/10.48550/arXiv.2101.08936.

[12] Delarue, F. (2002) "On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case", *Stochastic Processes and their Applications* **99**:209-286.

[13] Delarue, F., Menozzi, S. (2006) "A forward–backward stochastic algorithm for quasilinear PDEs", *The Annals of Applied Probability* **16(1)**:140-184.

[14] Noufel Frikha, Vincent Lemaire. Joint Modelling of Gas and Electricity spot prices. 2009. hal-00421289v2.

[15] Gobet, E., Lemor, J.-P.,Warin, X. (2005) "A regression-based Monte Carlo method to solve backward stochastic differential equations", *The Annals of Applied Probability* **15**:2172-2202.

[16] Karoui, E., Peng, S., Quenez, M. C. (1997) "Backward Stochastic Differential Equations in Finance", *Mathematical Finance* **7(1)**:1-71.

[17] Ma, J., Protter, P., Yong, J. (1994) "Solving forward-backward stochastic differential equations explicitly - a four step scheme", *Probab. Theory Relat. Fields* **98(3)**:339-359.

[18] Ma, J., Protter, P., Martín, J., Torres, S (2002) "Numerical Method for Backward Stochastic Differential Equations", *The Annals of Applied Probability* **12(1)**:302-316.

[19] Ma, J., Wu, Z., Zhang, D., Zhang, J. (2015) "On well-posedness of Forward-Backward SDEs - A Unified Approach", *The Annals of Applied Probability* **25(4)**:2168-2214.

[20] Pardoux, E., Peng, S. (1990) "Adapted solution of a backward stochastic differential equation", *Systems & Control Letters* **14(1)**:55-61.

[21] Peng, S. (1992) "A Nonlinear Feynman–Kac Formula and Applications", in *Proceedings of Symposium of System Sciences and Control Theory*, ed. Chen and Yong. Singapore: World Scientific, 173–184.

[22] Howison, S., Schwarz, D. (2015) "Risk-Neutral Pricing of Financial Instruments in Emission Markets: A Structural Approach", *SIAM Review* **57(1)**:95-127.

[23] Schwarz, D. (2012) "Price Modelling and Asset Valuation in Carbon Emission and Electricity Markets", Ph.D. Thesis, University of Oxford.

[24] Zhang, J. (2004) "A Numerical Scheme for BSDEs", *The Annals of Applied Probability* **14(1)**:459-488.

Appendix

Recall the mollified equation:

$$\begin{cases} dX_t^\varepsilon = \mu_X(t, X_t^\varepsilon)dt + \sigma_X(t, X_t^\varepsilon)dW_t + \varepsilon d\tilde{W}_t \\ dE_t^\varepsilon = f(X_t^\varepsilon, Y_t^\varepsilon)dt + \varepsilon dB_t \\ dY_t^\varepsilon = Z_t^\varepsilon dW_t + \tilde{Z}_t^\varepsilon d\tilde{W}_t + \Upsilon_t^\varepsilon dB_t. \end{cases} \tag{3}$$

Proposition A.1 (Proposition 2.4 of [5]). *Assume that the coefficients μ_X, σ_X and f are bounded in x. Then, for the mollified equation (3) we have:*

$$\forall (t,x,e) \in [0,T[\times \mathbb{R}^d \times \mathbb{R},\ 0 \leq \frac{\partial \theta^\varepsilon}{\partial e}(t,x,e) \leq \frac{1}{l_1(T-t)},$$

where $\theta^\varepsilon = Y_{t_0}^{\varepsilon,t_0,x,e}$ and l_1 is the same from Assumption (2).
Moreover, the L^∞-norm of $\frac{\partial \theta^\varepsilon}{\partial e}$ on the whole $[0,T] \times \mathbb{R}^d \times \mathbb{R}$ can be bounded in terms of L (from Assumption (2)) and the Lipschitz norm of the terminal condition ϕ.

Proposition A.2 (Proposition 2.6 of [5]). *Assume that the coefficients μ_X, σ_X and f are bounded in x. Then, there exists a constant C, depending on L and T only, such that, for any $(t,x,e) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}$, we have:*

$$\frac{\partial \theta^\varepsilon}{\partial x}(t,x,e) \leq C.$$

As a consequence, for any $\delta \in]0,T[$ and any compact set $K \subset \mathbb{R}^d$, the $\frac{1}{2}$-Hölder norm of the function $[0, T-\delta] \ni t \mapsto \theta^\varepsilon(t,x,e)$, $x \in K$ and $e \in \mathbb{R}$, is bounded in terms of δ, K, L and T only; the $\frac{1}{2}$-Hölder norm of the function $[0,T] \ni t \mapsto \theta^\varepsilon(t,x,e)$ (that is the same function but on the whole $[0,T]$), $x \in K$ and $e \in \mathbb{R}$, is bounded in terms of K, L, T and the Lipschitz norm of ϕ only.

Proposition A.3 (Proposition 2.8 of [5]). *Consider the mollified equation (3) with a non decreasing Lipschitz smooth terminal condition ϕ satisfying*

$$\inf_{x \in \mathbb{R}} \phi(x) = 0, \quad \text{and} \quad \sup_{x \in \mathbb{R}} \phi(x) = 1.$$

Then, for any $\rho > 0$ and $q \geq 1$, there exists a constant $C(\rho, q) > 0$, only depending on

ρ, L and T, such that for any $t \in [0, T[$, e, $\Lambda \in \mathbb{R}$ and $|p| \leq \rho$, we have:

$$e > \Lambda \Rightarrow \theta^{\varepsilon}(t,x,e) \geq \phi(\Lambda) - C(\rho,q)\left(\frac{e-\Lambda}{L(T-t)}\right)^{-q},$$

$$e < \Lambda \Rightarrow \theta^{\varepsilon}(t,x,e) \leq \phi(\Lambda) + C(\rho,q)\left(\frac{\Lambda-e}{L(T-t)}\right)^{-q}.$$

In particular, for any $t < T$ and $x \in \mathbb{R}^d$:

$$\lim_{e\to+\infty} \theta^{\phi}(t,x,e) = 1, \quad \lim_{e\to-\infty} \theta^{\phi}(t,x,e) = 0,$$

$\theta^{\phi}(t_0,x,e) := Y_{t_0}^{\phi,t_0,x,e}$ is the limit of θ^{ε} as ε tends to 0.