ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI
Corso di Laurea Magistrale in Matematica
Curriculum Generale e Applicativo

EQUAZIONI STOCASTICHE BACKWARD NEL MERCATO DELLE EMISSIONI

Tesi di Laurea in
Equazioni differenziali stocastiche

Relatore:
Chiar.mo Prof.
Andrea Pascucci

Laureando:
Matteo Tassinari

II Sessione
Anno Accademico 2014/2015

Introduction

The EU ETS (European Union Emission Trading System) is the system adopted from European Union to combat climate change and its key tool for reducing industrial greenhouse gas emissions cost-effectively. It was launched in 2005 to fight Global warming and is a major pillar of EU climate policy and it still is the biggest greenhouse gas emissions trading scheme in the world. Indeed, as of 2013, it covers more than 11,000 power stations, factories, airlines and other installations with a net heat excess of 20 MW in 31 countries: all 28 EU member states plus Iceland, Norway, and Liechtenstein. Altogether it covers around 45% of total greenhouse gas emissions from the 28 EU countries. The greenhouse gasses included in this scheme are carbon dioxide (CO_2), nitrous oxide (N_2O) and perfluorocarbons (PFCs). The scheme is based on the 'cap and trade', this means that it is set a cap on the total amount of greenhouse gasses emitted by firms, then 'Allowances' for emissions are then auctioned off or allocated for free, and can subsequently be traded on the market.Under the EU ETS, the governments of the EU Member States agree on national emission caps which have to be approved by the EU commission. Those countries then allocate allowances to their industrial operators, and track and validate the actual emissions in accordance with the relevant assigned amount. They require the allowances to be retired after the end of each year. This mechanism give a price to greenhouse gases emission and companies have to hand in enough allowances to cover their emissions. If they emit too much polluting gas they have to buy allowances and the price of these certificates increases on the market. Conversely, if a firm has performed well at reducing its emissions, it can sell its leftover credits. This allows the system to find the most cost-effective ways of reducing emissions without significant government intervention. The scheme has been divided into a number of 'trading periods'. The first ETS trading period lasted three years, from January 2005 to December 2007. The second trading period ran from January 2008 until December 2012, coinciding with the first commitment period of the Kyoto Protocol. The third trading period began in January 2013 and will span until December 2020. Compared to 2005, when the EU ETS was first implemented, the proposed caps for 2020 represents a 21% reduction of greenhouse gases. This target has been reached 6 years early as emissions in the ETS fell to 1812 millions tonnes in 2014. As result of this system, emissions of greenhouse gases from installations participating in the EU ETS are estimated to have decreased by at least 3% in 2013. In order to fortify the scheme, the EU ETS cap has been reduced, because the target is to reduce of 40% the level of greenhouse gas emission of 1990 within 2030. In particular, the cap will need to be lowered by 2.2% per year from 2021, compared with 1.74% currently.

In the first chapter we present the EU ETS mechanism in detail making some example to make clear how it works. In particular we take a simplified model in which are considered only two resources to produce electricity: the first intensively polluting (carbon) and the second environmental friendly (gas). Moreover we introduce how the scheme works in the case of multiple compliance periods but in our treatment we only consider the case of one compliance period. Our aim is to study this model, thus we define a mathematical model which allows us to associate equations to this scheme. Three variable quantity are mainly involved: demand of electricity, total emissions of greenhouse gases and the allowance certificates price. These variables aren't deterministic since they depend by random events. For example we can't know exactly the demand of electricity in the future and we can only predict it. Then we set a probability space and we model them with stochastic processes, and the system of equations we obtain is a system of SDE (Stochastic Differential Equations). In particular we have two forward equations (demand of electricity D_t and total emissions E_t) and a backward equation (allowance certificates price A_t) with discontinuous terminal condition depending on the random variable E_T. In the second chapter we analyze a simplified model with Lipschitz terminal condition for the process A_t and we associate a partial differential equation to

the BSDE (Backward Stochastic Differential Equation) thanks to Itô's formula and vanishing viscosity solution to parabolic PDEs theory. The PDE we obtain is non-linear since the equations related to the processes E_t and A_t are a coupled system of FBSDE (Forward Backward Stochastic Differential Equation) and we show existence and uniqueness of the solution in a key theorem. Then we generalize the problem to the relaxed terminal condition $\mathbb{P}\{\phi_-(E_T) \leq A_T \leq \phi_+(E_T)\} = 1$ and with an additional assumption we show that a solution to the PDE exist and it is unique. Finally we consider the case with singular terminal condition and we show that the non-standardness of this kind of equation arises from the degeneracy of the forward component E_t. We observe that the random variable E_T develops a Dirac mass at the cap E_t for any starting point. Thus conditionally to the event $E_t = E_{cap}$ we observe that the standard Markovian structure brakes down at terminal time, and the terminal value of the allowance certificates cannot be prescribed as the model would require. In chapter three we solve the problem numerically by an explicit scheme: we set the model parameters representative of a typical market and we solve the PDE associated to the stochastic equation. This way, we obtain a numerical approximation of the function which represents the allowance certificate as a function of demand, emission and time. Finally we simulate processes D_t, E_t and A_t with a Monte Carlo. The last chapter is dedicated to a tax fraud discovered in late 2008 by Europol. Criminals took advantage of a weakness in the market of allowance certificate and trading it between different EU member state they were able to evade VAT.

Contents

List of Figures

Chapter 1

The European Union Emission Trading System - EU ETS

1.1 The Cap and Trade Scheme

Agents in the market demand electricity, the production of which causes emission and firms can produce it using different technologies that vary in their costs of production and their emissions intensity. Emissions cause environmental problem and to limit the consequence of intensive emission European governments decided to introduce a system which regulates the market of electricity to the purpose to reduce emissions. The European Union Emission Trading System is a greenhouse gas emission trading scheme which reduces emission of GHGs (GreenHouse Gasses), and it is based on allowance certificate and compliance periods. Emissions of GHGs are measured in equivalent tonnes of CO_2 and from now on we write CO_2 instead of CO_2 equivalent GHGs. In every compliance period there is a limit of emissions, called "the cap of emissions", and government don't want firms to exceed this cap. In order to do this, firms submit allowance certificates, each one worth one EU Allowance Unit (EUA), and permit to emit one tonne of CO_2. At the end of the compliance period firms must offset their cumulative emissions by submitting a sufficient number of certificates. If they doesn't, they must pay a monetary penalty; this event is called non-compliance. For a suitable penalty we will show that it's not convenient for firms to exceed the fixed cap. This scheme lead to a trading of allowance certificate, and as always we need to establish a price for an allowance certificate at each time.

The purpose of this trading scheme is to reduce the emission of CO_2, and it is done in two different ways: load shifting and long term abatement measures. Load shifting is the event in which companies shift the production of electricity from carbon to pollution friendlier resources. The cost of carbon is represented by the allowance certificate and this makes convenient to produce electricity using less polluting resources (like, for example, solar panel or windmill blade), although they are obviously more expensive then oil or coal. Also long term abatement measures are leaded by the cost of carbon, which makes attractive for firms to invest in pollution friendly technology and resources, because if a firm hasn't sufficient allowance certificate to cover its emission at the end of the period, it has to pay a monetary penalty or to buy additional certificates from other firms which haven't used all the certificates they have submitted. This means that polluting friendly firms can sell certificates to other companies and take profit. This way, companies are induced to do long term investment in less polluting technology and renewable resources.

Indeed, nowadays, electricity is produced from fossil fuels or renewable resources like nuclear fission, solar panel, windmill blade, hydropower, geothermal heat or biomasses. Most of the supply of electricity comes from fossil fuel because the process has a better production efficiency than renewable resources, but at the same time it is more polluting. Therefore is important to identify which generators are used in the market at any point in time. To do this we introduce the bid stack, which aggregates the bidding behaviour of firms that supply electricity. Each firm can supply electricity at a specific price, in particular a bid is the amount of electricity that a firm can sell to the market at a specific price in a specific hour during the next trading day. For example between 14 and 15 a bid can be $(700MW, 90€), (400MW, 110€), (100MW, 125€)$, which mens that the generator can sell its fist $700MW$ at a price of $90e$, the next $400MW$ at a price of $110e$, and the last $100MW$ at the price of $125e$. Mathematically, each firm supplies an increasing step function that maps electricity supply to its marginal cost. Then the market administrator makes a ranking list from the cheapest bid to the expensive one, so electricity is supplied to satisfy the demand at the lowest price. The bids are decided by firms considering fixed and variable costs, which depend on their production cost and the most of it is represented by the used resource. Obviously fossil fuels allow to supply a lower bid in absence of cap and trade system, and this scenario is called business as usual. The problem is that fossil fuels are pollution intensive resources and this will cause environmental problem and global warming. In case of cap and trade system the cost of electricity produced by polluting resources grows up, so it is not convenient to use only fossil fuels any more. Governments want to reduce environmental problems and the market administrator must rearrange bids to preserve the increasing order and, as a result, environmentally friendly technologies are now called upon before pollution intensive ones, leading to cleaner production of electricity. Therefore firms have to construct an equilibrium scenario in which they can sell electricity to the market at a competitive price, keeping in mind that they have a restriction on CO_2 emission induced by allowance certificates. Generators have to utilize their certificate for compliance, and they can also sell unused certificate to decrease their bid. In case of non compliance, bids are increased by an amount equal to the marginal emission rate of the plant (measured in tonne of CO_2 per MWh) multiplied by the allowance price. This price setting applies directly to the day-ahead spot prices by uniform auctions, as in the case most exchanges today.

Example 1.1. For sake of simplicity we consider that electricity can be produced only from coal (intensive pollution fuel) and gas (environmental friendly fuel).
Initially, as shown in figure 1.1, the cost of carbon is lower than cost of gas so the generator start to produce electricity from coal. This way coal bids come first in the bid stack and the marginal emission relative to coal bids is higher the gas one. Then emission become more costly and bid levels of both resources increase, but the coal bid increases more rapidly. We reach a limit time in which the electricity produced from coal is more expensive than the gas one, so the gas bid comes first in the bid stack than the coal one. We have reach our target, that is the allowance certificate induce to product electricity from gas (the pollution friendly resource).

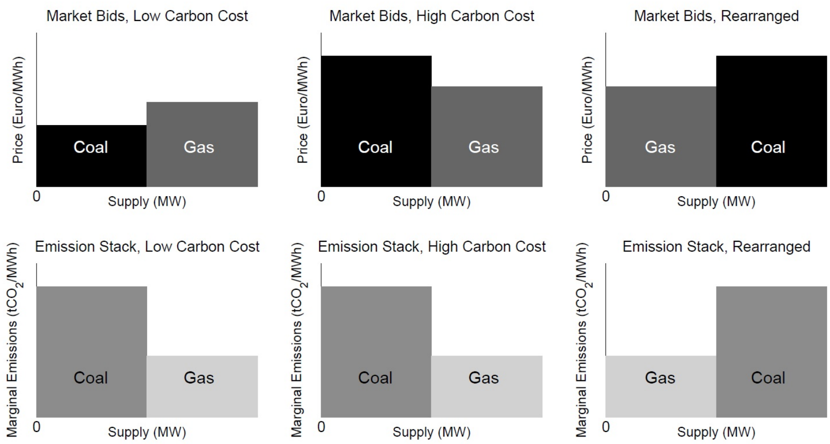


Figure 1.1: Rearrangement of the bid stack as the cost of carbon increases

1.2 The Mathematical Model

We have shown how the trade and cap system works, but we need to formalize it mathematically, so we can calculate allowance certificates price. First of all, we need to quantify the demand for electricity, which isn't a deterministic and fixed quantity, but it depends on consumers and it has a stochastic component. Then we introduce the electricity bid stack modelled as a continuous map from the supply of electricity to its marginal price and consequently we model the emission stack with a continuous map from the production of the last unity to the marginal emission. Later, we relate supply to demand, because we need to meet it at any point at any time and the total market emission rate, which depends on the technologies used to product electricity. Then, we introduce the cost of carbon and we show the difference between the business as usual system and the cap and trade one, and lastly we observe that the EU ETS scheme concretely reduces emission of CO_2 if compared with the classical system. Finally, we deduce the differential stochastic equations system from which we are able to calculate the risk-neutral price of allowance certificates.

As stated previously, the demand of electricity is not a deterministic quantity, but it has a random part at each time. Then we need to fix a filtered probability sparse on which we can study our model. We consider a time interval $[0, T]$ and let $(\Omega, \mathcal{F}, P, (\mathcal{F}_t))$ be a filtered probability space satisfying all usual assumptions, and let $(\mathcal{F}_t)_{t\in[0,T]}$ be the augmented filtration generated by a standard Brownian motion $(W_t)_{t\in[0,T]}$. From now to the end of the section we omit the time interval $[0, T]$, because we only refer to one compliance period which, without loss of generality, starts at time 0 and ends at time T. Each firm receive an initial allocation of allowance certificates and we assume them to be traded as liquidly financial products in which long and short position can be taken, because their cost of carry is negligible. Moreover we ignore the aggregation problem: we should consider the point of every agent on the market, but we only examine the point of view of the whole market. We can ignore this problem because we are looking for an arbitrage-free price for allowance certificates as a function of the aggregate forces that act on the market.

The consumers demand is represented by the $\mathcal{F}_t$-adapted stochastic process D_t. To supply this demand at any time $0 \leq t \leq T$ the aggregate of firms generates electricity, and we assume that the market uses only the currently available information to decide on its production level. This level is non-negative and below a constant value that expresses the maximum production capacity. We define the $\mathcal{F}_t$-adapted process ξ_t as the aggregate amount of electricity generated by all firms, and we assume:

$$0 \leq \xi_t \leq \xi_{max} \quad 0 \leq t \leq T \tag{1.1}$$

where $\xi_{max} \geq 0$ is the aggregate maximum production capacity of all firms. The market administrator ensures that the aggregate demand and aggregate supply of electricity are matched on a daily basis. This means that

$$D_t = \xi_t \quad 0 \leq t \leq T \tag{1.2}$$

The last assumption leads to

$$0 \leq D_t = \xi_{max} \quad 0 \leq t \leq T \tag{1.3}$$

and this means that there are always sufficient resources to meet demand. Usually, demand and supply are quoted in megawatts(MW). Now we want to model emissions of CO_2. We define the cumulative emission during the interval $[0, t]$, $t \leq T$ by E_t, and it is measured in tonne of CO_2. So we have an $\mathcal{F}_t$-adapted process E_t. The cumulative emission in every interval of the type $[0, t]$, $t \leq T$ is finite, then we have the constraint

$$0 \leq E_t \leq E_{max} \quad 0 \leq t \leq T \tag{1.4}$$

where E_{max} represent the maximum production of CO_2 on the interval $[0, t]$, $t \leq T$. In this trade and cap scheme, the regulator decides on an acceptable maximum level of cumulative emission and we call it E_{cap} and issues a corresponding number of allowance certificates. Obviously it must be

$$0 \leq E_{cap} \leq E_{max} \tag{1.5}$$

At the end of the compliance period firms have to balance the quantity of CO_2 emitted and the number of certificates allocated, because cumulative emissions in the market are offset against the initial allocation of allowances. In the case of one compliance period, certificates that are unused expire and lose their value. If the amount of emission for each firm exceeds the number of certificates submitted then they must pay a monetary penalty. This penalty is $\pi \geq 0$ for each tonne of CO_2 which is not offset against allowance certificates. This means that the monetary penalty is equal to π multiplied by the positive part of $E_T - E_{cap}$:

$$\text{Penalty} = \pi(E_T - E_{cap})^+ = \begin{cases} 0 & \text{if } E_T \leq E_{cap} \\ \pi(E_T - E_{cap}) & \text{if } E_T > E_{cap}. \end{cases} \tag{1.6}$$

Finally we define the $\mathcal{F}_t$-adapted process A_t, which represent the value of the allowance certificate. We want to remark that the allowance certificate are the traded asset in the market and we assume the existence of a risk-free asset, the so called bond, with a constant risk-free interest rate $r \geq 0$. Assuming that the simple interest is paid more and more frequently, we have the formula of continuous compounding with annual interest rate r:

$$B_t = B_0 e^{rt} \tag{1.7}$$

We have just defined the process involved in this scheme, and now we have to go on making some assumptions on the market which lead to a well posed definition on bid stack and summarize the action of the central market.

Assumption 1. The market administrator ensures that resources are used according to the merit order. This means that the cheapest production technologies are called upon to satisfy a given demand and hence electricity is supplied at the lowest possible price.

The price of electricity is strongly dependent on the technologies used to produce it, so the marginal price is strictly increasing as the demand grows up. Then we can model the bid stack as an increasing step function, and assuming that it has sufficiently many steps we can approximate it by a smooth function. First we define this function in absence of the cap-and-trade scheme, and we call it b^{BAU}, where BAU means Business As Usual.

Definition 1.2. The business-as-usual bid stack is represented by the bounded function

$$\begin{aligned} b^{BAU} : [0, \xi_{max}] &\rightarrow \mathbb{R}^+ \\ \xi &\mapsto b^{BAU}(\xi) \end{aligned} \tag{1.8}$$

Moreover $b^{BAU} \in C^1([0, \xi_{max}])$ and $\frac{db^{BAU}}{d\xi}(\xi) > 0$. The quantity $b^{BAU}(\xi)$ represent the bid level of the marginal of the marginal production unit measured in MWh, and ξ the supply of electricity measured in MW.

Roughly speaking, $b^{BAU}(\xi)$ denotes the total cost that arises when the quantity produced is incremented by one unit. That is, it is the cost of producing one more unit of electricity. In general terms, marginal cost at each level of production includes any additional costs required to produce the next unit. We remark that in reality business-as-usual bid stack is stochastic, because it depends on fuel price and other variable costs which fluctuate continuously. Nevertheless we ignore this fact ad we assume the bid stack to be a continuous deterministic function because we are only interested in the relative position of the different technologies in the bid stack, and moreover historic data observations show that it is only relevant in the long-run, but we only consider the one compliance period case.The emission intensive technologies are usually cheaper then the environmental friendly ones, so bids associated with a small level correspond to electricity produced from emission intensive generators. Conversely, bids at the right end of the interval $[0, \xi_{max}]$ stem mostly to environmental friendly technologies.

We now construct an emission stack, by creating a map from the supply of electricity to the marginal emission associated with the supply of the last unit.

Definition 1.3. The marginal emission stack is represented by the bounded function

$$\begin{aligned} e : [0, \xi_{max}] &\rightarrow \mathbb{R}^+ \\ \xi &\mapsto e(\xi) \end{aligned} \tag{1.9}$$

Moreover $e \in C^1([0, \xi_{max}])$. The quantity $b^{BAU}(\xi)$ associates with a specific supply of electricity ξ the emission rate of the marginal unit (measured in tonne of CO_2 per MWh).

Proposition 1.4. *The business-as-usual market emission rate* μ_e^{BAU} *is given by*

$$\mu_e^{BAU}(D) = k \int_0^D e(\xi) d\xi, \quad for \quad 0 \leq D \leq \xi_{max} \tag{1.10}$$

where k *is a strictly positive constant.*

Proof. Thanks to assumption 1.3 firms produce the exact amount of electricity consumers demand, and the generation capacity associated to the interval $[0, D]$ is used to meet demand. Then, the market emission rate per hour is obtained integrating the marginal emission stack over the interval $[0, D]$, i.e. the current level of demand. We now multiply this quantity for a strictly positive constant k in order to obtain the market emission rate μ_e^{BAU} measured on the same time that characterises T. We rescale it because the rate e is measured per hour, but we want μ_e^{BAU} to live in the same timescale as T (for example months or years). □

We are ready to introduce the cap-and trade scheme in the business-as-usual economy defined above and see its consequences. As seen before this scheme gives a price to carbon emission so it increases the production cost for firms. Then it leads to use environmental friendly technologies especially for firms which mostly rely on emission-intensive resources, because they have to buy additional allowance certificates to avoid penalization, and it makes their electricity more expensive. Moreover it follows that the level of their bid would be lower. Conversely, if a firm owns more certificate than it needs, it can sell it in the market and they have a profit. This way we can see the cost of carbon as an opportunity to increase incomes. In the long-run, these environmental friendly firms can reduce marginal cost and make lower level's bid so they increase competitiveness in the market. Nevertheless, we ignore the long-time behaviour of the long-term abatement process because we are only interested in the direct impact on the bid stack. We assume the cost of carbon is directly applied on the electricity price in the market in order to maintain constant the profit margin for firms. We observe that the cost of carbon, for each firm, is represented by allowance certificates, so we have to increase the business-as-usual bids by an amount equal to the allowance price multiplied by the marginal emission rate of the firm. Therefore, given an allowance price A the bid stack function becomes:

$$g(A,\xi) = b^{BAU}(\xi) + Ae(\xi), \quad \text{for} \quad 0 \le A < \infty, \quad 0 \le \xi < \xi_{max} \tag{1.11}$$

and we note that if $A = 0$ it is equal to the business-as-usual case. Generally, if $A \neq 0$ the bid stack function lose its monotonicity. Indeed, if the cost price of allowance certificate (the cost of carbon) becomes relatively more expensive, then bids associated with large marginal emission rates becomes relatively more expensive to produce for firms which relies on polluting resources.

Definition 1.5. We define the set of active generation units at a given allowance and electricity price P by

$$S(A,P) = \{\xi \in [0,\xi_{max}] \quad | \quad g(A,\xi) \le P\} \quad \text{for} \quad 0 \le A < \infty, \quad 0 \le P < \infty \tag{1.12}$$

We can define the map

$$\begin{aligned} \chi : [0,\infty[&\to \mathbb{R}^+ \\ P &\mapsto \lambda(S(\cdot,P)) \end{aligned} \tag{1.13}$$

where λ denotes the Lebesgue measure. This map, by the definition of sublevel set, is strictly increasing.

Assumption 2.

$$\lambda\left(\left\{\xi \in [0,\xi_{max}] \quad \middle| \quad \frac{\partial b^{BAU}}{\partial \xi}(\xi) + A\frac{\partial e}{\partial \xi}(\xi) = 0\right\}\right) = 0 \tag{1.14}$$

Under assumption 2 the function χ becomes continuous and therefore invertible. Then, using 1.12, the market bid stack is defined by

$$b(A,\xi) = (\lambda(S(A,\cdot)))^{-1}(\xi) \quad \text{for} \quad 0 \le A < \infty, \quad 0 \le \xi < \xi_{max} \tag{1.15}$$

And it implies the market price of electricity definition

$$P = b(A,D) \quad \text{for} 0 \le A < \infty, \quad 0 \le D < \xi_{max} \tag{1.16}$$

Proposition 1.6. *In presence of the cap and trade scheme and given an allowance price A and demand level D, the market emission rate is given by*

$$\mu_e(A,D) = k\int_{S_p(A,D)} e(\xi)d\xi, \quad \textit{for} \quad 0 \le A < \infty, \quad 0 \le D \le \xi_{max} \tag{1.17}$$

where $S_p(A,D) = S(A,b(A,D))$, and k is a positive time scaling constant.

Proof. The constant k is defined as in the proof of proposition 1.4. Under business-as-usual, demand D is satisfied by the generation capacity $[0, D]$, which is seen as a subset of the domain of the emission stack. The cap and trade scheme leads to a shift of this interval to the right or, depending on the shape of the marginal emission stack, split into multiple sets, and we call this effect load shifting. We define this new set as $S_p(A, D) = S(A, b(A, D))$, and the proof is complete. $\square$

Note that if $A = 0$ then $S_p(A, D) = [0, D]$. Now we show that the emission rate μ_e just defined has some properties which makes it well defined, that is its behaviour is what we intuitively expect to be from the real case. Moreover it is a regular function and it will be useful in the following. To explain this we have the lemma:

Lemma 1.7. *The market emission rate μ_e satisfies:*

(P1) *The map $D \mapsto \mu_e(\cdot, D)$ is:*

- **(i)** *strictly increasing*
- **(ii)** *Lipschitz continuous*

(P2) *The map $A \mapsto \mu_e(A, \cdot)$ is:*

- **(i)** *non increasing*
- **(ii)** *Lipschitz continuous*

(P3) *μ_e is bounded.*

Proof. **(P1)** **(i)** For $0 \leq D_1 < D_2 \leq \xi_{max}$, we have $S_p(\cdot, D_1) \subset S_p(\cdot, D_2)$ thanks to assumption 2. Moreover $e(\xi)$ is positive on the interval $[0, \xi_{max}]$, then

$$\mu_e(A, D_1) = k \int_{S_p(A,D_1)} e(\xi) d\xi \leq k \int_{S_p(A,D_2)} e(\xi) d\xi = \mu_e(A, D_2)$$

For all $0 \leq A < \infty$ fixed.

(ii) For $0 \leq D_1 < D_2 \leq \xi_{max}$ we define $\Delta^D S_p(D_2, D_1) = S_p(\cdot, D_2) \S_p(\cdot, D_1)$. Therefore

$$\begin{aligned} \mu_e(\cdot, D_2) - \mu_e(\cdot, D_1) = k \int_{\Delta^D S_p(D_2,D_1)} e(\xi) d\xi \leq \\ \leq K \lambda(\Delta^D S_p(D_2, D_1)) \max_\xi e(\xi) = \\ = k(D_2 - D_1) \max_\xi e(\xi) \end{aligned}$$

where λ is the Lebesgue measure. For $0 \leq D_2 < D_1 \leq \xi_{max}$ the proof is analogous.

(P2) **(i)** For $0 \leq A_1 < A_2 < \infty$, we define $\Delta^A S_p(A_1, A_2) = S_p(A_1, \cdot)\S_p(A_2, \cdot)$. Therefore

$$\mu_e(A_1, \cdot) - \mu_e(A_2, \cdot) = k \int_{\Delta^A S_p(A_1,A_2)} e(\xi) d\xi - k \int_{\Delta^A S_p(A_2,A_1)} e(\xi) d\xi$$

Fixed $0 \leq D \leq \xi_{max}$, we have:

- $\lambda(\Delta^A S_p(A_1, A_2)) = \lambda(\Delta^A S_p(A_2, A_1))$

- $e(\xi) = (g(A_2, \xi) - g(A-1, \xi))(A_2 - A_1)^{-1} > (b(A_2, D) - b(A-1, D))(A_2 - A_1)^{-1}$ on $\Delta^A S_p(A_1, A_2)$

- $e(\xi) = (g(A_2, \xi) - g(A-1, \xi))(A_2 - A_1)^{-1} \leq (b(A_2, D) - b(A-1, D))(A_2 - A_1)^{-1}$ on $\Delta^A S_p(A_2, A_1)$

then

$$\begin{aligned}\mu_e(A_1, \cdot) - \mu_e(A_2, \cdot) &> k\lambda(\Delta^A S_p(A_1, A_2))(b(A_2, D) - b(A-1, D))(A_2 - A_1)^{-1} - \\ &- k\lambda(\Delta^A S_p(A_1, A_2))(b(A_2, D) - b(A-1, D))(A_2 - A_1)^{-1} = \\ &= 0\end{aligned}$$

thus $\mu_e(A_1, \cdot) > \mu_e(A_2, \cdot)$ for $A_1 < A_2$.

(ii) Assume that $0 \leq A_1 < A_2 < \infty$. Since $e(\xi)$ is bounded and $\lambda(\Delta^A S_p(A_1, A_2)) = \lambda(\Delta^A S_p(A_2, A_1))$ we have

$$\begin{aligned}|\mu_e(A_1, \cdot) - \mu_e(A_2, \cdot)| &= k \int_{\Delta^A S_p(A_1,A_2)} e(\xi) d\xi - k \int_{\Delta^A S_p(A_2,A_1)} e(\xi) d\xi \leq \\ &\leq C_1 \lambda(\Delta^A S_p(A_1, A_2))\end{aligned}$$

with $C_1 \geq 0$. It is clear that $\Delta^A S_p(A_1, A_2)$ and $\Delta^A S_p(A_2, A_1)$ can be written as a finite union of intervals. As A_1 increases to A_2, there are three possibilities:

(a) existing intervals grow or shrink;

(b) new intervals appear or existing ones disappear;

(c) the intervals remain unchanged.

Now we differentiate the level curve $g(A, \xi) = b(A, D)$ for a given D:

$$\frac{\partial g}{\partial \xi} + \frac{\partial g}{\partial A} \cdot \frac{d\xi}{dA} = \frac{\partial b}{\partial A}$$

because ξ implicitly depends on A. And it follows:

$$\frac{d\xi}{dA} = -\frac{\left(\frac{\partial g}{\partial \xi} - \frac{\partial b}{\partial A}\right)}{\frac{\partial g}{\partial \xi}}$$

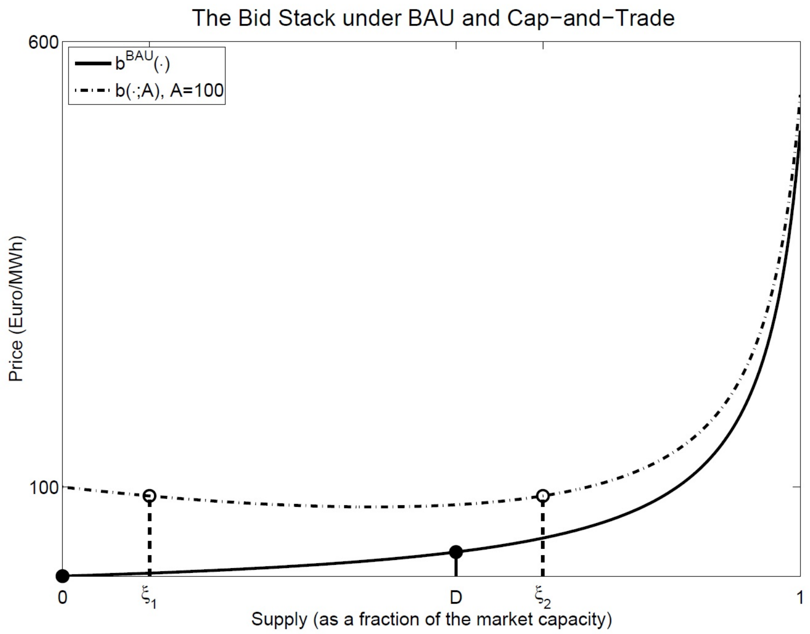
and thanks to assumption 2 it is bounded by a constant $C_2 \geq 0$. Then, in each case (a)-(c), as A changes, the intervals describing $\Delta^A S_p(A_2, A_1)$ don't move faster than $C_2(A_2 - A_1)$, and analogously it is true also for $\Delta^A S_p(A_1, A_2)$. Similarly we can obtain the same result for $A_1 > A_2$, and the statement is proved.

(P3) Since $S_p(A, D) \subseteq [0, \xi_{max}]$ for all $A \geq 0$ and $0 \leq D \leq \xi_{max}$ and keeping in mind that by definition $e(\xi)$ is bounded we obtain:

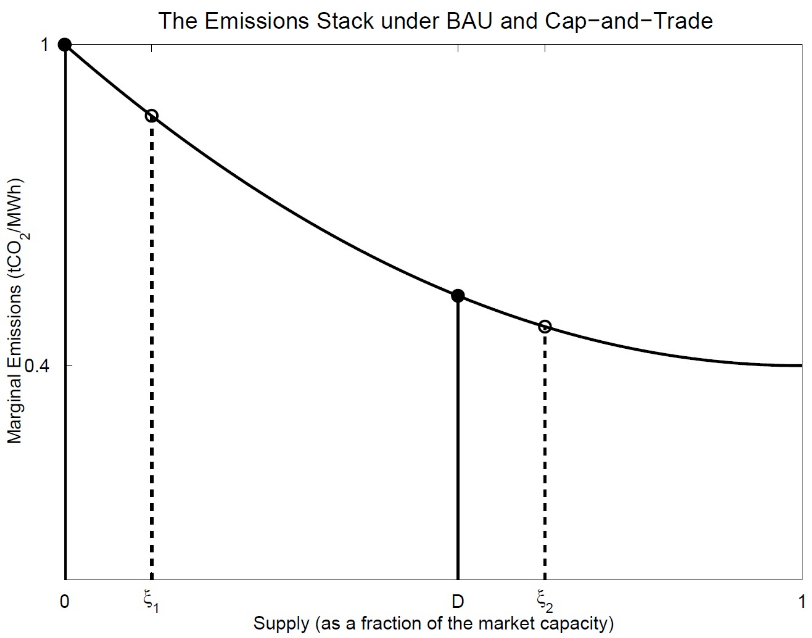
$$\begin{aligned} |\mu_e(A, D)| &= k \int_{S_p(A,D)} e(\xi) d\xi \leq \\ &\leq C_1 \lambda([0, \xi_{max}]) = \\ &= C_1 \xi_{max} < \infty \end{aligned}$$

□

We have just defined instantaneous emission, and to see the effect to load shifting and the following reduction of marginal emission we have to calculate cumulative emission. We can do this by integrating instantaneous emission up to the current time $0 \leq t \leq T$. In figure 1.2 we can see the difference between the business-as-usual market and the cap-and-trade scheme. Thanks to function b, the supply of resources to meet demand $[0, D]$ is led to be shifted to the interval $[\xi_1, \xi_2]$. Therefore,assuming that under BAU dirtier production technologies are placed further to the left in bid stack, instantaneous emission are now given by the smaller integral over the emission stack from ξ_1 to ξ_2. Then also cumulative emissions are smaller then the business-as-usual case.



(a) *Bid stacks b^{BAU} and b.*



(b) *Emissions stack e.*

Figure 1.2: The effect of cap-and-trade scheme on the bid stack and on the emission stack

Now we are ready to set up the problem of determining the arbitrage-free price of an allowance certificate. In order to do this we need to make the following assumption:

Assumption 3. There exist an equivalent martingale measure $Q \sim P$,under which, for $0 \leq t \leq T$, the discounted price of any tradable asset in the market is a martingale. We refer to Q as the risk-neutral measure.

This technical assumption guarantees that our market is arbitrage-free. Indeed, if assumption 3 is true it is verified the First Fundamental Theorem of Asset Pricing.

Theorem 1.8. *A market is arbitrage-free if and only if there exist at least one equivalent martingale measure.*

First of all, we want to make some additional assumption about the the cumulative emission process E_t an demand process D_t. At time $t = 0$, we assume to know the the demand of electricity, i.e. we assume to know D_0. Moreover we assume D_t to evolve according to an Itô diffusion. It means that, for each time $0 \leq t \leq T$, D_t is given by the stochastic differential equation

$$dD_t = \mu_d(D_t)dt + \sigma_d(D_t)dW_t, \quad D_0 = d \in (0, \xi_{max}) \tag{1.18}$$

We should write $\widetilde{W}_t$ instead of W_t, because W_t is an $\mathcal{F}_t$-adapted standard Brownian motion on the probability space $(\Omega, \mathcal{F}, P, (\mathcal{F}_t))$, but we have changed the probability measure and $\widetilde{W}_t$ is an $\mathcal{F}_t$-adapted standard Brownian motion on the probability space $(\Omega, \mathcal{F}, Q, (\mathcal{F}_t))$. This is only a technical clarification and from now, keeping in mind this remark, we write W_t instead of $\widetilde{W}_t$. Both coefficient μ_d and σ_d are functions of demand only, but in reality they should depend explicitly on time because demand of electricity has seasonal variation. Moreover we assume μ_d and σ_d to be positive, globally Lipschitz continuous and exhibits at most linear growth.These assumption will be relevant in the following.

Let's talk about cumulative emission process: cumulative emission are measured starting from the initial time 0, so it must be $E_0 = 0$. To calculate it we have to integrate instantaneous market emission rate μ_e on the interval $[0, t]$. Therefore we have

$$dE_t = \mu_e(A_t, D_t)dt, \quad E_0 = 0 \tag{1.19}$$

Clearly E_t is not decreasing as we intuitively expected it to be because it is a cumulative quantity. Finally, we have to characterise the process A_t which model the allowance price certificate. The problem is we do not know the value of A_t at time $t = 0$, so we can not represent this process with a forward stochastic differential equation as we do for processes E_t and D_t. In this case we know the value of the allowance certificate at the end of period. In the event of non-compliance, at the end of the period, the price of a certificate is equal to the monetary penalty π. This value is given by an arbitrage argument, because if $A_T > \pi$ we can build an arbitrage strategy, that is we can short sell certificates to the market making a sure risk-free profit equal to $A_T - \pi$. Thus $A_T \leq \pi$ and moreover, it can not be $A_T < \pi$ because of the penalization system. We conclude $A_T = \pi$ in case of non-compliance. On the other hand, clearly, in case of compliance A_T takes value 0.

$$A_T = \begin{cases} 0, & \text{if} \quad 0 \leq E_T \leq E_{cap} \\ \pi, & \text{if} \quad E_{cap} \leq E_T \leq E_{max} \end{cases} \tag{1.20}$$

We observe that, different from other processes, this is a backward stochastic problem and it must be solved with different techniques. In particular, we proceed in the same way we solved BSDE (Backward differential stochastic Equations) in section $B.2$ of appendix B. We know, thanks to assumption 3, that discounted allowance price is a martingale under measure Q. Then we can represent the process A_t at each time as the discounted conditional expectation of its terminal condition under measure Q.

$$A_t = E^Q \left[\pi e^{-r(T-t)} \mathbb{I}_{[E_{cap}, \infty)}\right], \quad 0 \leq t \leq T \tag{1.21}$$

or equivalently

$$e^{r(T-t)} A_t = E^Q \left[\pi \mathbb{I}_{[E_{cap}, \infty)}\right], \quad 0 \leq t \leq T \tag{1.22}$$

Note that the price process A_t takes value only in the interval $[0, \pi]$. The we have the following system of stochastic differential equations:

$$\begin{cases} dD_t = \mu_d(D_t)dt + \sigma_d(D_t)dW_t, & D_0 = d \in (0, \xi_{max}); \\ dE_t = \mu_e(A_t, D_t)dt, & E_0 = 0; \\ A_t = \pi e^{-r(T-t)} E^Q \left[\mathbb{I}_{[E_{cap},\infty)}(E_T) \middle| \mathcal{F}_t \right], & A_T = \pi \mathbb{I}_{[E_{cap},\infty)}(E_T). \end{cases} \tag{1.23}$$

On the other hand, the process A_t is a martingale under measure Q and the filtration $(\mathcal{F}_t)_t$ is natural, so we can apply the Martingale Representation Theorem. Then there exist an unique process $Z_t \in \mathbb{L}^2(\mathcal{F}_t)$ such that

$$d(e^{r(T-t)}A_t) = Z_t dW_t, \quad 0 \leq t \leq T \tag{1.24}$$

This means that we can represent A_t as an Itô integral respect to the Brownian motion W_t. It follows

$$\begin{aligned} Z_t dW_t &= d(e^{r(T-t)}A_t) = \\ &= -re^{-rt}A_t + e^{-rt}dA_t \end{aligned}$$

therefore

$$e^{-rt}dA_t = re^{-rt}A_t + Z_t dW_t$$

and finally

$$dA_t = rA_t dt + e^{rt}Z_t dW_t, \quad 0 \leq t \leq T \tag{1.25}$$

Thus we can rewrite system 1.23 as follows

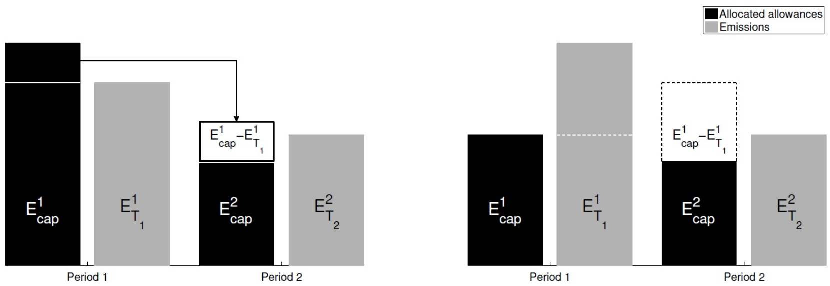
$$\begin{cases} dD_t = \mu_d(D_t)dt + \sigma_d(D_t)dW_t, & D_0 = d \in (0, \xi_{max}); \\ dE_t = \mu_e(A_t, D_t)dt, & E_0 = 0; \\ dA_t = rA_t dt + e^{rt}Z_t dW_t, & A_T = \pi \mathbb{I}_{[E_{cap},\infty)}(E_T). \end{cases} \tag{1.26}$$

Note that the problem 1.26 is a FBSDE (Forward Backward Stochastic Differential Equation) and in particular, we have two forward equations and one backward. Moreover the demand equation is independent from the others, but the equations for emission and allowance certificate price are coupled, and this is a complicated mathematical problem. For an introduction to this type of equations see appendix B and for a complete treatment see [6].

1.3 Multiple compliance periods

We have seen the scheme for a compliance period, but this model also considers the case of multiple compliance period. In this section we shortly expose how EU ETS scheme regulate the transition from a compliance period to the following one. We do this for completeness, but it's not theme of this thesis because we will solve the problem only in the case of one compliance period.

Governments have introduced three mechanism to regulate the transition between subsequently compliance periods and they go the by names of banking, borrowing and withdrawal. At the end of the first compliance period they may haven't used all certificates they have submitted, so these certificates become perfect substitutes for certificates issued in the second period. This feature is called banking and in the event of compliance the unused certificates are exchanged with certificates valid for the second period. Moreover, in case of compliance, it increases the price of allowance certificates at the end of the period from zero to the price of certificates of the sequent period because a firm can use in the second period a certificate saved in the first one. The purpose is to reduce emission of CO_2, so they introduce the mechanism called withdrawal, which makes banking stronger. Indeed, thanks to banking, in case of compliance, the price of the allowance certificate in the first period is equal to the price at the beginning of the second one. In case of non compliance at the end of the first period, each firm has to pay a monetary penalty and moreover the number of surplus certificates is withdrawn form the sequent allocation. It follows that there are fewer certificates in the second period and it propels firms to emit less tonne of CO_2. So the price of the certificate in the first period increases, because these two mechanism lead to the first period allowance certificate taking the value of the sum of the second period certificate and the penalization. This means that withdrawal can be seen as a supplement penalization. Roughly speaking, this is an incentive to save certificates; indeed, thanks to the withdrawal mechanism, there could be less certificates on the market and those saved could have a greater value. The third mechanism is borrowing, which, together with banking and withdrawal, keeps consecutive compliance periods



(a) *Banking.* **(b)** *Withdrawal.*

Figure 1.3: Banking and withdrawal mechanisms in an emission market with two periods

connected, and decreases the probability with which non compliance occurs. This feature allows firms to bring forward certificates from the second allocation and use them in the first period. This way the case of non compliance occurs only if the entire allocation of certificates of the second period is borrowed to supplement the aggregate supply during the first period. This mechanism does not affect the aggregate supply of certificates because the number of certificates borrowed is subtracted from the second period allocation. Putting together these three mechanism, we

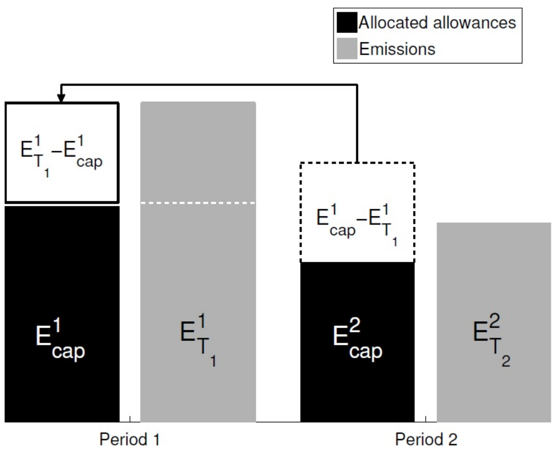


Figure 1.4: Borrowing mechanism in an emission market with two periods

have a model which keeps under control by authority the global emission of CO_2, and it is flexible enough for firms to manage the energy production and to supply the demand of electricity. Flexibility is very important because we have to keep in mind that companies have to sell electricity at a competitive price.

Chapter 2

Solution of the problem and associated PDE

In the previous chapter we have formulated mathematically the EU ETS scheme, and in this chapter we want to show that exist a solution to the stochastic system:

$$\begin{cases} dD_t = \mu_d(D_t)dt + \sigma_d(D_t)dW_t, & D_0 = d \in (0, \xi_{max}); \\ dE_t = \mu_e(A_t, D_t)dt, & E_0 = 0; \\ A_t = \pi e^{-r(T-t)} E^Q \left[\mathbb{I}_{[E_{cap},\infty)}(E_T) \middle| \mathcal{F}_t \right], & A_T = \pi \mathbb{I}_{[E_{cap},\infty)}(E_T). \end{cases} \tag{2.1}$$

Moreover we want to associate a Cauchy problem to the previous stochastic system. Guided by intuition, with a purely formal procedure, we can obtain the desired PDE, but we use this method only to show our target. Indeed, to do this we assume the existence of the solution of the stochastic system and of the Cauchy problem, but now we don't know it. In appendix B we show the connection between PDEs and linear SDEs and the hypothesis under which we can apply Itô's formula, but this time our problem isn't linear and we have to use different techniques. Assume that the the following Cauchy problem has a solution

$$\begin{cases} \frac{\sigma_d(x)^2}{2} \frac{\partial^2 u}{\partial x^2} + \mu_d(x) \frac{\partial u}{\partial x} + \mu_e(u, x) \frac{\partial u}{\partial y} + \frac{\partial u}{\partial t} - r = 0 \\ u(T, x, y) = \pi \mathbb{I}_{[E_{cap},\infty)}(y) \quad (x, y) \in [0, \xi_{max}] \times [0, E_{max}] \end{cases} \tag{2.2}$$

and assume that problem 2.1 has a solution $u = u(t, x, y) \in C^{1,2}$. In this hypothesis we can apply Itô's formula to the process $A_t = u(t, D_t, E_t)$:

$$\begin{aligned} dA_t =& \frac{\partial u}{\partial t}dt + \frac{\partial u}{\partial x}dD_t + \frac{\sigma_d^2}{2}\frac{\partial^2 u}{\partial x^2}dt + \frac{\partial u}{\partial x}dE_t = \\ =& \left(\frac{\partial u}{\partial t} + \frac{\sigma_d^2}{2}\frac{\partial^2 u}{\partial x^2} + \mu_d\frac{\partial u}{\partial x} + \mu_e\frac{\partial u}{\partial y}\right)dt + \sigma_d\frac{\partial u}{\partial x}dW_t \end{aligned} \tag{2.3}$$

Since A_t is the price of traded security with neutral risk drift r, we equate the drift of A_t to r and we obtain:

$$\frac{\partial u}{\partial t} + \frac{\sigma_d^2}{2}\frac{\partial^2 u}{\partial x^2} + \mu_d\frac{\partial u}{\partial x} + \mu_e\frac{\partial u}{\partial y} - r = 0 \tag{2.4}$$

Thus this semilinear PDE is the one associated to the process A_t. Following this intuition we want to study this relation formally.

2.1 A simplified problem - Lipschitz terminal condition

We first consider the simplified problem

$$\begin{cases} D_t = \int_0^t \mu_d(D_s)ds + \sigma_d(D_s)dW_s, & D_0 = d \in (0, \xi_{max}); \\ E_t = \int_0^t \mu_e(A_s, D_s)ds, & E_0 = 0; \\ A_t = E^Q\left[g(D_T, E_T)\middle|\mathcal{F}_t\right], & A_T = g(D_T, E_T). \end{cases} \tag{2.5}$$

where g is a globally Lipschitz continuous function. For shake of simplicity we omit the martingale measure Q, because we assume to have that measure unless we specificate it. Note that this is an easier case because $\pi e^{-r(T-t)}\mathbb{I}_{[E_{cap},\infty)}(E_T)$ is not a Lipschitz function since it is not even continuous, and this assumption is crucial in the proof which we are going to show. The first equation is independent form the other two, so we can solve it independently, and it is a typical SDE. Moreover, we know from general theory that in our regularity assumption on the coefficients μ_d and σ_d it has a solution. Thus, in the following, we assume the process D_t to be known and $D_t \in \mathbb{L}^2$. On the other hand, the other two equations are coupled, that is they must be solved together because E_t depends on the value of A_s until the time t, and A_t depends on the final value of E_t. In other words the process A_t depends on its own evolution. First of all, we want to associate a PDE to these stochastic equations as we have done for linear stochastic differential equations in appendix B. From above, intuitively, this PDE will not be linear because its coefficients depend on the solution itself. In particular we want to show that the considered PDE exist and it is:

$$\begin{cases} \frac{\sigma_d(x)^2}{2}\frac{\partial^2 u}{\partial x^2} + \mu_d(x)\frac{\partial u}{\partial x} + \mu_e(u,x)\frac{\partial u}{\partial y} + \frac{\partial u}{\partial t} = 0 \text{ in }]0,T] \times \mathbb{R}^2 \\ u(T,x,y) = g(x,y) \quad (x,y) \in \mathbb{R}^2 \end{cases} \tag{2.6}$$

The solution of this semilinear Cauchy problem represent the value of the process A_t and allow us to solve it numerically by a computer. We will show the existence of a viscosity solution of 2.6 in the sense of [10] and we characterize it introducing a vanishing viscosity solution of the regularized problem:

$$\begin{cases} \frac{\sigma_d(x)^2}{2}\frac{\partial^2 u}{\partial x^2} + \varepsilon^2\frac{\partial^2 u}{\partial y^2} + \mu_d(x)\frac{\partial u}{\partial x} + \mu_e(u,x)\frac{\partial u}{\partial y} + \frac{\partial u}{\partial t} = 0 \text{ in }]0,T] \times \mathbb{R}^2 \\ u(T,x,y) = g(x,y) \quad (x,y) \in \mathbb{R}^2 \end{cases} \tag{2.7}$$

with $\varepsilon \in]0,1]$. Then we will study the behaviour of the solution as $\varepsilon \to 0$. Note that in the regularized problem, the PDE is a parabolic second order differential equation, and we cam look for a solution as in [10]. Equation 2.6 is not parabolic because coefficients matrix of the second order derivative is singular, thus it is necessary to regularize the PDE to show the existence of a solution.

Now we use a probabilistic technique to show the existence of a viscosity solution, and it is based on the system of stochastic differential equations 2.5. To obtain the regularized Cauchy problem we introduced a Brownian perturbation, that is we consider a standard Brownian motion B_t $\mathcal{F}_t$-adapted on the filtered probability spaces $(\Omega, \mathcal{F}, P, (\mathcal{F}_t))$ defined in the previous chapter, and a positive constant $\varepsilon \in [0,1[$. Thus we have a perturbed backward stochastic system:

$$\begin{cases} E_t^\varepsilon = \int_0^t \mu_e(A_s^\varepsilon, D_s)ds + \varepsilon B_t, & E_0 = 0; \\ A_t^\varepsilon = E\left[g(D_T, E_T^\varepsilon)\middle|\mathcal{F}_t\right], & A_T = g(D_T, E_T). \end{cases} \tag{2.8}$$

In the following we often use properties of conditional expectation (see [1]) and

Hölder's inequality in the form

$$\left(\int_0^T f(t)dt\right)^2 = T\int_0^T (f(t))^2 dt$$

Theorem 2.1. *Let the foregoing hypothesis hold and let* $\varepsilon \in [0,1[$ *and* $Tk_1(k_2+1) < 1$*, where* k_1 *is the Lipschitz constant of* $\mu_e = \mu_e(x,y)$*, and* k_2 *is the Lipschitz constant of* $g = g(x,y)$ *respect to the second variable. Then there exists a unique solution* $(E_t^\varepsilon, A_t^\varepsilon) \in \mathbb{L}^2 \times \mathbb{L}^2$ *to the system 2.8.*

Proof. The key of this proof is Banach fixed-point theorem (also known as the contraction mapping theorem), thus we define an operator related to the system 2.8, and we show that under suitable condition it is contractive. It implies, thanks to Banach fixed-point theorem, existence and uniqueness of the solution of 2.8. We remark that the space $\mathbb{L}^2$ of adapted process X such that $\|X\|_{\mathbb{L}^2} = E\left[\left(\int_0^T |X_s|^2 ds\right)\right] < \infty$ is a Banach space under the previous norm. Moreover the product space $\mathbb{L}^2 \times \mathbb{L}^2$ is a Banach space under the norm $\|(X,Y)\|_{\mathbb{L}^2\times\mathbb{L}^2} = E\left[\left(\int_0^T (|X_s| + |Y_s|)^2\, ds\right)\right]$. Let (A,E) be in $\mathbb{L}^2 \times \mathbb{L}^2$. Consider the following operator:

$$\begin{aligned}\Lambda(E,A)_t = \begin{pmatrix} F(E,A)_t \\ G(E,A)_t \end{pmatrix} &= \\ &= \begin{pmatrix} \int_0^t \mu_e(A_s^\varepsilon, D_s)ds + \varepsilon B_t \\ E\left[g(D_T, F(E,A)_T)\middle|\mathcal{F}_t\right] \end{pmatrix}\end{aligned}$$

This operator is well defined from $\mathbb{L}^2 \times \mathbb{L}^2$ to itself, as the following shows:

$$\|\Lambda(E,A)_t\|^2_{\mathbb{L}^2\times\mathbb{L}^2} = E\left[\int_0^T \left(|F(E,A)_t| + |G(E,A)_t|\right)^2 dt\right] \tag{2.9}$$

$$\begin{aligned}|F(E,A)_t| &= \left|\int_0^t \mu_e(A_s, D_s)ds + \varepsilon B_t\right| \le \\ &\le \int_0^t |\mu_e(A_s, D_s)|ds + \varepsilon|B_t| \le \\ &\le \int_0^t |\mu_e(A_s, D_s)|ds + \varepsilon|B_t| = \\ &=\le \int_0^t |\mu_e(A_s, D_s) - \mu_e(0,0) + \mu_e(0,0)|ds + \varepsilon|B_t| \le \\ &\le \int_0^t \left(k_1|A_s| + k_1|D_s| + |\mu_e(0,0)|\right) ds + \varepsilon|B_t| \le \\ &\le T|\mu_e(0,0)| + k_1\int_0^t \left(|A_s| + |D_s|\right) ds + \varepsilon|B_t| \le\end{aligned} \tag{2.10}$$

$$
\begin{aligned}
|G(E,A)_t| &= \big|E\left[g(D_T,F(E,A)_T)\big|\mathcal{F}_t\right]\big| \leq \\
&\leq E\left[|g(D_T,F(E,A)_T)-g(D_T,0)+g(D_T,0)|\big|\mathcal{F}_t\right] \leq \\
&\leq E\left[k_2|F(E,A)_T|+|g(D_T,0)|\big|\mathcal{F}_t\right] = \\
&= E\left[k_2\Big|\int_0^T \mu_e(A_s,D_s)ds+\varepsilon B_t\Big|+|g(D_T,0)|\Big|\mathcal{F}_t\right] \leq \\
&\leq E\left[k_2T|\mu_e(0,0)|+k_1k_2\int_0^T\left(|A_s|+|D_s|\right)ds+k_2\varepsilon|B_t|+|g(D_T,0)|\Big|\mathcal{F}_t\right]
\end{aligned} \tag{2.11}
$$

putting 2.10 and 2.11 in 2.9 we obtain:

$$
\begin{aligned}
\|\Lambda(E,A)_t\|^2_{\mathbb{L}^2\times\mathbb{L}^2} \leq & \int_0^T E\Big[\Big(T|\mu_e(0,0)|+k_1\int_0^t\Big(|A_s|+|D_s|\Big)ds+\varepsilon|B_t|+ \\
&+E\Big[k_2T|\mu_e(0,0)|+ \\
&+k_1k_2\int_0^T(|A_s|+|D_s|)ds+k_2\varepsilon|B_t|+|g(D_T,0)|\Big|\mathcal{F}_t\Big]\Big)^2\Big]dt \leq \\
&\leq \int_0^T E\Big[\Big(T^2(k_2+1)^2|\mu_e(0,0)|^2+|g(D_T,0)|^2+\varepsilon^2(k_2+1)^2|B_t|^2+ \\
&+Tk_1^2(k_2^2+1)^2\int_0^T\Big(|A_s|+|D_s|\Big)^2ds\Big]dt<\infty
\end{aligned} \tag{2.12}
$$

Indeed, g is a continuous function and D_t and A_t are stochastic process in $\mathbb{L}^2$. Thus the operator $\Lambda:\mathbb{L}^2\times\mathbb{L}^2\to\mathbb{L}^2\times\mathbb{L}^2$ is well defined. Now we have to show it is a contraction. Let $(E^1,A^1),(E^2,A^2)\in\mathbb{L}^2\times\mathbb{L}^2$ be stochastic processes and we estimate:

$$
\begin{aligned}
|F(E^2,A^2)_t-F(E^1,A^1)_t| &= \Big|\int_0^t(\mu_e(A_s^2,D_s)-\mu_e(A_s^1,D_s))ds\Big| \leq \\
&\leq \int_0^t k_1\Big(|A_s^2-A_s^1|+|D_s-D_s|\Big)ds = \\
&= k_1\int_0^t|A_s^2-A_s^1|ds =
\end{aligned} \tag{2.13}
$$

$$
\begin{aligned}
|G(E^2,A^2)_t-G(E^1,A^1)_t| &= \Big|E[g(D_T,F(E^2,A^2))-g(D_T,F(E^1,A^1)_T)\Big|\mathcal{F}_t\Big]\Big| \leq \\
&\leq E\Big[\Big|g(D_T,F(E^2,A^2))-g(D_T,F(E^1,A^1)_T)\Big|\Big|\mathcal{F}_t\Big] \leq \\
&\leq k_2E\Big[\Big|F(E^2,A^2)_T-F(E^1,A^1)_T\Big|\Big|\mathcal{F}_t\Big] \leq \\
&\leq k_1k_2E\Big[\int_0^T|A_s^2-A_s^1|ds\Big|\mathcal{F}_t\Big]
\end{aligned} \tag{2.14}
$$

Thus, by conditional expectation property, we obtain

$$\begin{aligned}\|\Lambda(E^2,A^2)_t-\Lambda(E^1,A^1)_t\|^2_{\mathbb{L}^2\times\mathbb{L}^2} = &\int_0^T E\Big[\Big(|F(E^2,A^2)_T-F(E^1,A^1)_T|+\\ &+|G(E^2,A^2)_T-G(E^1,A^1)_T|\Big)^2\Big]dt \leq\\ &\leq \int_0^T E\Big[\Big(k_1(k_2+1)\int_0^T|A^2_s-A^1_s|ds\Big)^2\Big] \leq\\ &\leq k_1^2T^2(k_2+1)^2E\Big[\int_0^T\Big(|E^2_s-E^1_s|+|A^2_s-A^1_s|\Big)^2ds\Big] \leq\\ &\leq (k_2+1)^2k_1^2T^2\|(E^2,A^2)-(E^1,A^1)\|^2_{\mathbb{L}^2\times\mathbb{L}^2}\end{aligned} \tag{2.15}$$

Which imply

$$\begin{aligned}\|\Lambda(E^2,A^2)_t-\Lambda(E^1,A^1)_t\|_{\mathbb{L}^2\times\mathbb{L}^2} \leq (k_2+1)k_1T\\ |(E^2,A^2)-(E^1,A^1)\|_{\mathbb{L}^2\times\mathbb{L}^2}\end{aligned} \tag{2.16}$$

By hypothesis we have $Tk_1(k_2+1)<1$ and we conclude that Λ is a contraction, and we can apply Banach fixed-point theorem. Thus it exist an unique solution $(E^\varepsilon_t,A^\varepsilon_t)\in\mathbb{L}^2\times\mathbb{L}^2$. □

Remark 2.2. Analogously we can show existence and uniqueness of the solution of the unperturbed problem

$$\begin{cases} E_t=\int_0^t\mu_e(A_s,D_s)ds, & E_0=0;\\ A_t=E^Q\left[g(D_T,E_T)\middle|\mathcal{F}_t\right], & A_T=g(D_T,E_T).\end{cases} \tag{2.17}$$

because the the proof is the same except of the perturbation term εB_t, and this does not affect the scheme of our proof and the final result. Thus, if $Tk_1(k_2+1)<1$, then the coupled system of stochastic differential equations 2.17 has unique solution $(E_t,A_t)\in\mathbb{L}^2\times\mathbb{L}^2$.

Remark 2.3. Note that in the above theorem we assume that $Tk_1(k_2+1)<1$, then it must be $T<\frac{1}{k_1(k_2+1)}$, and it means we have an unique solution $(E_t,A_t)\in\mathbb{L}^2\times\mathbb{L}^2$ of problem 2.17 only for suitable small T. Thus we have study the problem only for time interval $[0,T]$ small enough.

Remark 2.4. The bound on the norm of the solution $(E_t^\varepsilon, A_t^\varepsilon) \in \mathbb{L}^2 \times \mathbb{L}^2$ of the perturbed problem 2.8 can be made independent of ε. Indeed we can write the following inequalities:

$$
\begin{aligned}
&|E_t^\varepsilon| \leq k_1 \int_0^t (|A_s^\varepsilon| + |D_s|)ds + \varepsilon|B_t| + T|\mu_e(0,0)| \\
&|A_t^\varepsilon| \leq E\Big[k_2 T|\mu_e(0,0)| + |g(D_T,0)| + k_2\varepsilon|B_T| + k_1k_2\int_0^T (|A_s^\varepsilon| + |D_s|)ds\Big|\mathcal{F}_t\Big]
\end{aligned}
\tag{2.18}
$$

Putting together the foregoing inequalities and since $\varepsilon < 1$ we have

$$
\begin{aligned}
|E_t^\varepsilon| + |A_t^\varepsilon| \leq E\Big[&T(k_2+1)|\mu_e(0,0)| + |g(D_t,0)| + |B_t| + k_2|B_T| + \\
&+ k_1(k_2+1)\int_0^T |D_s|ds + k_1(k_2+1)\int_0^T (|E_s| + |A_s|)ds\Big|\mathcal{F}_t\Big]
\end{aligned}
\tag{2.19}
$$

We remember the following equalities:

$$
\begin{aligned}
&E[\int_0^T |B_T|^2 dt] = T\cdot T = T^2 \\
&E[\int_0^T |B_t|^2 dt] = \int_0^T E[B_t]^2 dt = \int_0^T t dt = \frac{1}{2}T^2
\end{aligned}
\tag{2.20}
$$

and Schwartz inequality of the form

$$
(\alpha+\beta)^2 \leq (1+\frac{1}{a})\alpha^2 + (1+a)\beta^2 \tag{2.21}
$$

where $a > 0$ if a suitable large constant.

Squaring both sides of 2.19 and applying 2.21 we get

$$
\Big(|E_t^\varepsilon| + |A_t^\varepsilon|\Big)^2 \leq E\Big[(k_2+1)^2 k_1^2 T\Big(1+\frac{1}{a}\Big)\int_0^T \Big(|E_s^\varepsilon| + |A_s^\varepsilon|\Big)^2 ds + (1+a)\beta^2\Big|\mathcal{F}_t\Big]
\tag{2.22}
$$

where

$$
\beta^2 = \Big(T(k_2+1)|\mu_e(0,0)| + |g(D_t,0)| + |B_t| + k_2|B_T| + k_1(k_2+1)\int_0^T |D_s|ds\Big)^2
\tag{2.23}
$$

Integrating both sides of 2.22 from 0 to T and taking expectation we get

$$
E\Big[\int_0^T \Big(|E_t^\varepsilon| + |A_t^\varepsilon|\Big)^2 dt\Big] \leq E\Big[(k_2+1)^2 k_1^2 T^2\Big(1+\frac{1}{a}\Big)\int_0^T \Big(|E_s^\varepsilon| + |A_s^\varepsilon|\Big)^2 ds + (1+a)\beta^2\Big]
\tag{2.24}
$$

And it implies

$$E\Big[\int_0^T \Big(|E_t^\varepsilon|+|A_t^\varepsilon|\Big)^2 dt\Big] \le \frac{(1+a)}{1-(k_2+1)^2k_1^2T^2\Big(1+\frac{1}{a}\Big)} E\Big[\int_0^T \beta^2 dt\Big] \tag{2.25}$$

with

$$\begin{aligned} E\Big[\int_0^T \beta^2 dt\Big] \le E\Big[&(k_2+1)^2T^3|\mu_e(0,0)|^2 + T|g(D_t,0)|^2 + \Big(k_2^2+\frac{1}{2}\Big)T^2+ \\ &+ (k_2+1)^2k_1^2T^2\|D_t\|_{\mathbb{L}^2}^2\Big] \end{aligned} \tag{2.26}$$

and this estimate is independent of ε. Moreover, taking expectation of 2.22 and plugging 2.25 back into 2.22 and we have

$$\begin{aligned} E\Big[\Big(|E_t^\varepsilon|+|A_t^\varepsilon|\Big)^2\Big] &\le (k_2+1)k_1^2T\Big(1+\frac{1}{a}\Big)\frac{(1+a)}{1-(k_2+1)^2k_1^2T^2\Big(1+\frac{1}{a}\Big)}E\Big[\int_0^T \beta^2 dt\Big]+ \\ &+ (1+a)E\Big[\beta^2\Big] \le \\ &\le C\left(k_1,k_2,T,a,g,\mu_e,D_t,\frac{1}{1-\Big(1+\frac{1}{a}\Big)(k_2+1)^2k_1^2T^2}\right) \end{aligned} \tag{2.27}$$

In conclusion, thanks to Doob's inequality for submartingales (see [13]) we have also the following bound

$$\begin{aligned} E\Big[\sup_{0\le t\le T}\Big(|E_t^\varepsilon|+|A_t^\varepsilon|\Big)^2\Big] &\le 2\sup_{0\le t\le T} E\Big[\Big(|E_t^\varepsilon|+|A_t^\varepsilon|\Big)^2\Big] \le \\ &\le 2C\left(k_1,k_2,T,a,g,\mu_e,D_t,\frac{1}{1-\Big(1+\frac{1}{a}\Big)(k_2+1)^2k_1^2T^2}\right) \end{aligned} \tag{2.28}$$

which is independent of ε.

We have shown existence of an adapted solution $(E_t^\varepsilon, A_t^\varepsilon)$ of 2.8. We now apply Brownian Martingale representation Theorem which is exhibited in appendix A to the process A_t. We can do this because it is a Brownian martingale. Thus we can write the backward component of our system as two stochastic integrals of predictable processes Z_t^ε, where Z_t^ε is relative to the Brownian motion B_t and H_t^ε to W_t.

$$A_t^\varepsilon = g(D_T,E_T) - \int_t^T H_s^\varepsilon dW_s - \int_t^T Z_s^\varepsilon dB_s \tag{2.29}$$

Moreover

$$E\Big[\int_0^T [(H_s^\varepsilon)^2 + (Z_s^\varepsilon)^2]ds\Big] < \infty$$

With this representation, it follows the continuity in t of the process A_t, because the stochastic integrals in 2.29 are continuous in t.

As we can see in [3], under our hypothesis (Brownian environment and g deterministic function), the solution of a stochastic system is a Markov process. It follows that the solution $(E_t^\varepsilon, A_t^\varepsilon)$ of the system 2.8 is a couple of Markov processes. Hence, for all $(t,x,y) \in [0,T] \times \mathbb{R}^2$, the associated flows of solution

$$\begin{cases} D_s^{t,x} = x + \int_t^s \mu_d(D_r^{t,x})dr + \int_t^s \sigma_d(D_r^{t,x})dW_r \\ E_s^{\varepsilon,t,x,y} = y + \int_t^s \mu_e(A_r^{\varepsilon,t,x,y}, D_r^{t,x}) + \varepsilon(B_s - B_r) \\ A_s^{\varepsilon,t,x,y} = E\Big[g(D_T^{t,x}, E_T^{\varepsilon,t,x,y})\Big|\mathcal{F}_s\Big] \end{cases} \tag{2.30}$$

define a deterministic function

$$u^\varepsilon(t,x,y) = A_t^{\varepsilon,t,x,y} \tag{2.31}$$

We can define this deterministic function thanks to Blumenthal's $0-1$ law (C.8) and the Markov property of solution processes. In particular, Blumenthal's $0-1$ law is a dichotomic law which implies that $A_t^{\varepsilon,t,x,y}$ is a constant random variable $P^{(x,y)}$-almost surely on $(\Omega, \widetilde{\mathcal{F}}_t, P^{(x,y)})$, where $\widetilde{\mathcal{F}}_t$ is a σ-algebra of the universal filtration (see C). Thus, for any $t \in [0,T]$ and $(x,y) \in \mathbb{R}^2$ it is a deterministic function. The following proposition show the Hölder regularity of the function $u^\varepsilon(t,x,y)$.

Proposition 2.5. *Under the above hypothesis u^ε is globally Lipschitz in x, y and Hölder of order $\frac{1}{2}$ in t with constant C_0 independent of $\varepsilon \in [0,1[$. In compact form we write:*

$$\begin{aligned} |u^\varepsilon(t_2,x_2,y_2) - u^\varepsilon(t_2,x_1,y_1)| &\le C_0(|x_2 - x_1| + |y_2 - y_1|) \\ |u^\varepsilon(t_2,x_2,y_2) - u^\varepsilon(t_1,x_2,y_2)| &\le \widetilde{C_0}(1 + |(x_1,y_1)|)|t_2 - t_1|^{\frac{1}{2}} \end{aligned} \tag{2.32}$$

Proof. Take $t_1, t_2 \in [0,T]$, and without loss of generality assume that $t_1 \le t_2$, then consider $(x_1,y_1),(x_2,y_2) \in \mathbb{R}^2$. We can define the flows associated to these starting point. We can extend naturally the flows on the whole interval as follows

$$(D_s^{t_i,x_i}, E_s^{\varepsilon,t_i,x_i,y_i}, A_s^{\varepsilon,t_i,x_i,y_i}) = (D_{t_i}^{t_i,x_i}, E_{t_i}^{\varepsilon,t_i,x_i,y_i}, A_{t_i}^{\varepsilon,t_i,x_i,y_i})$$

for any $s \leq t_i$, $i = 1, 2$. Our target is to estimate $|A^{\varepsilon,t_2,x_2,y_2}_{t_2} - A^{\varepsilon,t_1,x_1,y_1}_{t_1}|$. For simplicity, we use the convenient notation $X^i = X^{\varepsilon,t_i,x_i,y_i}$ for any process that appears in this proof. Moreover we remember that we denote respectively by k_1,k_2,k_3,k_4 the Lipschitz constant of μ_e,g,μ_d,σ_d. For any $t \in [0, T]$ we have

$$\begin{aligned}
|D^2_t - D^1_t| &\leq |x_2 - x_1| + \int_{t_2}^{t_2\vee t} |\mu_d(D^2_s) - \mu_d(D^1_s)|ds + \int_{t_1\wedge t}^{t_2\wedge t} |\mu_d(D^1_s)|ds+ \\
&\quad + \int_{t_2}^{t_2\vee t} |\sigma_d(D^2_s) - \sigma_d(D^1_s)|dW_s + \int_{t_1\wedge t}^{t_2\wedge t} |\sigma_d(D^1_s)|dW_s \\
|E^2_t - E^1_t| &\leq |y_2 - y_1| + \int_{t_2}^{t_2\vee t} |\mu_e(A^2_s, D^2_s) - \mu_e(A^1_s, D^1_s)|ds + \int_{t_1\wedge t}^{t_2\wedge t} |\mu_e(A^1_s, D^1_s)|ds+ \\
&\quad + \varepsilon|W_{t_2\vee t} - W_{t_2} - W_{t_1\vee t} + W_{t_1}| \\
|A^2_t - A^1_t| &\leq E\Big[|g(D^2_T, E^2_T) - g(D^1_T, E^1_T)|\mathcal{F}_t\Big]
\end{aligned} \tag{2.33}$$

Putting together $|E^2_t - E^1_t|$ and $|A^2_t - A^1_t|$ and squaring both sides we get

$$\begin{aligned}
(|E^2_t - E^1_t| + |A^2_t - A^1_t|)^2 \leq \Big\{ E\Big[&(k_2 + 1)|y_2 - y_1| + k_1(k_2 + 1)\int_0^T |A^2_s - A^1_s|ds+ \\
&+ k_1(k_2 + 1)\int_0^T |D^2_s - D^1_s|ds + k_1(k_2 + 1)\int_{t_1}^{t_2} |\mu_e(A^1_s, D^1_s)|ds+ \\
&+ (k_2 + 1)\varepsilon|W_{t_2\vee t} - W_{t_2} - W_{t_1\vee t} + W_{t_1}|\Big|\mathcal{F}_t\Big]\Big\}^2
\end{aligned} \tag{2.34}$$

As we have done before in 2.24, we apply Shwartz inequality for a suitable large $a > 0$, take expected value and we integrate from 0 to T. Thus we obtain

$$E\Big[\int_0^T \Big(|E^2_t - E^1_t| + |A^2_t - A^1_t|\Big)^2 dt\Big] \leq \frac{(1 + a)T}{1 - (k_2 + 1)^2 k_1^2 T^2\Big(1 + \frac{1}{a}\Big)} E\Big[A^2\Big] \tag{2.35}$$

where

$$\begin{aligned}
E\Big[A^2\Big] &\leq (k_2 + 1)^2|y_2 - y_1|^2 + (k_2 + 1)^2(t_2 - t_1)E\Big[\int_{t_1}^{t_2} |\mu_e(A^1_s, D^1_s)|^2 ds\Big]+ \\
&\quad + (k_2 + 1)^2\varepsilon^2 T|B_{t_2} - Bt_1|^2 + Tk_1^2(k_2 + 1)^2 E\Big[\int_0^T |D^2_s - D^1_s|^2\Big]
\end{aligned} \tag{2.36}$$

From 2.34 we can estimate $E\Big[\int_0^T |D_s^2 - D_s^1|^2\Big]$

$$E\Big[\int_0^T |D_s^2 - D_s^1|^2\Big] \leq T|x_2 - x_1|^2 + T^2k_3^2E\Big[\int_0^T |D_s^2 - D_s^1|^2\Big] + Tk_4^2E\Big[\int_0^T |D_s^2 - D_s^1|^2\Big] + \\ + T(t_2 - t_1)E\Big[\int_{t_1}^{t_2} |\sigma_d(D_s^1)|^2\Big] \leq \\ \leq T|x_2 - x_1|^2 + T(t_2 - t_1)E\Big[\int_{t_1}^{t_2} |\sigma_d(D_s^1)|^2\Big] + \\ + T(Tk_3^2 + k_4^2)E\Big[\int_0^T |D_s^2 - D_s^1|^2\Big] \tag{2.37}$$

And this implies

$$E\Big[\int_0^T |D_s^2 - D_s^1|^2\Big] \leq \frac{T}{T(1 - Tk_3^2 + k_4^2)}\Big[|x_2 - x_1|^2 + (t_2 - t_1)E\Big[\int_{t_1}^{t_2} |\sigma_d(D_s^1)|^2\Big]\Big] \leq \\ \leq TC_1(k_3, k_4, T)|x_2 - x_1|^2 + TC_1(k_3, k_4, T)(t_2 - t_1)E\Big[\int_{t_1}^{t_2} |\sigma_d(D_s^1)|^2\Big] \tag{2.38}$$

Plugging the last inequality back in 2.36 we get

$$E\Big[A^2\Big] \leq (k_2 + 1)^2|y_2 - y_1|^2 + (k_2 + 1)^2(t_2 - t_1)E\Big[\int_{t_1}^{t_2} |\mu_e(A_s^1, D_s^1)|^2 ds\Big] + \\ + (k_2 + 1)^2\varepsilon^2T|B_{t_2} - Bt_1|^2 + C_1T^2k_1^2(k_2 + 1)^2|x_2 - x_1|^2 + \\ + C_1T^2k_1^2(k_2 + 1)^2E\Big[\int_{t_1}^{t_2} |\sigma_d(D_s^1)|^2\Big] \leq \\ \leq C_2(x_1, y_1, k_1, k_2, k_3, k_4, \mu_e, \sigma_d, T, a)(|t_2 - t_1| + |x_2 - x_1|^2 + |y_2 - y_1|^2) \tag{2.39}$$

Thus

$$E\Big[\int_0^T \Big(|E_t^2 - E_t^1| + |A_t^2 - A_t^1|\Big)^2 dt\Big] \leq \\ C_3\left(x_1, y_1, k_1, k_2, k_3, k_4, \mu_e, \sigma_d, T, a, \frac{1}{1 - \Big(1 + \frac{1}{a}\Big)(k_2 + 1)^2k_1^2T^2}\right) \cdot \\ \cdot (|t_2 - t_1| + |x_2 - x_1|^2 + |y_2 - y_1|^2) \tag{2.40}$$

Finally, using the properties of Brownian motion and the fact that $\varepsilon < 1$, proceeding as we have done before in remark 2.4 we obtain the following estimate

$$E\Big[\sup_{0\le t\le T}(|E_t^2-E_t^1|+|A_t^2-A_t^1|)^2\Big]\le$$
$$\le C_4\left(x_1,y_1,k_1,k_2,k_3,k_4,\mu_e,\sigma_d,T,a,\frac{1}{1-\left(1+\frac{1}{a}\right)(k_2+1)^2k_1^2T^2}\right)\cdot$$
$$\cdot(|t_2-t_1|+|x_2-x_1|^2+|y_2-y_1|^2) \tag{2.41}$$

Since the last estimate hold uniformly in t, it is also true for t_1, hence we get

$$\begin{aligned}
|u^\varepsilon(t_2,x_2,y_2)-u^\varepsilon(t_2,x_1,y_1)|^2 &\le \\
&\le |A_{t_2}^{\varepsilon,t_2,x_2,y_2}-A_{t_2}^{\varepsilon,t_2,x_1,y_1}|^2\le \\
&\le C_4(|x_2-x_1|^2+|y_2-y_1|^2)\le \\
&\le C_0^2(|x_2-x_1|+|y_2-y_1|)^2 \\
|u^\varepsilon(t_2,x_2,y_2)-u^\varepsilon(t_1,x_2,y_2)|^2 &\le \\
&\le |A_{t_2}^{\varepsilon,t_2,x_2,y_2}-A_{t_1}^{\varepsilon,t_1,x_2,y_2}|^2\le \\
&\le C_4|t_2-t_1|\le \\
&\le \widetilde{C_0}^2(1+|(x_1,y_1)|)^2|t_2-t_1|
\end{aligned} \tag{2.42}$$

And we conclude

$$\begin{aligned}
|u^\varepsilon(t_2,x_2,y_2)-u^\varepsilon(t_2,x_1,y_1)| &\le C_0(|x_2-x_1|+|y_2-y_1|) \\
|u^\varepsilon(t_2,x_2,y_2)-u^\varepsilon(t_1,x_2,y_2)| &\le \widetilde{C_0}(1+|(x_1,y_1)|)|t_2-t_1|^{\frac{1}{2}}
\end{aligned} \tag{2.43}$$

□

The next proposition shows that the function u^ε just defined is a viscosity solution of

$$\begin{cases}
\frac{\sigma_d(x)^2}{2}\frac{\partial^2 u}{\partial x^2}+\varepsilon^2\frac{\partial^2 u}{\partial y^2}+\mu_d(x)\frac{\partial u}{\partial x}+\mu_e(u,x)\frac{\partial u}{\partial y}+\frac{\partial u}{\partial t}=0 \text{ in } [0,T[\times\mathbb{R}^2 \\
u(T,x,y)=g(x,y) \quad (x,y)\in\mathbb{R}^2
\end{cases} \tag{2.44}$$

and to do this we use Itô's formula on u^ε.

Proposition 2.6. *Let $\varepsilon\in]0,1]$. Then the function u^ε is a viscosity solution of the problem 2.44.*

Proof. In the previous proposition we have shown the continuity of u^ε, thus now we only have to show it is both a viscosity subsolution and supersolution. We only show the subsolution case because the the proof of the supersolution case is analogous.

We have observed that $A_s^{\varepsilon,t,x,y}$ is a Markov process and from the pathwise uniqueness of the solution process of the problem 2.8 we have $u^\varepsilon(s, D_s^{t,x}, E_s^{\varepsilon,t,x,y})$ as surely for any $s \in [t, T]$.

Let us consider a point $(t, x, y) \in [0, T] \times \mathbb{R}^2$ and a function $\varphi \in C^{1,2}([0, T] \times \mathbb{R}^2)$, with bounded derivatives, and such that

$$u^\varepsilon(t, x, y) - \varphi(t, x, y) = 0$$

Moreover we assume, without loss of generality, that (t, x, y) is a global maximum for $u^\varepsilon(t, x, y) - \varphi(t, x, y)$. In this hypothesis, for any $\mathcal{F}_t$-stopping time τ, we have

$$u^\varepsilon(\tau, D_\tau^{t,x}, E_\tau^{\varepsilon,t,x,y}) - \varphi(\tau, D_\tau^{t,x}, E_\tau^{\varepsilon,t,x,y}) \leq 0 \tag{2.45}$$

In order to simplify the notation in the following of this proof we omit the superscript of u, D, E, A. Since φ is regular, it satisfy the hypothesis of itô's formula, and we apply it in the interval $[t, \tau]$, with τ $\mathcal{F}_t$-stopping time. Thus we get

$$\begin{aligned} \varphi(\tau, D_\tau, E_\tau) = \varphi(t, x, y) + \int_t^\tau \Big(& \varphi_t(\tau, D_r, E_r) + \mu_d(D_r)\varphi_x(\tau, D_r, E_r) + \\ & + \mu_e(u(r, D_r, E_r), D_r)\varphi_y(\tau, D_r, E_r) + \frac{\sigma_d^2(D_r)}{2}\varphi_x x(\tau, D_r, E_r) + \\ & + \frac{\varepsilon^2}{2}\varphi_y y(\tau, D_r, E_r)\Big) dr + \int_t^\tau \sigma_d(D_r)\varphi_x(\tau, D_r, E_r) dW_r + \\ & + \int_t^\tau \varepsilon\varphi_y(\tau, D_r, E_r) dB_r \end{aligned} \tag{2.46}$$

and thanks to the Brownian Martingale Representation Theorem we have

$$\begin{aligned} u(t, x, y) = A_t = A_\tau - \int_t^\tau Z_r dW_r = \\ = u(\tau, D_\tau^{t,x}, E_\tau) - \int_t^\tau Z_r dW_r \end{aligned} \tag{2.47}$$

which imply

$$u(\tau, D_\tau^{t,x}, E_\tau) = u(t, x, y) + \int_t^\tau Z_r dW_r \tag{2.48}$$

Substituting equalities 2.46 and 2.48 in 2.45 we obtain

$$\begin{aligned} u^{\varepsilon}(\tau, D_\tau, E_\tau) - \varphi(\tau, D_\tau, E_\tau) = u(t,x,y) - \varphi(t,x,y) + \int_t^\tau Z_r dW_r - \\ - \int_t^\tau \Big(\varphi_t(\tau, D_r, E_r) + \mu_d(D_r)\varphi_x(\tau, D_r, E_r) + \\ + \mu_e(u(r, D_r, E_r), D_r)\varphi_y(\tau, D_r, E_r) + \\ + \frac{\sigma_d^2(D_r)}{2}\varphi_x x(\tau, D_r, E_r) + \\ + \frac{\varepsilon^2}{2}\varphi_y y(\tau, D_r, E_r)\Big) dr - \frac{\varepsilon^2}{2}\varphi_y y(\tau, D_r, E_r)\Big) dr + \\ + \int_t^\tau \sigma_d(D_r)\varphi_x(\tau, D_r, E_r) dW_r - \\ - \int_t^\tau \varepsilon\varphi_y(\tau, D_r, E_r) dB_r \leq 0 \end{aligned} \tag{2.49}$$

Now we take expectations in the previous equality, and since the martingale part do not give any contribute we get

$$E\Big[\Phi(\tau, D_r, E_r)\Big] \geq 0 \tag{2.50}$$

where

$$\Phi = \frac{\sigma_d^2(\cdot)}{2}\varphi_x x + \frac{\varepsilon^2}{2}\varphi_y y + \mu_d(x)\varphi_x + \mu_e(u,\cdot)\varphi_y(\tau, D_r, E_r) + \varphi_t \tag{2.51}$$

Since the equality is verified at time T, because of the definition of A, to show that u is a viscosity subsolution of 2.44 we must verify $\Phi(t,x,y) \geq 0$ (see remark 2.7). Note that the coefficient of the second order derivative respect to the variable y is $\frac{\varepsilon}{2}$ instead of ε, but imposing $\widetilde{\varepsilon} = \frac{\varepsilon}{\sqrt{2}}$ we obtain the same equation.
By contradiction we assume $\Phi(t,x,y) < 0$, that is we assume there exist an $\delta < 0$ such that $\Phi(t,x,y) < \delta$. We define the $\mathcal{F}_t$-stopping time τ_1 as follows

$$\tau_1 = \inf\left\{ r > t \middle| \Phi(r, D_r, E_r) \geq \frac{\delta}{2} \right\} \wedge T$$

By construction $\tau_1 > t$ almost surely, and since Inequality 2.50 holds for any stopping time, we have

$$0 > \frac{\delta}{2}E(\tau_1 - t) \geq E\Big[\Phi(\tau, D_r, E_r)\Big] \geq 0$$

and this is clearly a contradiction. Thus $\Phi(t,x,y) \geq 0$ and u is a viscosity subsolution of 2.44. Analogously it can be shown that u is a viscosity supersolution of 2.44, hence u is a viscosity solution of 2.44 and this complete the proof. □

Remark 2.7. In [10], the function $\Phi(t,x,y)$ is asked to be ≤ 0, but in the hypothesis of the viscosity subsolution definition the sign of the of the second order derivatives is " $-$ ". In our case the sign is " $+$ ", thus to show that u is a viscosity subsolution we have to change sign and we have to verify $\Phi \geq 0$.

We have just proved the existence of a viscosity solution of the problem 2.44, and in the next statement we will show uniqueness of this viscosity solution.

Proposition 2.8. *Le $\varepsilon \in [0,1[$. If u is a subsolution and v a is a supersolution of problem 2.44 such that both verify Hölder estimate 2.32, then $u \leq v$.*

Proof. We set $S_\varrho =]0,\varrho[\times\mathbb{R}^2$ and consider the function

$$H(t,h) = exp\left(\frac{|h|^2}{1-(2\varrho)^{-1}t} - \sigma t\right) \quad (t,h) \in \widetilde{S_\varrho} \tag{2.52}$$

We compute the following derivatives:

$$\begin{aligned}
H_t &= H\left(\frac{x^2+y^2}{2\varrho(1-(2\varrho)^{-1}t)^2} - \sigma\right)\\
H_x &= H\left(\frac{2x}{1-(2\varrho)^{-1}t}\right)\\
H_y &= H\left(\frac{2y}{1-(2\varrho)^{-1}t}\right)\\
H_{xx} &= H\left(\frac{4x^2}{1-(2\varrho)^{-1}t} + \frac{2}{1-(2\varrho)^{-1}t}\right)\\
H_{yy} &= H\left(\frac{4y^2}{1-(2\varrho)^{-1}t} + \frac{2}{1-(2\varrho)^{-1}t}\right)
\end{aligned}$$

and we have

$$\begin{aligned}
\widehat{H}(\varepsilon,t,h) = &\frac{\frac{\sigma_d^2}{2}H_{xx} + \varepsilon^2 H_{yy} + (\mu_e(u,x)+\mu_e(v,x))H_y + \mu_d(x)H_x + H_t}{H} =\\
&= \frac{\sigma_d^2}{2}\left(\frac{4x^2}{1-(2\varrho)^{-1}t} + \frac{2}{1-(2\varrho)^{-1}t}\right) + \varepsilon^2\left(\frac{4y^2}{1-(2\varrho)^{-1}t} + \frac{2}{1-(2\varrho)^{-1}t}\right) +\\
&+ (\mu_e(u,x)+\mu_e(v,x))\left(\frac{2y}{1-(2\varrho)^{-1}t}\right) + \mu_d(x)\left(\frac{2x}{1-(2\varrho)^{-1}t}\right) +\\
&+ \frac{x^2+y^2}{2\varrho(1-(2\varrho)^{-1}t)^2} - \sigma
\end{aligned} \tag{2.53}$$

Since μ_e and μ_d are globally Lipschitz and u, v verify estimates 2.32,it is possible to choose sufficiently large positive constants ϱ^{-1},σ such that for every $\varepsilon \in [0, 1[$

$$\sup_{S_\varrho} \widehat{H}(\varepsilon, t, h) < 0 \tag{2.54}$$

We want to prove that $u \leq v$ in $\widehat{H}(S_\varrho)$, thus we suppose, by contradiction, that there exist $\overline{z} \in S_\varrho$ such that $u(\overline{z}) - v(\overline{z}) > 0$.
We define the following functions defined on $[0, \varrho[\times\mathbb{R}^2$

$$w = \frac{u}{H} - \frac{\delta}{\varrho - t}$$
$$\omega = \frac{v}{H} + \frac{\delta}{\varrho - t}$$

and we take $\delta > 0$ suitably small such that $w(\overline{z}) - \omega(\overline{z}) = \frac{u(\overline{z})-v(\overline{z})>0}{H(\overline{z})} - \frac{2\delta}{\varrho - t} > 0$.
We have

$$\lim_{|h|\to\infty} (w - \omega)(t, h) = -\frac{2\delta}{\varrho - t} < 0 \tag{2.55}$$

and

$$\lim_{|h|\to\varrho^-} (w - \omega)(t, h) = -\infty \tag{2.56}$$

uniformly in $h \in \mathbb{R}^2$. Now we can double the number of spatial variables and, for $\alpha > 0$, consider the function

$$\Phi_\alpha(t, h, h') = w(t, h) - \omega(t, h') - \frac{\alpha}{2}|h - h'|^2 \tag{2.57}$$

Let $(t_\alpha, h_\alpha, h'_\alpha)$ be a maximum point of Φ_α in $[0, \varrho[\times\mathbb{R}^2$. The maximum is achieved in view of upper semicontinuity of $w - \omega$, compactness of $[0, \varrho] \times \mathbb{R}^2$ and 2.55,2.55 (see [4]). Moreover we have

$$0 < w(\overline{z}) - \omega(\overline{z}) \leq \Phi_\alpha(t_\alpha, h_\alpha, h'_\alpha) \leq \sup_{S_\varrho}(w - \omega) < +\infty \tag{2.58}$$

By proposition D.2 in appendix D, we get

$$\lim_{\alpha\to\infty} \alpha|h_\alpha - h'_\alpha|^2 = 0 \tag{2.59}$$

so that, by 2.55 and 2.58,there exist a compact subset M of $\mathbb{R}^2$ such that $h_\alpha, h'_\alpha \in M$ for every $\alpha > 0$. Hence we may suppose that there exist the limit

$$\lim_{\alpha\to\infty} (t_0, h_0, h'_0) \in [0, \varrho] \times \mathbb{R}^2 \times \mathbb{R}^2 \tag{2.60}$$

If $t_0 = 0$, then $\Phi_\alpha(t_\alpha, h_\alpha, h'_\alpha) \to -2\delta\varrho^{-1}$ and this contradicts 2.58. Hence $t_\alpha > 0$ if α is large. Analogously, by 2.56 and 2.58, $t_0 < \delta$. Then D.2 in appendix D yields

$$\lim_{\alpha\to\infty} \Phi_\alpha(t_\alpha, h_\alpha, h'_\alpha) = w(t_0, h_0) - \omega(t_0, h_0) = \sup_{[0,\varrho[\times\mathbb{R}^2} \tag{2.61}$$

Thus we may apply theorem D.6 in appendix D to infer that there exist $a \in \mathbb{R}$ and some matrices X^w, Y^ω such that

$$\begin{aligned} &(a, \alpha(h_\alpha - h'_\alpha), X^w) \in P^{2,+}_{S_\varrho} w(t_\alpha, h_\alpha) \\ &(a, \alpha(h_\alpha - h'_\alpha), Y^\omega) \in P^{2,-}_{S_\varrho} \omega(t_\alpha, h'_\alpha) \\ &X^w \le Y^\omega \end{aligned} \tag{2.62}$$

Since

$$\begin{aligned} u &= \left(w + \frac{\delta}{\varrho - t}\right) H \\ v &= \left(\omega - \frac{\delta}{\varrho - t}\right) H \end{aligned} \tag{2.63}$$

by theorem D.4 in appendix D we deduce that

$$\begin{aligned} &(d^u_t, (d^u_x, d^u_y), X^u) \in P^{2,+}_{S_\varrho} w(t_\alpha, h_\alpha) \\ &(d^v_t, (d^v_x, d^v_y), Y^v) \in P^{2,-}_{S_\varrho} w(t_\alpha, h'_\alpha) \end{aligned} \tag{2.64}$$

where

$$\begin{aligned} d^u_t &= \left(\left(a - \frac{\delta}{(\varrho - t)^2}\right) H + \frac{u}{H} H_t\right)(t_\alpha, h_\alpha) \\ (d^u_x, d^u_y) &= \left(\alpha(h_\alpha - h'_\alpha) H + \frac{u}{H} D_h H\right)(t_\alpha, h_\alpha) \\ X^u &= \left(X^w H + 2\alpha(h_\alpha - h'_\alpha) \otimes D_h H + \frac{u}{H} D^2_h H\right)(t_\alpha, h_\alpha) \end{aligned} \tag{2.65}$$

and

$$\begin{aligned} d^u_t &= \left(\left(a - \frac{\delta}{(\varrho - t)^2}\right) H + \frac{v}{H} H_t\right)(t_\alpha, h'_\alpha) \\ (d^v_x, d^v_y) &= \left(\alpha(h_\alpha - h'_\alpha) H + \frac{v}{H} D_h H\right)(t_\alpha, h'_\alpha) \\ Y^v &= \left(Y^\omega H + 2\alpha(h_\alpha - h'_\alpha) \otimes D_h H + \frac{v}{H} D^2_h H\right)(t_\alpha, h'_\alpha) \end{aligned} \tag{2.66}$$

Next, since u is a subsolution of 2.44, we get

$$\begin{aligned} &\frac{\sigma^2_d(x_\alpha)}{2} X^u_{11} + \varepsilon^2 X^u_{22} + \mu_d(x_\alpha) d^u_x + \mu_e(u(t_\alpha, h_\alpha), x_\alpha) d^u_y + d^u_t + \\ &\quad + \mu_e(u(t_\alpha, h_\alpha), x_\alpha) d^v_y \le \mu_e(u(t_\alpha, h_\alpha), x_\alpha) d^v_y \end{aligned} \tag{2.67}$$

or, in other terms

$$\frac{\sigma_d^2(x_\alpha)}{2}X_{11}^w + \varepsilon^2 X_{22}^w + 2\alpha(x_\alpha - x'_\alpha)\left[\frac{\sigma_d^2}{2}(x_\alpha)\frac{H_x}{H}(t_\alpha, h_\alpha) + \frac{\mu_d(x_\alpha}{2})\right] +$$
$$+ \alpha(y_\alpha - y'_\alpha)\left[\frac{2\varepsilon^2 H_y}{H}(t_\alpha, h_\alpha) + \mu_e(u(t_\alpha, h_\alpha), x_\alpha)\right] + a + \frac{\delta}{(\varrho - t_\alpha)^2} +$$
$$+ \frac{u}{H^2}(t_\alpha, h_\alpha)\Big[\frac{\sigma_d^2}{2}(x_\alpha)H_{xx}(t_\alpha, h_\alpha) + \varepsilon^2 H_{yy}(t_\alpha, h_\alpha) + \mu_e(u(t_\alpha, h_\alpha), x_\alpha)H_y(t_\alpha, h_\alpha) +$$
$$+ \mu_d(x_\alpha)H_x(t_\alpha, h_\alpha) + H_t(t_\alpha, h_\alpha)\Big] +$$
$$+ \mu_e(u(t_\alpha, h_\alpha), x_\alpha)\left[\alpha(y_\alpha - y'_\alpha) + \frac{v}{H^2}(t_\alpha, h'_\alpha)\right] \leq \mu_e(u(t_\alpha, h_\alpha), x_\alpha)d_y^v \tag{2.68}$$

On the other hand, since v is a supersolution of 2.44, analogously, we get

$$\frac{\sigma_d^2(x'_\alpha)}{2}X_{11}^\omega + \varepsilon^2 X_{22}^\omega + 2\alpha(x_\alpha - x'_\alpha)\left[\frac{\sigma_d^2}{2}(x'_\alpha)\frac{H_x}{H}(t_\alpha, h'_\alpha) + \frac{\mu_d(x'_\alpha}{2})\right] +$$
$$+ \alpha(y_\alpha - y'_\alpha)\left[\frac{2\varepsilon^2 H_y}{H}(t_\alpha, h'_\alpha) + \mu_e(u(t_\alpha, h'_\alpha), x'_\alpha)\right] + a - \frac{\delta}{(\varrho - t_\alpha)^2} +$$
$$+ \frac{v}{H^2}(t_\alpha, h'_\alpha)\Big[\frac{\sigma_d^2}{2}(x'_\alpha)H_{xx}(t_\alpha, h'_\alpha) + \varepsilon^2 H_{yy}(t_\alpha, h'_\alpha) + \mu_e(v(t_\alpha, h'_\alpha), x'_\alpha)H_y(t_\alpha, h'_\alpha) +$$
$$+ \mu_d(x'_\alpha)H_x(t_\alpha, h'_\alpha) + H_t(t_\alpha, h'_\alpha)\Big] +$$
$$+ \mu_e(u(t_\alpha, h_\alpha), x_\alpha)\left[\alpha(y_\alpha - y'_\alpha) + \frac{v}{H^2}(t_\alpha, h'_\alpha)\right] \geq \mu_e(u(t_\alpha, h_\alpha), x_\alpha)d_y^v \tag{2.69}$$

Finally, subtracting 2.68 from 2.70, for $\alpha > 0$, we obtain

$$I_\alpha + J_\alpha \geq 0 \tag{2.70}$$

where

$$\widehat{I}_\alpha = \alpha\Big\langle h_\alpha - h'_\alpha, \frac{\sigma_d^2(x'_\alpha)}{2}\frac{H_x}{H}(t_\alpha, h'_\alpha - \frac{\sigma_d^2(x_\alpha)}{2}\frac{H_x}{H}(t_\alpha, h_\alpha) + \frac{\mu_d(x'_\alpha)}{2} - \frac{\mu_d(x_\alpha)}{2},$$
$$\frac{2\varepsilon^2 H_y}{H}(t_\alpha, h'_\alpha) - \frac{2\varepsilon^2 H_y}{H}(t_\alpha, h_\alpha) + \mu_e(u(t_\alpha, h'_\alpha), x'_\alpha) - \mu_e(u(t_\alpha, h_\alpha), x_\alpha)\Big\rangle -$$
$$- \alpha(y_\alpha - y'_\alpha)\frac{2\delta}{(\varrho - t)^2} \tag{2.71}$$

and

$$J_\alpha = \frac{u}{H^2}(t_\alpha, h_\alpha)\Big[\frac{\sigma_d^2(x_\alpha)}{2}H_{xx}(t_\alpha, h_\alpha) + \varepsilon^2 H_{yy}(t_\alpha, h_\alpha) + \mu_e(u(t_\alpha, h_\alpha), x_\alpha)H_y(t_\alpha, h_\alpha) + \\ + \mu_d(x_\alpha)H_x(t_\alpha, h_\alpha) + H_t(t_\alpha, h_\alpha)\Big] - \frac{v}{H^2}(t_\alpha, h'_\alpha)\Big[\sigma_d^2(x'_\alpha)H_{xx}(t_\alpha, h'_\alpha) + \\ + \varepsilon^2 H_{yy}(t_\alpha, h'_\alpha) + \mu_e(u(t_\alpha, h'_\alpha), x'_\alpha)H_y(t_\alpha, h'_\alpha) + \mu_d(x'_\alpha)H_x(t_\alpha, h'_\alpha) + H_t(t_\alpha, h'_\alpha)\Big] \tag{2.72}$$

thus, putting $I_\alpha = \widehat{I}_\alpha + \alpha(y_\alpha - y'_\alpha)\frac{2\delta}{(\varrho - t_\alpha)^2}$, we get

$$I_\alpha + J_\alpha \geq \frac{2\delta}{(\varrho - t_\alpha)^2} > 0 \tag{2.73}$$

As α goes to infinity, we have $I_\alpha \to 0$ and

$$J_\alpha \to \\ \frac{u-v}{H}(t_0, h_0)\frac{\sigma_d^2 H_{xx} + \varepsilon^2 H_{yy} + (\mu_e(u, .) + \mu_e(v, .))H_y + \mu_d H_x + H_t}{H}(t_0, h_0) = \\ = \frac{u-v}{H}(t_0, h_0)\widehat{H}(\varepsilon, t_0, h_0) \tag{2.74}$$

Finally, since $\frac{u-v}{H}(t_0, h_0) > 0$ and $\sup_{S_\varrho} \widehat{H}(\varepsilon, t, x) < 0$ we have

$$0 < I_\alpha + J_\alpha \to \frac{u-v}{H}(t_0, h_0)\widehat{H}(\varepsilon, t_0, h_0) < 0 \tag{2.75}$$

and this is clearly a contradiction. Thus we have proved $u \leq v$ in S_ϱ. Repeating this procedure finitely many times, we conclude the proof. □

Remark 2.9. The uniqueness of the solution follows directly from the last proposition because if u and v are both viscosity solution of 2.44,by definition, they are both viscosity subsolution and supersolution. Hence, from the proposition, if we take v as subsolution and u as supersolution, we have $v \leq u$. Analogously we obtain $u \leq v$, thus we conclude $u = v$.

From general theory of second order parabolic PDEs we it is well known that, in our hypothesis (in particular Lipschitz continuity of σ_d, μ_d and μ_e, and Hölder continuity of u^ε), there exist a function $v \in C^{1+\frac{\alpha}{2}, 2+a}(S) \cap C(S \cap \widetilde{\partial} S)$ classical solution of the linear Cauchy problem

$$\begin{cases} \frac{\sigma_d(x)^2}{2}v_{xx} + \varepsilon^2 v_{yy} + \mu_d(x)u_x + \mu_e(u^\varepsilon, x)v_y + v_t = 0, & \text{in } S \\ v = u^\varepsilon, & \text{in } \widetilde{\partial} S \end{cases} \tag{2.76}$$

for $\varepsilon > 0$ and

$$S = \{(t, x, y)|x^2 + y^2 < R^2, t \in]0, T[\}$$
$$\widetilde{\partial} S = \partial S \cap \{t > 0\} \tag{2.77}$$

with $R > 0$ fixed. For the proof of this result see, for example, [2] or [11]. Moreover, by the comparison principle for viscosity solutions D.7 in appendix D, we have $u^\varepsilon = v$ in S.

Theorem 2.10. *If $\varepsilon > 0$, then u^ε is a solution of the problem 2.44 in classical sense. Moreover $u^\varepsilon \in C^\infty([0, T] \times \mathbb{R}^2)$.*

Proof. From the discussion above, since R is arbitrary, we have that u^ε is a solution of the problem 2.44 in classical sense. Moreover we have $u^\varepsilon \in C^{1+\frac{\alpha}{2},2+a}([0, T] \times \mathbb{R}^2)$. Now we use the following bootstrap argument: we apply the same procedure again with the new regularity of u^ε in the equation 2.76, thus we obtain the solution $v \in C^{3+\frac{\alpha}{2},3+a}(S)$, and again we get $u^\varepsilon \in C^{3+\frac{\alpha}{2},3+a}([0, T] \times \mathbb{R}^2)$. Applying this procedure infinitely many times we have $u^\varepsilon \in C^\infty([0, T] \times \mathbb{R}^2)$. □

Finally, we can prove the main result.

Theorem 2.11. *Let $0 < T < \overline{T}$, with $\overline{T} = (k_1(k_2 + 1))^{-1}$. Let respectively denote by k_1, k_2, k_3, k_4 the Lipschitz constant of μ_e, g, μ_d, σ_d. Then there exist a unique viscosity solution u of problem*

$$\begin{cases} \frac{\sigma_d(x)^2}{2}\frac{\partial^2 u}{\partial x^2} + \mu_d(x)\frac{\partial u}{\partial x} + \mu_e(u, x)\frac{\partial u}{\partial y} + \frac{\partial u}{\partial t} = 0 \text{ in } [0, T[\times \mathbb{R}^2 \\ u(T, x, y) = g(x, y) \quad (x, y) \in \mathbb{R}^2 \end{cases} \tag{2.78}$$

such that

$$|u^(t_2, x_2, y_2) - u^(t_2, x_1, y_1)| \leq C_0(|x_2 - x_1| + |y_2 - y_1|)$$
$$|u^(t_2, x_2, y_2) - u^(t_1, x_2, y_2)| \leq \widetilde{C_0}(1 + |(x_1, y_1)|)|t_2 - t_1|^{\frac{1}{2}} \tag{2.79}$$

for every $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$,$t_1, t_2 \in [0, T]$, where $C_0, \widetilde{C_0}$ are positive constant which depend on k_1, k_2, k_3 and k_4. For every $\varepsilon \in]0, 1[$, the regularized problem

$$\begin{cases} \frac{\sigma_d(x)^2}{2}\frac{\partial^2 u}{\partial x^2} + \varepsilon^2\frac{\partial^2 u}{\partial y^2} + \mu_d(x)\frac{\partial u}{\partial x} + \mu_e(u, x)\frac{\partial u}{\partial y} + \frac{\partial u}{\partial t} = 0 \text{ in } [0, T[\times \mathbb{R}^2 \\ u(T, x, y) = g(x, y) \quad (x, y) \in \mathbb{R}^2 \end{cases} \tag{2.80}$$

has an unique classical solution u^ε for which 2.79 holds with $C_0, \widetilde{C_0}$ independent of ε. Moreover u^ε converges to u as ε goes to zero, uniformly on compacts of $[0,T] \times \mathbb{R}^2$.

Proof. We have already proof existence, estimate 2.32 and uniqueness of the solution $u^{varepsilon} \in C^\infty([0,T] \times \mathbb{R}^2)$ of the regularized problem 2.80. To complete the proof we have to show that u is a vanishing viscosity solution in the sense that u is the limit of u^ε, uniform on compacts as $\varepsilon \to 0^+$. The technique is the same used in proposition 2.8, thus we only sketch the proof. Fix $\varrho > 0$ suitably small so that the function

$$H(t,h) = exp\left(\frac{|h|^2}{1-(2\varrho)^{-1}t} - \sigma t\right) \quad (t,h) \in \widetilde{S_\varrho} \tag{2.81}$$

with $S_\varrho =]0, \varrho[\times\mathbb{R}^2$ is such that

$$\hat{k} = \sup_{\varepsilon\in]0,1[} \sup_{S_\varrho} \frac{\frac{\sigma_d^2}{2} H_{xx} + (\mu_e(u^\varepsilon, x) + \mu_e(u,x))H_y + \mu_d(x)H_x + H_t}{H} < 0 \tag{2.82}$$

We have to show that $\forall R, \gamma > 0$ exist $\varepsilon_0 > 0$ such that $|u^\varepsilon(z) - u(z)| \leq \gamma$ $\forall z \in [0, \varrho[\times B(0,R)$ and $\varepsilon \in]0, \varepsilon_0[$, where $B(0,R)$ represent the Euclidean ball in $\mathbb{R}^2$. By contradiction, we assume that for some $R, \gamma > 0$ and every $\varepsilon > 0$ there exist $z^\varepsilon \in [0, \varrho[\times B(0,R)$ such that $(u^\varepsilon - u)(z^\varepsilon) > \gamma$. We consider the following functions defined on $[0, \varrho[\times\mathbb{R}^2$:

$$w^\varepsilon = \frac{u^\varepsilon}{H} - \frac{\delta}{\varrho - t}$$
$$\omega = \frac{u}{H} + \frac{\delta}{\varrho - t}$$

and we choose $\delta > 0$ suitably small and independent of ε, so that

$$w^\varepsilon(z^\varepsilon) - \omega(z^\varepsilon) > 0$$

Proceeding as in the proof of proposition 2.8 we can show the existence of a global maximum $(t_0^\varepsilon, h_0^\varepsilon)$ of $w^\varepsilon - \omega$, and analogously we obtain I_α^ε and J_α^ε and the following inequality:

$$I_\alpha^\varepsilon + J_\alpha^\varepsilon \geq \frac{2\delta}{\varrho - t_\alpha} > 0$$

with

$$I_\alpha^\varepsilon \to 0$$

and

$$J_\alpha^\varepsilon \to \left(\frac{u^\varepsilon - u}{H} \frac{\frac{\sigma_d^2}{2} H_{xx} + (\mu_e(u^\varepsilon, x) + \mu_e(u, x)) H_y + \mu_d(x) H_x + H_t}{H} + \varepsilon^2 \frac{u^\varepsilon H_{yy}}{H} \right) (t_0^\varepsilon, h_0^\varepsilon)$$

as $\alpha \to +\infty$. Setting

$$\overline{k} = \sup_{S_\varrho} \left| \frac{u^\varepsilon H_{yy}}{H} \right| < \infty$$

as $\alpha \to +\infty$ we get

$$0 \leq \hat{k} \frac{u^\varepsilon - u}{H} (t_0^\varepsilon, h_0^\varepsilon) + \varepsilon^2 \overline{k} < \frac{\hat{k}\gamma}{H(t_0^\varepsilon, h_0^\varepsilon)} + \varepsilon^2 \overline{k} \tag{2.83}$$

Since

$$\lim_{|h| \to \infty} (w^\varepsilon - \omega)(t, h) = -\frac{2\delta}{\varrho - t} < 0$$

uniformly in ε, we have

$$\sup_{\varepsilon \in]0,1[} |h_0^\varepsilon| < \infty$$

Therefore, by the last inequality, 2.83 contradicts the fact that $\varepsilon > 0$ is arbitrarily small and this concludes the proof. □

Remark 2.12. The bound on T comes from the consideration we have done in remark 2.3.

We have just proved, under suitable hypothesis, the existence and the uniqueness of solution of problem 2.78, which is associated to the stochastic system 2.5. Thus we can solve the differential problem to obtain a solution of the backward stochastic differential equation, and we can conclude that for any $d \in (0, \xi_{max})$, there exist an unique solution. This result is obtained in the case of Lipschitz continuity of the function g, but to give a price to allowance certificates the terminal condition function is the indicator function $\pi \mathbb{I}_{[E_{cap}, +\infty)}(E_t)$. In the next section we will investigate on the solution of problem with this singular terminal condition.

2.2 Singular terminal condition

We will show that there is no way to construct a solution to 2.5 with terminal condition $\pi \mathbb{I}_{[E_{cap}, +\infty)}(E_t)$ which preserve the flow property and the expected Markovian structure at terminal time T, and we will prove it trough the degeneracy of the inviscid Burger's equation. Before to do this we want to study the case of the relaxed terminal condition:

$$\mathbb{P}\{\phi_-(E_T) \le A_T \le \phi_+(E_T)\} = 1 \tag{2.84}$$

with

$$\begin{aligned} \phi(x) &= \mathbb{I}_{[E_{cap},+\infty)}(x) \\ \phi_-(x) &= \sup_{x'<x} \phi(x') \\ \phi_+(x) &= \inf_{x'>x} \phi(x') \end{aligned} \tag{2.85}$$

Note that, for shake of simplicity, we take the penalty π equal to one. For the general case, it is enough to take $\pi\phi$ instead of ϕ and we obtain the same result. It can be shown the following theorem (see [15])

Theorem 2.13. *Assume there exist two constants $l_1, l_2 > 0$, and $1/L \le l_1 \le l_2 \le L$ with $L \ge 1$ the Lipschitz constant of μ_e such that*

$$l_1|A - A'| \le |\mu_e(D, A) - \mu_e(D, A')| \le l_2|A - A'| \quad A, A' \in \mathbb{R}$$

Given any initial condition $(d, e) \in \mathbb{R}^2$, there exist a unique progressively measurable 4-tuple $(D_t, E_t, A_t, Z_t)_{0\le t\le T}$ satisfying

$$\begin{cases} dD_t = \mu_d(D_t)dt + \sigma_d(D_t)dW_t \\ dE_t = \mu_e(A_t, D_t)dt \\ A_t = rA_t dt + e^{rt} Z_t dW_t \end{cases} \tag{2.86}$$

and the relaxed terminal condition 2.84. Moreover there exist a constant C depending on L and T only, such that almost surely $|Z_t| \le C$ for $t \in [0, T]$.

Proof. The complete proof can be seen on the article [15], we only give an idea of this proof. We consider a perturbed stochastic system (called mollified system) similarly to what we have done in the previous section and to associate the solution of the stochastic problem to a semilinear PDE. Then, for a Lipschitz smooth terminal condition ϕ we have a solution u^ϕ which depends on the terminal condition.

Now assume ϕ to be a non-decreasing function, ϕ_+ and ϕ_- as in 2.84. Notice that ϕ_+ is a cumulative distribution function as a non-decreasing right continuous function matching 0 at $-\infty$ and 1 at $+\infty$. Notice also that ϕ_- is the left-continuous version of ϕ_+. Then we can construct two mollifying sequences for ϕ. let j being a C^∞ function with compact support which represent the density of a positive random variable, and let ξ and ϑ be positive independent random variables, ξ with ϕ as cumulative distribution function and ϑ with j as density. For each integer $n \geq 1$, denote by ϕ_+^n and ϕ_-^n the cumulative distribution functions of the random variables $\xi - n^{-1}\vartheta$ and $\xi + n^{-1}\vartheta$ respectively. Then, the functions ϕ_+^n and ϕ_-^n are non decreasing with values in $[0, 1]$. They are $C^{+\infty}$, with bounded derivatives of any order. Moreover $\phi_+^n \geq \phi$ and $\phi_-^n \leq \phi$ and the sequences $(\phi_+^n)_{n\geq1}$ and $(\phi_-^n)_{n\geq1}$ converge pointwise towards ϕ_+ and ϕ_- respectively as n tends to $+\infty$. Finally

$$\int_{\mathbb{R}} |\phi_+^n(e) - \phi_+(e)|de \leq \int_{\mathbb{R}\times\mathbb{R}^+} \mathbb{P}(e \leq \xi \leq e + t/n)j(t)dtde = \frac{1}{n}\int_{\mathbb{R}^+} tj(t)dt \to 0$$

as n tends to $+\infty$, so that the convergence of $(\phi_+^n)_{n\geq1}$ towards ϕ_+ holds in $L^1(\mathbb{R})$ as well. Analogously the convergence of $(\phi_-^n)_{n\geq1}$ towards ϕ_- holds in $L^1(\mathbb{R})$. Then, for each $n \geq 1$, we obtain two solutions $v^{\phi_+^n}$ and $v^{\phi_-^n}$ to the PDE associated to the problem 2.86 with terminal condition ϕ_+^n and ϕ_-^n respectively. The sequences $(v^{\phi_+^n})_{n\geq1}$ and $(v^{\phi_-^n})_{n\geq1}$ converge uniformly on compact subset of $[0, T) \times \mathbb{R} \times \mathbb{R}$, therefore

$$\lim_{n\to+\infty} v^{\phi_+^n}(t, d, e) = \lim_{n\to+\infty} v^{\phi_-^n}(t, d, e)$$

for any $(t, d, e) \in [0, T) \times \mathbb{R} \times \mathbb{R}$. By construction the limit matches the continuous function $v(t, d, e)$ such that $A_t = v(t, D_t^d, E_t^{d,e})$ for any $t < T$ and

$$\phi_-(E_T^{d,e}) \leq \lim_{t\nearrow T} v(t, D_t^d, E_t^{d,e}) \leq \phi_+(E_T^{d,e}) \quad \mathbb{P}\text{-almost surely}$$

where the limit exist as the almost-surely limit of a non-negative martingale. Moreover $v(t, d, e)$ is Lipschitz continuous respect to d and e, and it is a $[0, 1]$-valued martingale with respect to the complete filtration generated by the Brownian motion W_t. The integral martingale representation of $(v(t, D_t^d, E_t^{d,e}))_{0\leq t<T}$ is bounded by a constant dependent on T and C (the Lipschitz constant respect to the variable e) only. The existence of such function should be proved, and it is done in [15]. □

Remark 2.14. The statement is true also for the initial time $t_0 \in [0,T)$ instead of 0.

In the following we always refer to v as the function defined in the proof of theorem 2.13, and we empathise that the function $\mu_e(a,d)$ is not increasing as a function of the variable a. Now we discuss the existence of a solution in the case $\phi(x) = \mathbb{I}_{[E_{cap},+\infty)}(x)$ and in order to do this we give some additional hypothesis:

- For any $d \in \mathbb{R}$, the function $y \hookrightarrow \mu_e(a,d)$ is differentiable with respect to a and there exist $\alpha \in (0,1]$ such that, for any $(d,d',a,a') \in \mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}$, we have
$$|\partial_y\mu_e(a,d) - \partial_y\mu_e(a',d')| \le L(|d'-d|^\alpha + |y'-y|^\alpha)$$
- the function μ_d and σ_d are bounded by L.

For shake of convenience, we switch from the degenerate component E of the forward process to a process $\overline{E}$ which has the same terminal value, hence leaving the terminal condition of the backward process unchanged, and which will be easier to manipulate. We introduce the modified process

$$\overline{E}_t = E_t - E\left[\int_t^T \mu_e(0,D_s)ds\middle|\mathcal{F}_t\right] \tag{2.87}$$

Hence the process $\overline{E}_t$ gives an approximation of E_T given $\mathcal{F}_t$, and in particular $\overline{E}_T = E_T$. Moreover, since μ_d and σ_d are Lipschitz continuous we have that $w : [0,T]\times\mathbb{R} \to \mathbb{R}$ is a deterministic function, and if μ_d, σ_d and μ_e have bounded derivatives of any order, w is a classical solution of the PDE:

$$\partial_t w(t,d) + \frac{\sigma_d^2(d)}{2}\partial_{dd}w(t,d) + \mu_d(d)\partial_d w(t,d) - \mu_e(d,0) = 0 \tag{2.88}$$

with terminal condition w(T,d)=0. Consequently $\overline{E}_t$ is an Itô process and

$$\begin{aligned} d\overline{E}_t &= dE_t + d[w(t,D_t)] = \\ &= [\mu_e(A_t,D_t) - \mu_e(0,D_t)]dt + \sigma_d(D_t)\partial_d w(t,D_t)dW_t \end{aligned} \tag{2.89}$$

If the coefficients μ_d, σ_d and μ_e haven't bounded derivative of any order, the equality 2.89 still holds thanks to Itô's formula and to Martingale Representation Theorem. But $\sigma_d(D_t)\partial_d w(t,D_t)$ may not exists, and in this case the integrand of martingale part is given by the Martingale Representation Theorem. However we still write $\sigma_d(D_t)\partial_d w(t, D_t)$ for the integrand appearing in the stochastic integral respect to W. In any case this integrand is bounded.

Lemma 2.15. *There exist a constant C, depending on L and T only, such that*

$$\forall(t,d,d') \in [0,T) \times \mathbb{R} \times \mathbb{R}, |w(t,d') - w(t,d)| \leq C(T-t)|d'-d|$$

In particular, when it exist, the function $\partial_d w(t,\cdot)$ is uniformly bounded from above by C(T-t). And, in any case, the representation term $(\mu_d(D_t)\partial_d w(t,D_t))_{t_o \leq t \leq T}$ is bounded by $CL(T-t)$ provided σ_d is bounded by L.

Proposition 2.16. *There exist a constant C and an exponent $\beta \in (0,1)$, depending on α, L and T only such that*

$$\forall(t_0,d,e) \in [0,T) \times \mathbb{R} \times \mathbb{R}, \left| v(t_0,d,e) - \psi\left(\frac{\overline{e} - E_{cap}}{l(t_o,d,e)|T-t_0|}\right)\right| \leq C(T-t_0)^{\beta}$$

where

$$l(t_0,d,e) = \int_0^1 \frac{\partial \mu_e}{\partial a}(\lambda v(t_0,d,e),d)d\lambda$$

and $\overline{e} = e + w(t,d)$. Moreover the function

$$\psi(e) = e\mathbb{I}_{[0,1]}(e) + \mathbb{I}_{(1,+\infty)}(e)$$

is the solution of the inviscid Burger's equation

$$\partial_t u(t,e) - u(t,e)\partial_e u(t,e) = 0, \quad (t,e) \in [0,T) \times \mathbb{R}$$

with terminal condition

$$u(T,\cdot) = \mathbb{I}_{[0,+\infty)}$$

Note that by definition we have $v(t_0,d,e)l(t_0,d,e) = \mu_e(d, v(t_0,d,e)) - \mu_e(d,0)$.

Remark 2.17. The function $\psi\left(\frac{\overline{e}-E_{cap}}{l|T-t_0|}\right)$ with $l \in [l_1, l_2]$ constant satisfies the inviscid Burger's equation

$$\partial_t u(t,e) - lu(t,e)\partial_e u(t,e) = 0, \quad (t,e) \in [0,T) \times \mathbb{R}$$

with terminal condition

$$u(T,\cdot) = \mathbb{I}_{[E_{cap},+\infty)}$$

The choice of l in proposition 2.16 is done to consider a function μ_e of general form, and not only in the affine form $\mu_e(a,d) = \mu_e(0,d) - la$. Thus for any $d \in \mathbb{R}$ the function $\psi\left(\frac{\bar{e}-E_{cap}}{l(t_o,d,e)|T-t_0|}\right)$ solves the equation:

$$\partial_t u(t,e) - l(t_0,d,e)u(t,e)\partial_e u(t,e) = 0, \quad (t,e) \in [t_0,T) \times \mathbb{R}$$

with terminal condition

$$u(T,\cdot) = \mathbb{I}_{[E_{cap},+\infty)}$$

The next proposition shows the existence of a Dirac Mass at terminal time T.

Proposition 2.18. *There exist a constant $c \in (0,1)$ depending on α and L only, such that, if $T - t_0 \le c$, $p \in \mathbb{R}$ and $\frac{\bar{e}-E_{cap}}{|T-t_0|} \in \left[\frac{l_1}{4}, \frac{3l_1}{4}\right]$, then:*

$$\mathbb{P}\left\{E_T^{t_0,p,e} = E_{cap}\right\} \ge c$$

Remark 2.19. In particular, in the case of the allowance certificates we have $t_0 = 0$.

Proof. Given an initial condition $(t_0,d,e) \in [0,T) \times \mathbb{R} \times \mathbb{R}$ for the process (D_t,E_t), we consider the stochastic differential equations:

$$\begin{aligned} d\overline{E}_t^{\pm} &= \left(l(t,D_t,E_t)\psi\left[l^{-1}(t,D_t,E_t)\frac{E_t^{\pm}-E_{cap}}{T-t}\right] \pm C'(T-t)^{\beta}\right) + \\ &+ \sigma_d(D_t)\partial_d w(t,D_t)dW_t \end{aligned} \tag{2.90}$$

with $E_{t_0}^{\pm} = \bar{e}$ as initial conditions, the constant C' being chosen later on. Notice that the process appearing in l and l^{-1} above is E and not $\overline{E}^{\pm}$.From 2.89 and the definition of l it follows that

$$d\overline{E}_t = l(t,D_t,E_t)v(t,D_t,E_t)dt + \sigma_d(D_t)\partial_d w(t,D_t)dW_t, \quad t \in [t_0,T) \tag{2.91}$$

with

$$\left| l(t,D_t,E_t)v(t,D_t,E_t) - l(t,D_t,E_t)\psi\left[l^{-1}(t,D_t,E_t)\frac{E_t^{\pm}-E_{cap}}{T-t}\right]\right| \le LC(T-t)^{\beta}, \quad t \in [t_0,T)$$

where C is given by proposition 2.16. We now choose $C' = LC$. By the comparison theorem for one-dimensional SDE, we deduce

$$\overline{E}_t^{-} \le \overline{E}_t \le \overline{E}_t^{+}, \quad t \in [t_0,T)$$

Next we introduce the bridge equations

$$d\overline{Z}_t^{\pm} = \left(\frac{\overline{Z}_t^{\pm} - E_{cap}}{T-t} \pm C'(T-t)^{\beta} \right) dt + \sigma_d(D_t)\partial_d w(t, D_t) dW_t, \quad \overline{Z}_{t_0}^{\pm} = \overline{e} \quad (2.92)$$

The solution is given by

$$\overline{Z}_t^{\pm} = E_{cap} + (T-t) \left[\frac{\overline{e} - E_{cap}}{T-t} \pm C' \int_{t_0}^{t} (T-s)^{\beta-1} ds + \int_{t_0}^{t} (T-s)^{-1} \sigma_d(D_s)\partial_d w(s, D_s) dW_s \right] \quad (2.93)$$

so that $\overline{Z}_t^{\pm} \to E_{cap}$ as $t \to T$. The stochastic integral is well defined up to time T since $\sigma_d(D_s)\partial_d w(s, D_s)$ is bounded. Now, we choose $\overline{e}$ such that $\overline{e} - E_{cap}/(T-t_0) \in [l_1/4, 3l_1/4]$ and t_0 such that $C' \int_{t_0}^{t} (T-s)^{\beta-1} ds \in [0, l_1/16]$, and we introduce the stopping time

$$\tau = inf \left\{ t \geq t_0 : \left| \int_{t_0}^{t} (T-s)^{-1} \sigma_d(D_s)\partial_d w(s, D_s) dW_s \right| \geq \frac{l_1}{16} \right\} \wedge T$$

for any $t \in [t_o, \tau)$, so that

$$\frac{\overline{Z}_t^{\pm} - E_{cap}}{T-t} = l(t, D_t, E_t)\psi \left[l^{-1}(t, D_t, E_t) \frac{\overline{Z}_t^{\pm} - E_{cap}}{T-t} \right], \quad t_0 \leq t \leq \tau$$

in other words,$(\overline{Z}_t^{\pm})_{t_0 \leq t \leq \tau}$ and $(\overline{E}_t^{\pm})_{t_0 \leq t \leq \tau}$ coincide. We deduce that, on the event $F = \{\tau = T\}$,

$$\overline{Z}_t^{\pm} = \overline{E}_t^{\pm}, \quad t \in [t_0, T]$$

Finally, by Markov inequality and lemma 2.15, the probability of the event F is strictly greater than zero for $T - t_0$ small enough. And this complete the proof. □

Remark 2.20. If we consider the non-viscous case, that is when E_t depends only on A_t, the dynamics of E_t coincides with the dynamics of the characteristic of the associated inviscid equation of conservation law. The simplest example is the SDE:

$$\begin{cases} dE_t = -u(t, E_t) dt \\ E_{t_0} = e; \end{cases} \quad (2.94)$$

where u satisfy the Burger's equation in the form

$$\begin{cases} \frac{\partial u}{\partial t}(t, x) - u(t, x) \frac{\partial}{\partial x}(t, x) = 0 \\ u(T, x) = \mathbb{I}_{[E_{cap}, +\infty)}(x) \end{cases} \quad (2.95)$$

Note that this PDE is the first order non-viscous version of the PDE

$$\frac{\sigma_d(x)^2}{2}\frac{\partial^2 u}{\partial x^2} + \mu_d(x)\frac{\partial u}{\partial x} + \mu_e(u,x)\frac{\partial u}{\partial y} + \frac{\partial u}{\partial t} = 0 \tag{2.96}$$

where we take $\mu_e(u,x) = -u$ since $\mu_e(u,x)$ is decreasing in the first variable. We want to solve the equation 2.95 with the characteristic curves method. We can rewrite the equation as follows

$$\begin{cases} -\frac{\partial u}{\partial t}(t,x) + u(t,x)\frac{\partial u}{\partial x}(t,x) = 0 \\ u(T,x) = \mathbb{I}_{[E_{cap},+\infty)}(x) = \phi(x) \end{cases} \tag{2.97}$$

Now we consider curves $\gamma(s,r) = (t(s,r), x(s,r))$ along which the solution of the PDE is constant, thus we have

$$\frac{du}{ds}(t(s,r), x(s,r)) = 0 \tag{2.98}$$

and we obtain

$$\frac{\partial u}{\partial t}\frac{dt}{ds} + \frac{\partial u}{\partial x}\frac{dx}{ds} = 0 \tag{2.99}$$

Comparing the last equation with we get the following system of ODE

$$\begin{cases} \frac{du}{ds} = 0 \\ \frac{dx}{ds} = u \\ \frac{dt}{ds} = -1 \end{cases} \tag{2.100}$$

with initial conditions

$$\begin{cases} u(T,r) = \phi(r) \\ x(0,r) = r \\ t(0,r) = T \end{cases} \tag{2.101}$$

We emphasize that u in the last equation is constant because we are studying it along a curve on which it is constant. We see that the solution is given by

$$\begin{cases} u(s,r) = \phi(r) \\ x(s,r) = us + r \\ t(s,r) = T - s \end{cases} \tag{2.102}$$

and we obtain $x(t,r) = \phi(r)(T-t) + r$ and the characteristic curves are of the form $\gamma(s,r) = (T-s, \phi(r)s + r)$. If $r < E_{cap}$, curves are $\gamma(s,r) = (T-s, r)$, that

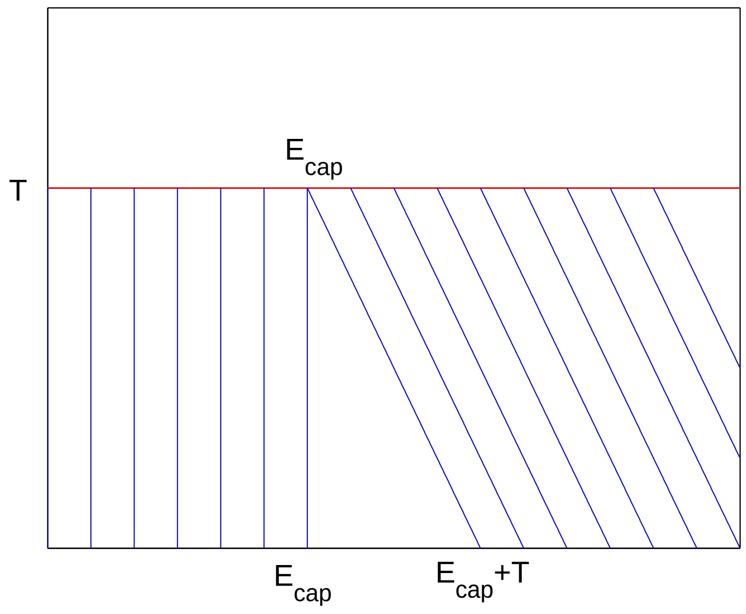


Figure 2.1: Characteristics

is they are vertical lines along which the solution is equal to 0. On the other hand if $r \geq E_{cap}$, characteristics are $\gamma(s, r) = (T - s, s + r)$, that is they are lines along which the solution is equal to 1. We can resume this result in picture 2.1.

And we have the solution

$$u(t, x) = \begin{cases} 0 & \text{if } x < E_{cap} \\ 1 & \text{if } x \geq E_{cap} + t. \end{cases} \tag{2.103}$$

with $(t, x) \in [0, T] \times \mathbb{R}$. However we have a problem, in fact there is a region on which we don't have enough information and we have to define a solution also in this cone. We define rarefaction solution in the interval $E_{cap} \leq x < E_{cap} + T$. Here the characteristic curves have equation $x(t) = c(T - t) + E_{cap}$ with $0 < c < 1$ because they are forced to pass trough E_{cap} at time T. Along these curves the solution is constant and equal to $u(t, x) = c$, thus we obtain $u(t, x) = c = f\left(\frac{x - E_{cap}}{T - t}\right)$ for a suitable function f, which will be defined later, and for $t < T$. We derive u respect to t and x and we obtain

$$\frac{\partial u}{\partial t} = f'(\frac{x - E_{cap}}{T - t})\left(-\frac{x - E_{cap}}{(T - t)^2}\right) \tag{2.104}$$

$$\frac{\partial u}{\partial x} = f'(\frac{x - E_{cap}}{T - t})\left(\frac{1}{T - t}\right) \tag{2.105}$$

Substituting the last equalities in 2.95 we have

$$f'(\frac{x-E_{cap}}{T-t})\left[f\left(\frac{x-E_{cap}}{T-t}\right)-\frac{x-E_{cap}}{T-t}\right]=0 \tag{2.106}$$

The case $f'\left(\frac{x-E_{cap}}{T-t}\right)$ don't satisfy Rankine-Hugoniot jump conditions because it implies $f\left(\frac{x-E_{cap}}{T-t}\right) = constant$. Thus we have $f\left(\frac{x-E_{cap}}{T-t}\right) = \frac{x-E_{cap}}{T-t}$. We can conclude that the solution to the PDE 2.95 is

$$u(t,x) = \begin{cases} 0 & \text{if } x < E_{cap} \\ \frac{x-E_{cap}}{T-t} & \text{if } E_{cap} < x \leq E_{cap} + t \text{ and } t < T \\ 1 & \text{if } x \geq E_{cap} + t. \end{cases} \tag{2.107}$$

and the characteristics are represented in picture 2.2.

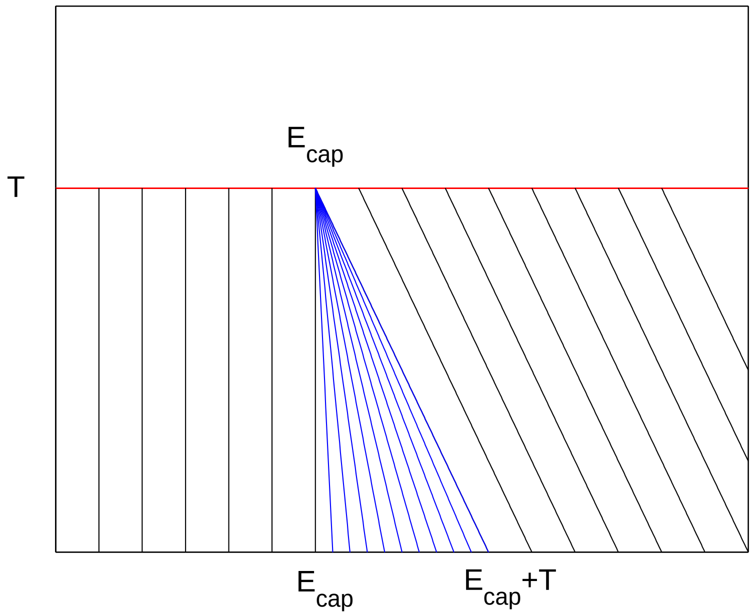


Figure 2.2: Characteristics with rarefaction

Consequently the process E_t in 2.94 with initial condition $E_{t_0} = e$ and $0 \leq t_0 < T$ and $e \in \mathbb{R}^{\geq 0}$ is

$$E_t = \begin{cases} e & \text{if } e < E_{cap} \\ e - \frac{e-E_{cap}}{T-t_0}(t-t_0) & \text{if } E_{cap} < x \leq E_{cap} + (T-t_0) \\ e-(t-t_0) & \text{if } e \geq E_{cap} + (T-t_0). \end{cases} \tag{2.108}$$

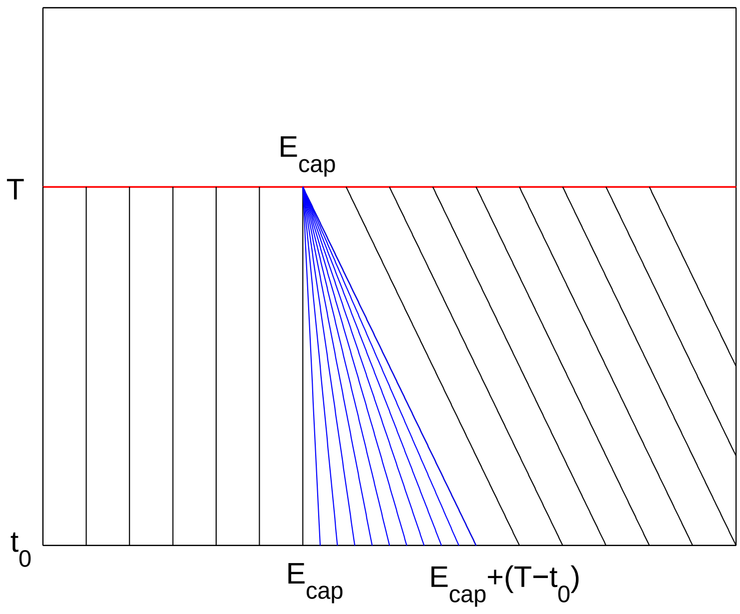


Figure 2.3: Trajectories of the emission process

and this correspond to the picture 2.3.

As we can see in this example there is a cone of initial conditions (t_0, d, e) for which the distribution of the random variable $E_T^{t_0,d,e}$ has a Dirac mass at a singular point E_{cap}, but we empathise that's not true for all initial point. Indeed, we have seen that this happens only if $E_{cap} \leq e < E_{cap} + (T - t_0)$ in this example. Thus there is a non-zero event scenarii for which the terminal conditions $\phi_-(E_T^{t_0,d,e})$ and $\phi_+(E_T^{t_0,d,e})$ differ, and this makes the relaxation of terminal condition $\mathbb{P}\{\phi_-(E_T) \leq A_T \leq \phi_+(E_T)\} = 1$ meaningful.

Theorem 2.21. *Let the foregoing hypothesis hold and assume that*

- $\sigma_d^2(d) \geq L^{-1}$, $d \in \mathbb{R}$
- *For any* $d \in \mathbb{R}$, $|\partial_d f(d, 0)| \geq L^{-1}$, *and the function* $p \to \partial_d f(d, 0)$ *is uniformly continuous.*

Then, for any starting point $(t_0, d, e) \in [0, T) \times \mathbb{R}^2$ *we have* $\mathbb{P}\{E_T^{t_0,d,e} = E_{cap}\} > 0$ *and the topological support of the conditional law of* $A_T^{t_0,d,e}$ *given* $E_T^{t_0,d,e} = E_{cap}$ *is* $[0, 1]$.

The last theorem shows that Dirac mass exist for all initial conditions $(t_0, d, e) \in [0, T) \times \mathbb{R}^2$. Moreover the fact that the topological support of the conditional law of $A_T^{t_0,d,e}$ given $E_T^{t_0,d,e} = E_{cap}$ is $[0, 1]$ means that all the values between $0 = \phi_-(E_{cap})$ and $1 = \phi_+(E_{cap})$ may be observed in the relaxed terminal condition $\mathbb{P}\{\phi_-(E_T) \leq A_T \leq \phi_+(E_T)\} = 1$. This implies that the σ-algebra $\sigma(A_T^{t_0,d,e})$ is not included into the σ-algebra $\sigma(E_T^{t_0,d,e})$ and the Markovian structure brakes down a terminal time. This fact has bad consequences for the emission market model, because a price for the allowance certificates exist and it is unique in such a model, but its terminal value cannot be prescribed as the model would require.

Chapter 3

Numerical Solution

This chapter is dedicate to the numerical solution of the problem and to the simulation of the processes involved. First of all we have to give a characterizations of the functions involved in the model: the Marginal Emission Stack $e(\xi)$, the Business as Usual Bid Stack b^{BAU}, and functions μ_d and σ_d of the Demand process. Then we specify the model's parameters and we present the necessary boundary conditions. Finally we graphically present the obtained results.

The Marginal Emission Stack $e(\xi)$ is define as follows

$$e(\xi) = \underline{e} + \left(\frac{\overline{e} - \underline{e}}{\xi^{\theta_2} max} \right) \xi^{\theta_2} \tag{3.1}$$

with $0 \leq \xi \leq \xi_{max}$, $\underline{e}, \overline{e} \geq 0$ and $0 \leq \theta_2 < 1$. With this choice e is strictly convex and strictly decreasing on its domain of definition. The parameters $\overline{e}, \underline{e}$ represent respectively the maximum and the minimum marginal emission rate in the market. In the assumption to have only two generators (gas and coal for example), $\underline{e}$ correspond to the marginal emission rate of the more environmental friendly and $\overline{e}$ to the marginal emission rate of the dirtier one. The parameter θ_2 controls the fuel mix in the market.The smaller the value of θ_2, the smaller portion of the market capacity that is served by the pollution intensive technology.

The Business as Usual Bid Stack b^{BAU} is taken in the form

$$b^{BAU}(\xi) = \underline{b} + \left(\frac{\overline{b} - \underline{b}}{\xi^{\theta_1} max} \right) \xi^{\theta_1} \tag{3.2}$$

with $0 \leq \xi \leq \xi_{max}$. Moreover $\underline{b}, \overline{b} \geq 0$ and $2 < \theta_1 < infty$. Under this assumption, b^{BAU} is strictly increasing and strictly convex and the parameters $\overline{b}, \underline{b}$ represent respectively the maximum and the minimum prices of electricity the model can produce. The range of the allowed bids in many auction based electricity market is well known, thus we can take them form historical data and they are relatively easy to infer in practice. The parameter θ_1 controls the stepness of the stack and in particular how quickly marginal costs of generators increase. We note that these functions respect all assumptions we have made in the model definition, and this choice of b^{BAU} and e makes the function

$$g(A, \xi) = b^{BAU}(\xi) + Ae(\xi), \quad \text{for} \quad 0 \leq A < \infty, \quad 0 \leq \xi < \xi_{max} \tag{3.3}$$

to be strictly convex. Moreover the set $S_p(\cdot, \cdot)$ is always of the form $[\xi_1, \xi_2]$ for $0 \leq \xi_1 \leq \xi_2 \leq \xi_{max}$.

The functions which define the process D_t are taken in the form

$$\begin{aligned} \mu_d(D_t) &= -\eta(D_t - \overline{D}) \\ \sigma_d(D_t) &= \sqrt{2\eta\overline{\sigma_d}D_t(\xi_{max} - D_t)} \end{aligned} \tag{3.4}$$

where $\overline{D}, \eta, \overline{\sigma_d} > 0$. With this definition the process

$$\begin{cases} dD_t = -\eta(D_t - \overline{D})dt + \sqrt{2\eta\overline{\sigma_d}D_t(\xi_{max} - D_t)}dW_t \\ D_0 = d \in (0, \xi_{max}) \end{cases} \tag{3.5}$$

ia Jacobi diffusion process and it has a linear, mean-reverting drift component and degenerates on the boundary. Moreover, subject to $\overline{D} \in (0, \xi_{max})$ and $min(\overline{D}, \xi_{max} - \overline{D})$, the process remains within the interval $(0, \xi_{max})$, and its stationary distributions is a beta distribution and its mean is given by $\overline{D}$ how we can see in [7].

For the choice of the model's parameters we don't take it from a particular example of electricity market, but they can be considered representative of a medium sized market whose fuel mix predominantly consist of coal and gas generators. In particular we take the following parameters for the bid and emission stack:

- $\overline{b} = 200$

- $\underline{b} = 0$

- $\theta_1 = 10$

- $\overline{e} = 1.2$

- $\underline{e} = 0.4$

- $\theta_2 = 0.4$

- $\xi_{max} = 30000$

Moreover in the function μ_e the constant k we use to set the timescale is taken equal to the number of production hours in a year because we want to solve the problem with the time interval set to 1 year. That is

- $k = 25 * 365 = 8760$

With these parameters and the processes $A_t = 0$ and $D_t = \xi_{max}$ for $0 \leq t \leq T$ we find $E_T = 1.6519 \times 10^8$. This is the maximum value that the process E_t can assume, thus we take

- $E_{max} = 1.6519 \times 10^8$

In the following list we give the parameters of the demand process

- $\eta = 10$

- $\overline{D} = 21000$

- $\overline{\sigma_d} = 0.05$

- $r = 0.05$

Now, calculating the cumulative emissions for $A_t = 0$ and demand at its mean level $D_t = \overline{D}$, for $0 \leq t \leq T$, we find that $E_T = 1.2961 \times 10^8$. This leads to choose the cup slightly below this level, in order to incentivise a reduction in emissions. The parameters characterising the emissions trading scheme are

- $E_{cap} = 1.17 \times 10^8$

- $\pi = 100$
- $T = 1$

and we remember that time is measured in years.

To solve numerically the Cauchy problem

$$\begin{cases} \frac{\sigma_d(x)^2}{2}\frac{\partial^2 u}{\partial x^2} + \varepsilon^2\frac{\partial^2 u}{\partial y^2} + \mu_d(x)\frac{\partial u}{\partial x} + \mu_e(u,x)\frac{\partial u}{\partial y} + \frac{\partial u}{\partial t} = 0 \text{ in } [0,T[\times\mathbb{R}^2 \\ u(T,x,y) = g(x,y) \quad (x,y)\in\mathbb{R}^2 \end{cases} \tag{3.6}$$

which represent the allowance certificate price we have to give some necessary boundary conditions. First of all, we need to understand at which boundary points we need to specify boundary conditions in addition to the terminal condition and what conditions make sense given the stochastic problem

$$\begin{cases} dD_t = \mu_d(D_t)dt + \sigma_d(D_t)dW_t, & D_0 = d \in (0, \xi_{max}); \\ dE_t = \mu_e(A_t, D_t)dt, & E_0 = 0; \\ A_t = \pi e^{-r(T-t)} E^Q \left[\mathbb{I}_{[E_{cap},\infty)}(E_T)\big|\mathcal{F}_t\right], & A_T = \pi \mathbb{I}_{[E_{cap},\infty)}(E_T). \end{cases} \tag{3.7}$$

We can do this thanks to Fichera's function f (see [12]). Defining $n = (n_d, n_e)$ to be the inward normal vector to the boundary, Fichera's function for the operator

$$\frac{\partial}{\partial t} + \frac{\sigma_d^2(D)}{2}\frac{\partial^2}{\partial D^2} + \mu_d(D)\frac{\partial}{\partial D} + \mu_e(\cdot, D)\frac{\partial}{\partial E} - r \tag{3.8}$$

is

$$f(t,D,E) = \left(\mu_d(D) - \frac{\sigma_d^2(D)}{2}\frac{\partial}{\partial D}\right) n_d + \mu_e(u(t,D,E),D)n_e \quad \text{on } \partial U_T \text{ or } \partial U_{T_i} \tag{3.9}$$

In the case of the coefficients μ_d and σ_d are of the form prescribed before we have

$$f(t,D,E) = \eta\left((\overline{D} - \overline{\sigma_d}\xi_{max}))(2\overline{\sigma_d} - 1)\right) n_d + \mu_e(u(t,D,E),D)n_e \\ \text{on } \partial U_T \text{ or } \partial U_{T_i} \tag{3.10}$$

At points were $f \geq 0$ information is outward flowing and no boundary conditions have to be specified. Conversely when $f < 0$ the information is inward flowing and boundary conditions are necessary. In the parts of the boundary when corresponding to $D = 0$ and to $D = \xi_{max}$, we have that $f \geq 0$ if and only if $min(\overline{D}, \xi_{max} - \overline{D}) \geq \xi_{max}\overline{\sigma_d}$, which is the same condition prescribed to guarantee that the process D_t stays within the interval $(0, \xi_{max})$. Thus we do not have to give boundary conditions in these parts of the boundary. At points of the boundary corresponding to $E = 0$, we find that $f \geq 0$ always.On the part of the boundary on which $E = E_{max}$, $f < 0$ except at the point $(D, E) = (0, E_{max})$, where $f = 0$, an ambiguity which would could be resolved smoothing the domain. Therefore, no boundary conditions are necessary except when $E = E_{max}$. In the case of one compliance period, the boundary condition at $E = E_max$ takes the form

$$u(t, D, E) = \pi e^{-r(T-t)}, \quad [0, T) \times (0, \xi_{max}) \times \{E = E_{max}\} \tag{3.11}$$

This condition follows from the fact that, as soon as the cumulative emission surpass the cap, every additional tonne of CO_2 is penalised at a rate π at time T.

Now we will show hot we have to discretized our domain $\overline{U}_T$ and the derivatives in the PDE. We choose mesh widths ΔD,ΔE and a time step Δt. The discrete mesh points (D_i, E_j, t_k) are then defined by

$$\begin{aligned} D_i &= i\Delta D \\ E_j &= j\Delta E \\ t_k &= k\Delta t \end{aligned} \tag{3.12}$$

The parameters we have we have choose to define the the mesh satisfy the Courant-Friedrichs-Lewy (see [14])condition for the convergence of explicit schemes. In particular we take

$$\begin{aligned} D_{max}/\Delta D &= 12 \\ E_{max}/\Delta E &= 200 \\ 1/\Delta t &= 440 \end{aligned} \tag{3.13}$$

The finite difference scheme we employ produces approximations $u^k_{i,j}$ which are assumed to converge to the true solution u as the mesh tends to zero.

We choose a backward scheme in order to work with an explicit scheme because the partial differential equation is posed backward in time. In the E-direction we are approximating a conservation law PDE with discontinuous terminal condition. The first derivative in the E-direction,relating to the non-linear part of the PDE, is discredited against the drift direction using a one-sided upwind difference. The characteristic information is propagating in the direction of decreasing E, this one sided difference is also used to calculate the value of the approximation on the part of the boundary corresponding to $E = 0$. In D-direction, the equation is parabolic everywhere except on the boundary, where it degenerates. Hence we use central differences to discretize the first and second order derivative. At the boundaries corresponding to $D = 0$ and $D = \xi_{max}$, where the second derivative vanishes and and no boundary conditions need to be specified, we use a one-sided difference in our numerical scheme. In figures 3.1, 3.2, 3.3, 3.4, 3.5 we display the obtained results for different times.

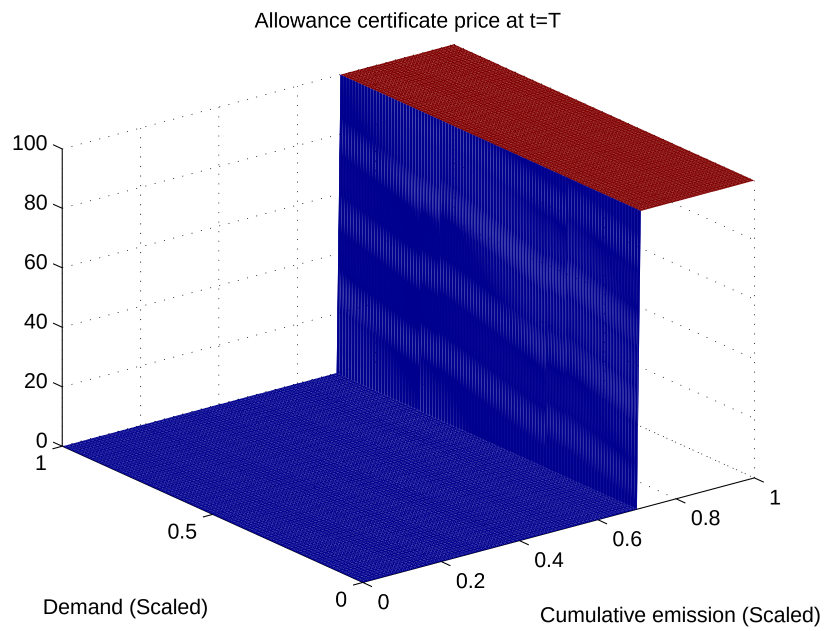


Figure 3.1: Allowance certificate price at t=T

For $t < T$, the allowance price depends on the cumulative emissions to date and the current level of demand. For each fixed level of emission, the function $u(t, D, E)$ is increasing in D, and it correspond to the intuitive idea, since for higher levels of demand, the corresponding market emission rat is greater and consequently it is more likely that the cap will be reached. Similarly, fixing D, the function $u(t, D, E)$

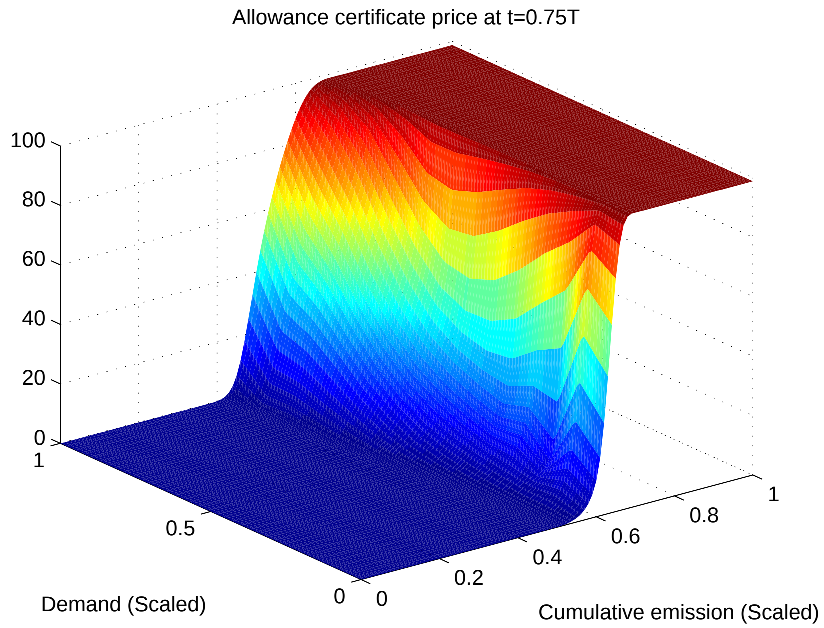


Figure 3.2: Allowance certificate price at t=0.75T

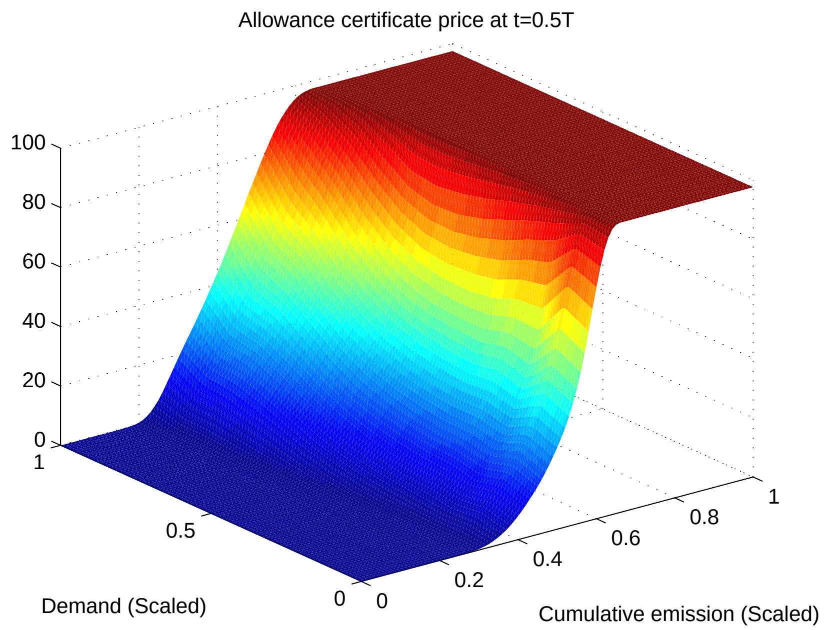


Figure 3.3: Allowance certificate price at t=0.50T

is increasing in D is an increasing function of E. In particular , we can think of the current level of cumulative emissions determining an interval for the allowance price and the demand for electricity setting the exact price within this interval. Further, we notice that the allowance price equals the discounted penalty, if cumulative

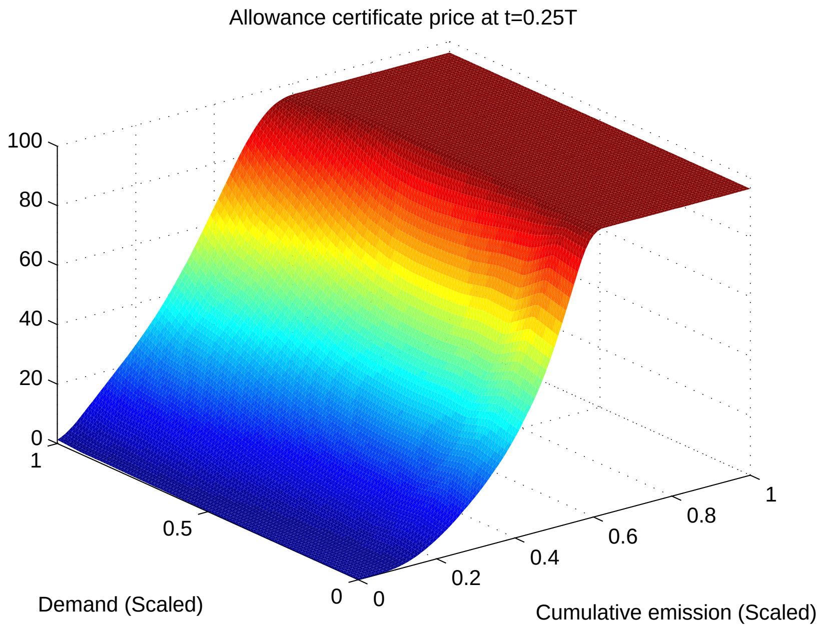


Figure 3.4: Allowance certificate price at t=0.25T

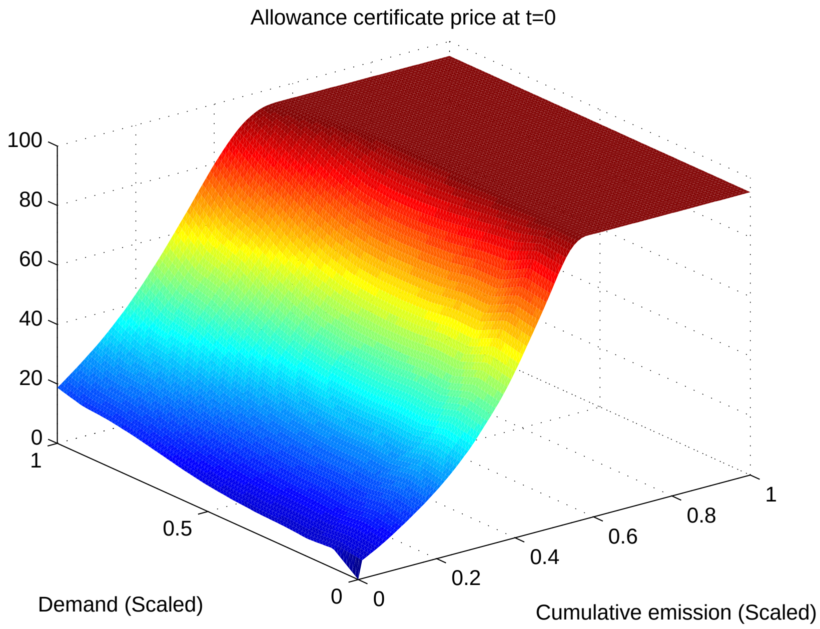


Figure 3.5: Allowance certificate price at t=0

emissions exceed the cap.

Moreover we simulate the processes D_t, E_t and A_t with Mote-Carlo method and we display any of these simulation in figures 3.6, 3.7 and 3.8 As expected the cumulative emission process E_t is strictly increasing and in these simulation it stay

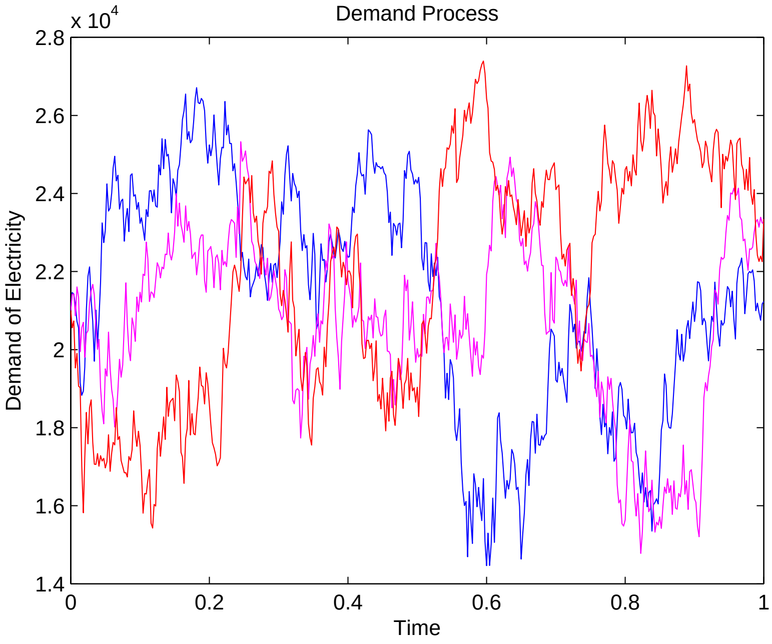


Figure 3.6: Simulation of three paths of D_t

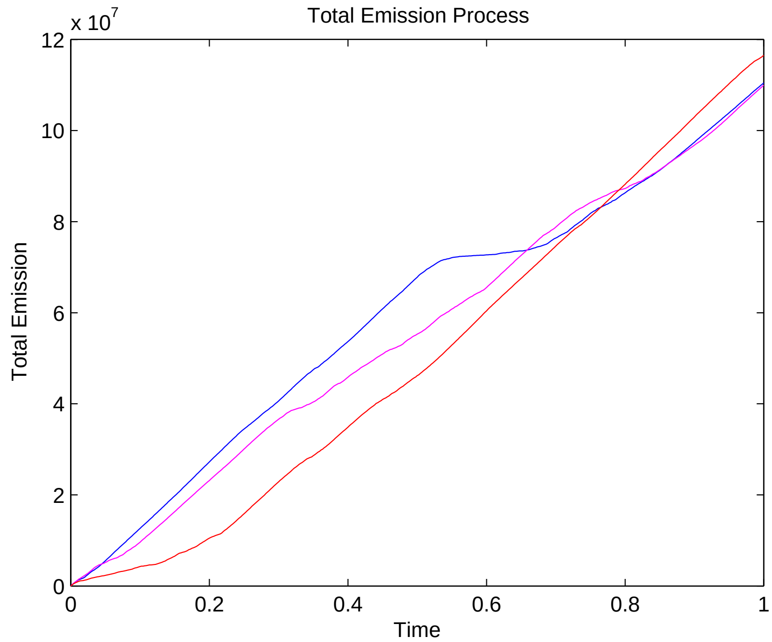


Figure 3.7: Simulation of three paths of E_t

under the cap at terminal time $t = T$. This leads to the value of the process A_t at terminal time, indeed it is equal to zero. This matches the intuitive idea that the certificate has no value if at terminal time total emissions are under the fixed cap E_{cap}.

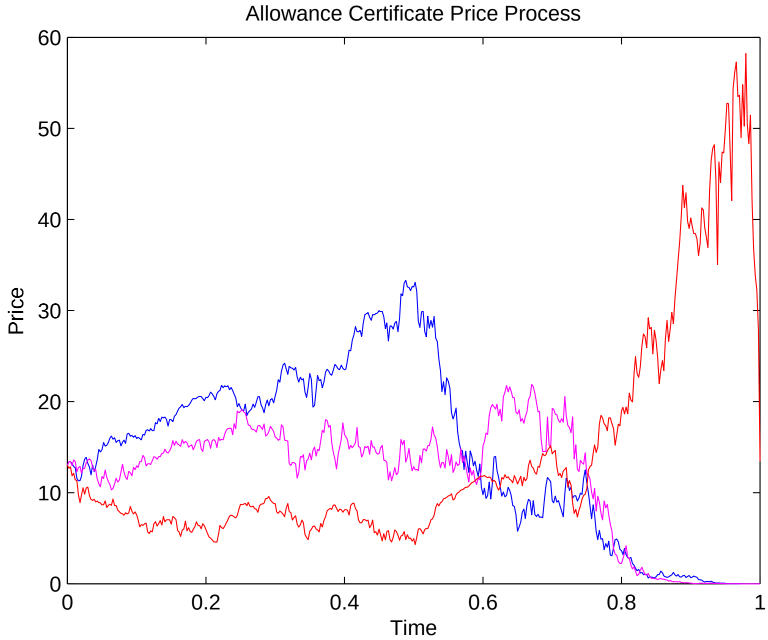


Figure 3.8: Simulation of three paths of A_t

Finally we simulate the the process E_t for different values of the penalty π. Figure 3.9 shows that cumulative emissions at terminal time are a decreasing as the penalty increase. As expected to higher value of the penalty correspond a more aggressive strategy in order to reduce polluting emissions. Moreover, more aggressive regulation now only leads to small reductions in the cumulative emissions, thus our analysis confirms the well known stylized fact that emissions trading cannot incentivise firms to reduce cumulative emissions far below the cap.

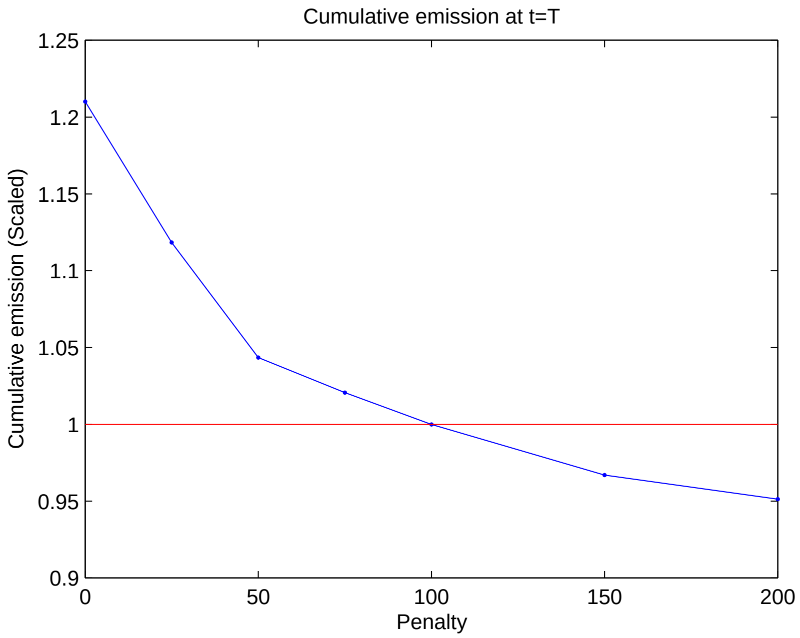


Figure 3.9: Cumulative emissions related to the value of the penalty

Chapter 4

Tax Fraud in EU Emission Trading Scheme

In the EU Emission Trading Scheme the volume of certificated exchanged on the market every year is at least 90 billion euros. In late 2008 Europol launched its inquiry because suspicious trading activities appeared and in announcing its investigations agents said that as much as 90 percent of the entire market volume on emissions exchanges was caused by fraudulent activities, undermining the very viability of the ETS just as the EU is touting a similar scheme for the rest of the world. The peak of exchanges was registered in May, when several hundreds of certificates were bought form brokers in France and Denmark and the price of one credit doubled. Europol estimated that this fraud cost to government coffers about 5 billion euros. The system was simple and very profitable. When anyone resident in one EU member state buys an allowance certificate in a different country, he doesn't have to pay VAT. Thus, thanks to this detail, criminals establish themselves in one EU member state and open a trading account with the national carbon credit registry. Every country has a carbon credit registry which is coordinated by the Cilt (Community independent transaction log) of European Commission. Then, they buy carbon credits in a different country, which makes them exempt from VAT. These are then sold to buyers in the original country, but with VAT attached on, although the VAT just disappears along with the trader and the money never arrives in government coffers. Certificates owned by criminals were very attractive for firms, since its price was lower. Thus firms bought these certificates and criminal organizations made large profit since they didn't pay VAT.

Some member states including France, the Netherlands, Spain and the UK, but not all, changed their taxation rules on such transactions to prevent further losses, and as soon as this particular loophole was closed in the few member states that did deal with the problem, as much as 90 percent of the trading volume disappeared. The European Commission for its part said that while the Europol report needed to be looked into, it was aware of existing faults in the transfer of greenhouse gas credits and that at a recent meeting of EU finance ministers, a general approach to tackle the matter was considered.

Appendix A

Brownian martingale representation theorem

Let $(\Omega, \mathcal{F}, P, (\mathcal{F}_t))$ be a filtered probability space, $(W_t)_{t\in[0,T]}$ a d-dimensional Brownian motion and $u \in \mathbb{L}^2(F_t^W)$ a stochastic process, where $\mathcal{F}_t^W$ is the Brownian filtration induced by W_t. It's well known that the process:

$$M_t = M_0 + \int_0^t u_s dW_s \tag{A.1}$$

with $t \in [0,T]$ is a F_t^W-martingale. In this section we want to prove that given a$\mathcal{F}_t^W$-martingale, using suitable hypothesis, it can be represented by a stochastic process $u \in \mathbb{L}^2(\mathcal{F}_t^W)$ trough an integral as we see in (A.1).

We first recall the definition of martingale:

Definition A.1. Let M be an integrable adapted stochastic process on the filtered probability space $(\Omega, \mathcal{F}, P, (\mathcal{F}_t))$. We say that M is a martingale with respect to $\mathcal{F}_t$ and the measure P if:

$$M_s = E[M_t \| \mathcal{F}_s] \quad \text{for every } 0 \leq s \leq t$$

Definition A.2. Let $(W_t)_{t\in[0,T]}$ be a d-dimensional Brownian motion on the filtered probability space $(\Omega, \mathcal{F}, P, (\mathcal{F}_t))$ and $\lambda \in \mathbb{L}^2_{loc}(\mathcal{F}_t)$. We define the exponential martingale associated to λ as.

$$Z_t^\lambda = exp\left(-\int_0^t \lambda_s dW_s - \frac{1}{2}\int_0^t |\lambda|^2 ds\right) \tag{A.2}$$

Remark A.3. Applying Itô's formula to $Z_t^\lambda wehave$:

$$dZ_t^\lambda = -Z_t^\lambda \lambda_t dW_t \tag{A.3}$$

and therefore Z_t^λ is a local $\mathcal{F}_t$-martingale since $\lambda \in \mathbb{L}^2_{loc}(\mathcal{F}_t)$

Remark A.4. From now until the end of this section, unless otherwise stated, we will always work on the spaces defined in A.2.

We now exhibit these results, leaving out the proof, useful for our purpose.

Proposition A.5. *Let W be a d-dimensionale Brownian motion and $\sigma \in \mathbb{L}^2$ a $N \times d$-matrix such that*

$$\int_0^T |\sigma_s \sigma_s^*| ds \leq k$$

with $k > 0$. Therefore, taking

$$X_t = \int_0^t \sigma_s dW_s$$

For all $\lambda > 0$ we have

$$P\left(\sup_{0 \leq t \leq T} |X_t| \geq \lambda\right) \leq 2Ne^{\frac{\lambda^2}{2kN}}$$

Proposition A.6. *Let X be a random variable on $(\Omega, \mathcal{F}, P)$ e $f \in C^1(\mathbb{R}_{\geq \not{\vdash}})$ such that $f' \geq 0$ or $f' \in L^2(\mathbb{R}_{\geq \not{\vdash}}, P^{|X|})$. Then*

$$E[f(|X|)] = f(0) + \int_0^{+\infty} f'(\lambda) P(|X| \geq \lambda) d\lambda$$

Lemma A.7. *If exist $C \in \mathbb{R}$ such that*

$$\int_0^T |\lambda_t| dt \leq C q.s.$$

then Z^λ is a martingale and

$$E\left[\sup_{0 \leq t \leq T} \left(Z_t^\lambda\right)^p\right] < +\infty \quad \forall p \geq 1$$

In particular $Z^\lambda \in \mathbb{L}^p(\Omega, P)$ $\forall p \geq 1$

Proof. Define

$$\hat{Z}_T = \sup_{0 \leq t \leq T} Z_t^\lambda$$

For all $\xi > 0$, from $A.5$ it follows

$$\begin{aligned} P\left(\hat{Z}_t \geq \xi\right) &\leq P\left(\sup_{0\leq t\leq T} exp\left(-\int_0^t \lambda_s dW_s\right) \geq \xi\right) = \\ &= P\left(\sup_{0\leq t\leq T}\left(-\int_0^t \lambda_s dW_s\right) \geq log(\xi)\right) \leq \\ &\leq c_1 e^{-c_2(log(\xi))^2} \end{aligned}$$

If we take the function $f(x) = x^p$ and we apply proposition A.6 to $\hat{Z}_T$:

$$E\left[\hat{Z}_T^p\right] = p\int_0^{+\infty} \xi^{p-1} P\left(\hat{Z}_T \geq \xi\right)$$

Therefore $Z^\lambda \in \mathbb{L}^p(\Omega, P)$ $\forall p \geq 1$ and Z^λ is a martingale. □

To prove next lemma we need some classical results on random variables. We cite these two results without their proofs.

Proposition A.8. *Let* $(\Omega, F, P, (\mathcal{F}_n))$ *be a filtered probability space and* $X \in L^p(\Omega, P)$ $p > 1$ *a random variable. Then, set* $F_\infty = \sigma(\mathcal{F}_n, n \in \mathbb{N})$*, we have:*

$$\lim_{n\to+\infty} E[X|\mathcal{F}_n] = E[X|F_\infty] \quad \textit{in } L^p$$

Proposition A.9. *Let* X*,*Y *be random variables on* $(\Omega, \mathcal{F})$*. Then* X $\sigma(Y)$*-measurable* $\Leftrightarrow \exists f$ $\mathcal{B}$*-measurable such* $X = f(Y)$*.*

Lemma A.10. *Let* $\{t_n\}_{n\in\mathbb{N}}$ *be dense in* $[0, T]$ *with usual topology. Therefore the collection of random variables* $\varphi(W_{t_1}, \ldots, W_{t_n})$ *with* $\varphi \in C_0^\infty(\mathbb{R}^n)$ *is dense in* $L^2(\Omega, \mathcal{F}_T^W)$*.*

Proof. Set $\mathcal{F}_n = \sigma(W_{t_1}, \ldots, W_{t_n})$, $n \in \mathbb{N}$ a discrete filtration and observe that

$$\mathcal{F}_T^W = \sigma(\mathcal{F}_n, n \in \mathbb{N})$$

Let $X \in L^p(\Omega, \mathcal{F}_T^W, P)$ and we take the discrete martingale

$$X_n = E[X|\mathcal{F}_n] \quad n \in \mathbb{N}$$

Frow proposition it follows

$$\lim_{n\to+\infty} X_n = \lim_{n\to+\infty} E[X|\mathcal{F}_n] = E[X|\mathcal{F}_T^W] \quad \text{in } L^2$$

Now, we have X_n $\mathcal{F}_n$-measurable and $\sigma(W_{t_1}, \dots, W_{t_n})$-measurable, so by proposition $\exists \varphi^n$ $\mathcal{B}$-measurable such that

$$X_n = \varphi^n(W_{t_1}, \dots, W_{t_n})$$

Because of the density, φ^n can be approximated in $L^2(\mathbb{R}^n)$ by a sequence $\varphi^n_k \in C_0^\infty(\mathbb{R}^n)$. Then

$$\lim_{k\to+\infty} \varphi^n_k(W_{t_1}, \dots, W_{t_n}) = X_n$$

We can conclude because we have shown that $\forall X \in L^2(\Omega, \mathcal{F}_T^W)$ $\exists$ a sequence of random variables $\varphi^n_k(W_{t_1}, \dots, W_{t_n}$ with $\varphi^n_k \in C_0^\infty(\mathbb{R}^n)$ such that

$$\lim_{n,k\to+\infty} \varphi^n_k(W_{t_1}, \dots, W_{t_n} = X \quad \text{in } L^2$$

□

Lemma A.11. *Take* $\lambda \in L^\infty([0,T], \mathbb{R}^d)$ *and*

$$Z_t^\lambda = exp\left(-\int_0^t \lambda_s dW_s - \frac{1}{2}\int_0^t |\lambda|^2 ds\right)$$

Then the space of linear combinations of random variables Z^λ *is dense in* $L^2(\Omega, \mathcal{F}_T^W, P)$

Proof. For simplicity we only prove the statement in the case $d = 1$. Prove this claim is equivalent to show that the following equation is true:

$$\left(< Z^\lambda, X >_{L^2(\Omega)} = \int_\Omega X Z^\lambda dP = 0 \Rightarrow X = 0 \quad \text{q.s.}\right) \tag{A.4}$$

We start considering a piecewise function

$$f(\xi) = e^{\xi_1 W_{t_1}, \dots, \xi_n W_{t_n}} \quad \xi \in \mathbb{R}^n, \quad t_1, \dots, t_n \in [0,T], \quad n \in \mathbb{N}$$

By $A.4$ we obtain

$$F(\xi) = < f(\xi), X >_{L^2(\Omega)} = 0 \quad \xi \in \mathbb{R}^n$$

We can extend F on $\mathbb{C}^n$:

$$F(z) = < f(z), X >_{L^2(\Omega)} = 0 \quad z \in \mathbb{C}^n$$

and by the analytic continuation principle we have $F(z) = 0$ on $\mathbb{C}^n$. Now let $\varphi^n(W_{t_1}, \ldots, W_{t_n}) \in C_0^\infty(\mathbb{R}^n)$ and applying the inverse Fourier transform:

$$\begin{aligned}\int_\Omega \varphi^n(W_{t_1}, \ldots, W_{t_n}) X dP &= \int_\Omega \left(\frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{\xi_1 W_{t_1}, \ldots, \xi_n W_{t_n}} \hat{\varphi}(-\xi) d\xi \right) X dP = \\ &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \hat{\varphi}(-\xi) \int_\Omega e^{\xi_1 W_{t_1}, \ldots, \xi_n W_{t_n}} X dP d\xi = \\ &= 0\end{aligned}$$

From lemma $A.10$it follows that $\{\varphi^n(W_{t_1}, \ldots, W_{t_n})\}$ is dense in $L^2(\Omega, \mathcal{F}_T^W, P)$ and we obtain $X = 0$. □

We are ready to show the key result of this section, which states that is possible to represent a random variable $X \in L^2(\Omega, \mathcal{F}_T^W, P)$ by its expectation and the stochastic integral of a process $u \in \mathbb{L}^2$.

Theorem A.12. *For each random variable $X \in L^2(\Omega, \mathcal{F}_T^W, P)$ exist a process $u \in \mathbb{L}^2(\mathcal{F}^W)$ such that*

$$X = E[X] = \int_0^T u_t dW_t$$

Moreover this process is unique.

Remark A.13. The uniqueness of the process u is in the sense of the $m \otimes P$-equivalence (m represent the Lebesgue measure on $[0, T]$), i.e. $u = v$ $m \otimes P$-a.s.. $\Leftrightarrow m \otimes P\left(\{(t, \omega) | u_t(\omega) = v_t(\omega)\}\right) = 0$.

Proof. **Uniqueness.** Let $u, v \in \mathbb{L}^2(\mathcal{F}^W)$ such that

$$X = E[X] = \int_0^T u_t dW_t$$

$$X = E[X] = \int_0^T v_t dW_t$$

Subtracting these two equations we obtain

$$0 = \int_0^T (u_t - v_t) dW_t$$

Now, from the Itô isometry it follows

$$\begin{aligned}E\left[\int_0^T (u_t - v_t) dt\right] &= E\left[\int_0^T (u_t - v_t) dW_t\right] = \\ &= 0\end{aligned}$$

Then $u - v = 0$ $m \otimes P$-a.s. $\Rightarrow u = v$ $m \otimes P$-a.s.

Existence. First of all the existence will be proved in the case of the deterministic function Z_T^λ con $\lambda \in L^\infty([0,T], \mathbb{R}^d)$ f, then it we are going to extend it for all $X \in L^2(\Omega, \mathcal{F}_T^W, P)$. Assume X of the form

$$Z_T^\lambda = exp\left(- \int_0^T \lambda_t dW_t - \frac{1}{2} \int_0^T |\lambda|^2 dt \right)$$

By what we have seen until now, thanks to Itô's formula, we have

$$dZ_t^\lambda = -Z_t^\lambda \lambda_t dW_t$$

and by lemma $A.7$ it follows $\lambda_t Z_t^\lambda \in \mathbb{L}^2$ because λ is bounded. Then

$$X = 1 - \int_0^T \lambda_t Z_t^\lambda dW_t$$

This proves the claim for the class $X = Z_T^\lambda$. Now let $X \in L^2(\Omega, \mathcal{F}_T^W)$, by lemma $A.11$ X it can be approximated in $L^2(\Omega, \mathcal{F}_T^W)$ by a sequence of random variables $(X_n)_{n \in \mathbb{N}}$, where X_n is a linear combination of random variables of the form Z_T^λ con $\lambda \in L^\infty([0,T], \mathbb{R}^d)$. Therefore exist a process $u^n \in \mathbb{L}^2(\mathcal{F}^W)$ such that

$$X_n = E[X_n] + \int_0^T u_t^n dW_t$$

For $n, m \in \mathbb{N}$ we evaluate

$$\begin{aligned} E[(X_n - X_m)^2] = E\left[\left(E\left[X_n - X_m\right] + \int_0^T \left(u_t^n - u_t^m\right) dW_t \right) \right] = \\ = E\left[X_n - X_m\right]^2 + E\left[\int_0^T \left(u_t^n - u_t^m\right)^2 dt \right] + \\ + 2E\left[E\left[X_n - X_m\right] \int_0^T \left(u_t^n - u_t^m\right) dW_t \right] = \\ = E\left[X_n - X_m\right]^2 + E\left[\int_0^T \left(u_t^n - u_t^m\right)^2 dt \right] + \\ + 2E\left[X_n - X_m\right] E\left[\int_0^T \left(u_t^n - u_t^m\right) dW_t \right] = \\ = E\left[X_n - X_m\right]^2 + E\left[\int_0^T \left(u_t^n - u_t^m\right)^2 dt \right] \end{aligned}$$

$$\Rightarrow E\left[\int_0^T \left(u_t^n - u_t^m\right)^2 dt \right] = E[(X_n - X_m)^2] - E\left[X_n - X_m\right]^2 \xrightarrow[n,m \to +\infty]{} 0$$

Then $\{u^n\}$ is a Cauchy sequence in $\mathbb{L}^2(\mathcal{F}^W)$ and the following limit exists:

$$\lim_{n \to +\infty} u^n = u \in \mathbb{L}^2(\mathcal{F}^W)$$

In that case

$$X_n = E[X_n] + \int_0^T u_t^n dW_t \xrightarrow[n \to +\infty]{} X = E[X] + \int_0^T u_t dW_t$$

and the claim is completely proved. □

Theorem A.14. *Let $(M_t)_{t \in [0,T]}$ be a $\mathcal{F}^W$-martingale such that $M_T \in \mathbb{L}^2(\mathcal{F}_T^W)$. Then $\exists! u \in \mathbb{L}^2(\mathcal{F}^W)$ such that*

$$M_t = M_0 + \int_0^t u_s dW_s \quad q.s. \quad \forall t \in [0,T]$$

Remark A.15. As we have seen in the previous theorem, the uniqueness has to be intended in the sense of remark $A.13$

Proof. Let $M_T \in \mathbb{L}^2(\mathcal{F}_T^W)$ then, by theorem $A.12$, $\exists! u \in \mathbb{L}^2(\mathcal{F}^W)$ such that

$$M_T = E[M_T] + \int_0^T u_s dW_s \quad \text{q.s.}$$

We have $E[M_T] = M_0$ because M_T is a martingale and

$$M_T = M_0 + \int_0^T u_s dW_s \quad \text{q.s.}$$

Now, fix $t \leq T$

$$\begin{aligned} M_t &= E\left[M_T \middle| \mathcal{F}_t^W\right] = \\ &= E\left[M_0 + \int_0^T u_s dW_s \middle| \mathcal{F}_t^W\right] = \\ &= M_0 + \int_0^t u_s dW_s + E\left[\int_t^T u_s dW_s \middle| \mathcal{F}_t^W\right] = \\ &= M_0 + \int_0^t u_s dW_s \end{aligned}$$

□

The above result can also be shown if $(M_t)_{t \in [0,T]}$ is a local martingale, i.e. we have the following theorem

Theorem A.16. *Let $(M_t)_{t \in [0,T]}$ be a local $\mathcal{F}^W$-martingale. Then $\exists! u \in \mathbb{L}^2_{loc}(\mathcal{F}^W)$ such that*

$$M_t = M_0 + \int_0^t u_s dW_s \quad q.s. \quad \forall t \in [0,T]$$

Appendix B

Linear SDE and BSDE

B.1 Linear Stochastic Differential Equations

In this section we are going to show some important results of linear SDE (Stochastic Differential Equations) without provide to demonstrate them. For a complete report on this topic see [1].
Considered the filtered probability space $(\Omega, \mathcal{F}, P, \mathcal{F}_t)$. By linear SDE in $\mathbb{R}^N$ we mean an equation of the form

$$dX_t = (B(t)X_t + b(t))dt + \sigma(t)dW_t \tag{B.1}$$

where W_t is a d-dimensional Brownian motion ($d \leq N$) and B, b, σ are functions in L^∞_{loc} such that

$$B : \mathbb{R}^{\geq 0} \to \mathbb{R}^N \times \mathbb{R}^N$$

$$b : \mathbb{R}^{\geq 0} \to \mathbb{R}^N \times \mathbb{R}$$

$$\sigma : \mathbb{R}^{\geq 0} \to \mathbb{R}^N \times \mathbb{R}^d$$

Moreover it's given an initial condition $X_0 = Z$. Furthermore we assume that the following conditions are satisfied:

- $Z \in L^2(\Omega, P)$ and $\mathcal{F}_0$-measurable
- $B(t)X_t + b(t)$ is locally Lipschitz respect to X_t and uniformly continuous respect to t.

- σ uniformly continuous respect to t.

- $B(t)X_t + b(t)$ has linear growth respect to X_t.

From the SDE general theory, it follows that the linear SDE with initial conditions $X_0 = Z$ admit an unique solution. The uniqueness of the solution has to be meant in the sense of indistinguishable process.
Let's introduce the Cauchy problem

$$\begin{cases} \Phi'(t) = B(t)\Phi(t) \\ \Phi(t_0) = I_N \end{cases} \tag{B.2}$$

where I_N is the identity matrix $N \times N$.

Proposition B.1. *Let $X_0^x = x \in \mathbb{R}^N$ be the initial condition associated to the equation B.1. Then the solution of that equation is*

$$X_t^x = \Phi(t)\left(x + \int_0^t \Phi^{-1}(s)b(s)ds + \int_0^t \Phi^{-1}(s)\sigma(s)dW_s\right) \tag{B.3}$$

Moreover the process X_t^x has multi-normal distribution for all $t > 0$ where

$$m_x(t) = E\left[X_t^x\right] = \Phi(t)\left(x + \int_0^t \Phi^{-1}(s)b(s)ds\right) \tag{B.4}$$

$$C(t) = cov(X_t^x) = \Phi(t)\left(\int_0^t \Phi^{-1}(s)\sigma(s)\left(\Phi^{-1}(s)\sigma(s)\right)^*\right)\Phi^*(t) \tag{B.5}$$

The matrix $C(t)$ is positive semi-definite because $d \leq N$. By definition $X_t^x \sim \mathcal{N}_{m_x(t),C(t)}$ and it means thatX_t^x has the same characteristic function of $\mathcal{N}_{m_x(t),C(t)}$, i.e.

$$\varphi_{X_t^x}(\xi) = exp\left(i < \xi, m_x(t) > -\frac{1}{2} < C(t)\xi, \xi >\right)$$

where $C(t)$ is a symmetric semi-definite positive matrix. Therefore, generally speaking, it hasn't normal density. For fixed $t > 0$, the matrix $C(t)$ must to be symmetric definite in order that X_t^x has normal distribution, and in this case we have the function of the variable $y \in \mathbb{R}^N$:

$$\Gamma(0, x; t, y) = \frac{(2\pi)^{-\frac{N}{2}}}{\sqrt{det(C(t))}} exp\left(-\frac{1}{2} < C^{-1}(t)(y - m_x(t)), (y - m_x(t)) >\right) \tag{B.6}$$

for all $x \in \mathbb{R}^N$, $t \in [0,T]$. The function Γ is said to be the transition density of the process X_t^x. More generally we give the following definition of Γ with $t < T$:

$$\begin{aligned}\Gamma(t,x;T,y) =& \frac{(2\pi)^{-\frac{N}{2}}}{\sqrt{det(C(T-t))}} \cdot \\ & \cdot exp\left(-\frac{1}{2} < C^{-1}(T-t)(y - m_x(T-t)), (y - m_x(T-t)) >\right)\end{aligned} \tag{B.7}$$

The transition density of the SDE isrelated to the solution of the Cauchy problem

$$\begin{cases} Lu = f, & \text{in } \mathcal{S}_T =]0,T[\times\mathbb{R}^N \\ u(T,\cdot) = \varphi \end{cases} \tag{B.8}$$

where f, φ are given functions, $c_{ij} = \sigma\sigma^*$ and $\mathcal{A}$ is the characteristic operator of the SDE obtained employing Itô's formula. In the case of linear SDE we have:

$$L = \frac{1}{2}\sum_{i,j=1}^{N} c_{ij}(t)\partial_{x_i x_j} + < b(t) + B(t)x, \nabla > + \partial_t \tag{B.9}$$

where $\nabla = (\partial_{x_i}, \ldots, \partial_{x_N})$. The relation between fundamental solution of the operator L and transition density is given by the following theorem:

Theorem B.2. *If operator L admit fundamental solution, then it ie equal to the transition density of the SDE B.1.*

In the previous theorem we suppose the existence of a fundamental solution, so it's natural try to understand under which hypothesis L has a fundamental solution.

Definition B.3. The operator L is said to be uniformly parabolic if exist a positive number $\lambda > 0$ such that

$$\lambda^{-2}|\xi|^2 \leq \sum_{i,j=1}^{N} c_{ij}(t)\xi_i\xi_j \leq \lambda|\xi|^2 \tag{B.10}$$

where $t \in \mathbb{R}^{\geq 0}$, $\xi \in \mathbb{R}^N$

Definition B.4. Let $\alpha \in]0,1]$ and O a subset of $\mathbb{R}^{N+1}$.$C_\alpha^P(O)$ is the space of functions u which are bounded on O such that

$$|u(t,x) - u(s,y)| \leq C\left(|t-s|^{\frac{\alpha}{2}} + |x-y|^\alpha\right) \tag{B.11}$$

where $(t,x),(s,y) \in O$. A function u is said to be bounded and Hölder continuous if $u \in C_\alpha^P(O)$ for some $\alpha \in]0,1]$.

Theorem B.5. *Under the hypothesis:*

- *L uniformly parabolic in $\mathbb{R}^{N+1}$;*
- *$c_{ij}, \bar{b}_j$ bounded and Hölder continuous for all $1 \le i, j \le N$, where $\bar{b}_j(t, x) = b_i(t) + B_{ij}(t)x_j$;*

the operator L has fundamental solution $\Gamma = \Gamma(t, x; s, y)$ for $x, y \in \mathbb{R}^N$ e $t > s$. Moreover f if φ are continuous functions and there exist two positive constants c e $\gamma < 2$ such that:

- $|\varphi(x)| \le ce^{c|x|^{\gamma}}$ $x \in \mathbb{R}^N$
- $|f(t, x)| \le ce^{c|x|^{\gamma}}$ $(t, x) \in \mathcal{S}_T$
- *f locally Hölder continuous respect to the variable x and uniformly continuous respect to t.*

Le function u defined by

$$u(t, x) = \int_{\mathbb{R}^N} \Gamma(t, x; T, y)\varphi(y)dy + \int_t^T \int_{\mathbb{R}^N} \Gamma(t, x; s, y)f(s, y)dyds \tag{B.12}$$

with $(t, x) \in \mathcal{S}_T$ e $u(T, x) = \varphi(x)$ is classical solution of the Cauchy problem B.8

The property to be uniform parabolic of the operator L implies that $C(t)$ is positive definite, but, in general,the conversely isn't true; Indeed, as shown in the following example, exist linear SDE whose operator isn't uniformly parabolic, but the SDE has transition density anyway.

Example B.6. Consider this simplified form of the Langervin equation $\mathbb{R}^2$

$$\begin{cases} dX_t^1 = dW_t, \\ dX_t^2 = X_t^1 dt, \end{cases}$$

which describes the trajectory of a particle in the phase space. In particular X_t^1 represent the velocity and X_t^1 the position. This SDE is clearly linear with $d = 1$ and $N = 2$, and the coefficient matrices are:

$$B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

$$\sigma = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$b = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

The solution of the Cauchy problem B.2 is $\Phi(t) = e^{tB} = \sum_{n=0}^{\infty} \frac{(tB)^n}{n!}$ Observe that B è nilpotent, indeed we have $B^2 = 0$. then

$$\Phi(t) = \begin{pmatrix} 1 & 1 \\ t & 0 \end{pmatrix}$$

Moreover

$$c = \sigma\sigma^* = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

By proposition B.1, stated $x = (x_1, x_2)$, we can compute expectation and covariance matrix:

$$m_x(t) = e^{tB}x = (I + tB)x = (x_1, x_2 + tx_1)$$

$$C(t) = \begin{pmatrix} t & \frac{t^2}{2} \\ \frac{t^2}{2} & \frac{t^3}{3} \end{pmatrix}$$

Subsequently $det(C(t)) = \frac{t^4}{12} > 0$ with $t > 0$, and it follows that C is positive definite for $t > 0$. This fact implies that the SDE has transition density and we can compute it to have and explicit function. We obtain:

$$\begin{aligned} \Gamma(t, x; T, y) = & \frac{\sqrt{3}}{\pi(T-t)^2} \cdot \\ & \cdot\, exp\left(-\frac{y_1 - x_1}{2(T-t)} - \frac{3(2y_2 - 2x_2 - (T-t)(y_1 + x_1))}{2(T-t)^3} > \right) \end{aligned} \tag{B.13}$$

Now we can compute the differential operator L associated to the SDE and it results

$$L = \frac{1}{2}\partial_{x_1 x_2} + x_1 \partial_{x_2} + \partial t$$

Since the second derivative coefficient matrix $c = \sigma\sigma^*$ is degenerate, the operator L isn't uniformly parabolic. Nevertheless the SDE has transition density and L has Gaussian fundamental solution $B.13$.

Our aim is to find necessary and sufficient conditions to establish when the covariance matrix associated to the process X_t is definite positive, independently form the uniform parabolicity of the operator L. Applying control theory arguments, it can be shown that Kalman e Hörmander conditions are equivalent to say that $C(t)$ sia positive definite. Hereafter, for shake of simplicity, we assume the matrices B and σ to be constant and independent of time. Moreover we assume σ to have maximal rank d.

Remark B.7. Note that the covariance matrix $B.5$ doesn't depend on $b(t)$

Definition B.8. The pair (B, σ) is said to verify Kalman condition if the block matrix of dimension $(N \times (Nd))$ defined by

$$(\sigma \ B\sigma \ B^2\sigma \ \cdots \ B^{N-1}\sigma)$$

has maximal rank, i.e. it has rank N.

Theorem B.9 (Kalman rank condition)**.** *The matrix $C(t)$ is positive definite with $t > 0$ if and only if the pair (B, σ) satisfy Kalman condition.*

Remark B.10. Kalman rank condition is independent of t.

Definition B.11. Let X, Y be vector fields from $\mathbb{R}^N$ to $\mathbb{R}^N$. The commutator of X with Y is:

$$[X, Y] = XY - XY$$

Remark B.12. The commutator of two vector fields is still a vector field. For a proof of this fact and further results see [9].

Theorem B.13 (Hörmander operator condition)**.** *Let $\partial_{x_1}, \ldots, \partial_{x_d}$ vector fields, $Y =< Bx, \nabla >$ and consider the Kolmogorov type operator with constant coefficients*

$$L = \frac{1}{2}\Delta_{\mathbb{R}^d} + < b + Bx, \nabla > + \partial_t$$

associated to the linear SDE

$$dX_t = (BX_t + b)dt + \sigma dW_t \tag{B.14}$$

Then $C(t) > 0$ with $t > 0$ is equivalent to: "For all $x \in \mathbb{R}^N$ the dimension of the vector space generated by $\partial_{x_1}, \ldots, \partial_{x_d}, Y =< Bx, \nabla >$ and and their commutators is maximal, i.e. N".

Remark B.14. Hörmander condition is a condition on the differential operator L. This criterion has been introduced in the PDE study, see [8].

Remark B.15. Kalman and Hörmander are equivalent for the SDE $B.14$. The proof follows immediately because they are both equivalent to $C(t) > 0$.

B.2 Backward Stochastic Differential Equations

Throughout this section we suppose to have a d-dimensional Brownian motion W_t on the filtered probability space $(\Omega, \mathcal{F}, P, \mathcal{F}_t)$.

Consider a Cauchy problem. Fix an initial or a terminal condition for the equation is conceptually the same, because the method to solve the problem is the same. This isn't true for stochastic differential equations, indeed the initial condition is a point $x \in \mathbb{R}^N$, but the final one is a stochastic process $\xi \in \mathbb{L}^2(\mathcal{F}_T)$. We are looking for a solution of the SDE which is an adapted process, and this is the reason why we need a different method for the backward problem. As we can see in the following basic example, if we use a standard approach to the problem we wont obtain an adapted solution.

Example B.16. Consider the backward problem

$$\begin{cases} dY_t = 0, \quad t \in [0, T], \\ Y_T = \xi, \end{cases}$$

where $\xi \in \mathbb{L}^2(\mathcal{F}_T)$ is a stochastic process. The unique solution is $Y_t = \xi \ \forall t \in [0, T]$, but ξ isn't necessarily a random variable $\mathcal{F}_t$-measurable $\forall t \in [0, T]$ because we only know that $\xi \in \mathbb{L}^2(\mathcal{F}_T)$, i.e ξ is $\mathcal{F}_T$-measurable. Then the solution process isn't adapted to the filtration. To solve this problem one way to proceed is the following: we modify the solution setting

$$Y_t = E\left[\xi | \mathcal{F}_t\right] \quad t \in [0, T]$$

This way $Y_T = \xi$ and Y_t is adapted to the filtration. Moreover Y_t is an $\mathcal{F}_t$-martingale. Suppose that $\mathcal{F}_t$ is a Brownian filtration. If it isn't true we can extend the the probability space $(\Omega, \mathcal{F}, P, \mathcal{F}_t)$ to the space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P}, \widetilde{\mathcal{F}}_t)$ on which the this hypothesis is satisfied, see [5]. From now we assume to work in the probability space just defined. Then for the representation theorem for Brownian martingale $A.14$ exist $Z_t \in \mathbb{L}^2$ such that

$$Y_t = Y_0 + \int_0^t Z_s dW_s \quad \text{q.s.} \quad \forall t \in [0, T] \tag{B.15}$$

Now we can reformulate the problem in differential form:

$$\begin{cases} dY_t = Z_t dW_t \quad t \in [0, T], \\ Y_T = \xi, \end{cases}$$

Then by solution of $B.16$ we mean a pair of adapted process (Y_t, Z_t), and thanks to this trick it's possible to find an adapted solution to problem $B.16$. Roughly speaking, change the definition of solution, adding a new component Z_t, allow us to find an adapted solution.

We can represent a BSDE in a different way through an integral formulation. Sice what we seen until now, we can rewrite $Y_T = \xi$ as

$$Y_T = Y_0 + \int_0^T Z_s dW_s$$

and we can deduce:

$$\begin{aligned} Y_0 &= Y_T - \int_0^T Z_s dW_s \\ &= \xi - \int_0^T Z_s dW_s \end{aligned}$$

Putting this equation back into $B.15$ we have:

$$\begin{aligned} Y_t &= Y_0 + \int_0^t Z_s dW_s \\ &= \xi - \int_t^T Z_s dW_s \quad \forall t \in [0, T] \end{aligned} \tag{B.16}$$

The last stochastic integral isn't a backward Itô's integral, but a usual one. Therefore it represent an usual stochastic differential equation. Applying Itô's isometry to $|Y_t|^2$ and keeping in mind expectation's property it results:

$$E\left[|\xi|^2\right] = E\left[|Y_t|^2\right] + E\left[\int_t^T |Z_s|^2 ds\right] \quad \forall t \in [0, T] \tag{B.17}$$

form which follow that if we have two solutions (Y_t, Z_t), $(\widetilde{Y}_t, \widetilde{Z}_t)$ of the same equation with the same terminal value ξ then they are indistinguishable. Indeed we have:

$$0 = E\left[|\xi - \xi|^2\right] = E\left[|Y_t - \widetilde{Y}_t|^2\right] + E\left[\int_t^T |Z_s - \widetilde{Z}_s|^2 ds\right] \quad \forall t \in [0,T] \tag{B.18}$$

therefore $Y_t = \widetilde{Y}_t$ ans $Z_t = \widetilde{Z}_t$. Since equation $B.16$ is linear we have the uniqueness of solution. At the end of this example we observe that ifξ is a non constant random variable then by uniqueness of the solution we have $Y_t = \xi$ and $Z_t = 0$ solution of the problem $B.16$ because they satisfy $B.16$. Note that we achieve our purpose modifying the definition of solution to obtain an adapted solution of the BSDE.

Generally we have to solve a system composed by two stochastic differential equation: one forward and one backward. This pair of equations is called FBSDE (Forward Backward Stochastic Differential Equation) and we can consider it like a generalization of a backward problem. Until now we only give an intuitive idea of a solution of a BSDE; Our aim is to give a formal definition of solution of a FBSDE. To this purpose, we need to establish notations and set some spaces:

- $\mathcal{L}^2_{\mathcal{F}}(\Omega; C([0,T]); \mathbb{R}^n)$= space of all continuous stochastic processes X_t and $\mathcal{F}_t$-adapted which take value in $\mathbb{R}^n$ such that $E\left[\sup_{t\in[0,T]} |X_t|\right] < \infty$.

- $\mathcal{L}^2_{\mathcal{F}}(0,T; W^{1,\infty}(M,N))$=set of all functions $f : [0,T] \times M \times N \times \Omega \to N$ (where M, N Euclidean spaces) such that, $\forall \theta \in M$ fixed, the mapping $(t,\omega) \mapsto f(t,\theta,\omega)$ define an $\mathcal{F}_t$-adapted process with $f(t,0,\omega) \in \mathbb{L}^2_{\mathcal{F}}([0,T]; N)$. Moreover f must be Lipschitz respect to the variable θ almost surely.

- $\mathcal{L}^2_{\mathcal{F}_T}(\Omega; W^{1,\infty}(\mathbb{R}^n, \mathbb{R}^m))$= set of all functions $g : \mathbb{R}^n \times \Omega \to \mathbb{R}^m$ such that $\forall x \in \mathbb{R}^n$ fixed $\omega \mapsto (x,\omega)$ be $\mathcal{F}_T$-measurable and g be uniformly Lipschitz in $\mathbb{R}^n$. Moreover it must be $g(0,\omega) \in \mathbb{L}^2_{\mathcal{F}}$.

- $\mathcal{M}[0,T] = \mathcal{L}^2_{\mathcal{F}}(\Omega; C([0,T]); \mathbb{R}^n) \times \mathcal{L}^2_{\mathcal{F}}(\Omega; C([0,T]); \mathbb{R}^m) \times \mathbb{L}^2_{\mathcal{F}}([0,T]; \mathbb{R}^l)$

Consider a FBSDE in his general form:

$$\begin{cases} dX_t = b(t, X_t, Y_t, Z_t)dt + \sigma(t, X_t, Y_t, Z_t)dW_t, \\ dY_t = h(t, X_t, Y_t, Z_t)dt + \widehat{\sigma}(t, X_t, Y_t, Z_t)dW_t, \\ X_0 = x, \\ Y_T = g(X_T). \end{cases} \tag{B.19}$$

with $x \in \mathbb{R}^n$ and $b, h, \sigma, \widehat{\sigma}, g$ functions such that, set $M = \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^l$, satisfy the following hypothesis:

- $b \in \mathcal{L}^2_{\mathcal{F}}(0, T; W^{1,\infty}(M, \mathbb{R}^n))$;
- $\sigma \in \mathcal{L}^2_{\mathcal{F}}(0, T; W^{1,\infty}(M, \mathbb{R}^{n\times d}))$;
- $h \in \mathcal{L}^2_{\mathcal{F}}(0, T; W^{1,\infty}(M, \mathbb{R}^m))$;
- $\widehat{\sigma} \in \mathcal{L}^2_{\mathcal{F}}(0, T; W^{1,\infty}(M, \mathbb{R}^{m\times d}))$;
- $g \in \mathcal{L}^2_{\mathcal{F}_T}(\Omega; W^{1,\infty}(\mathbb{R}^n, \mathbb{R}^m)$;

Definition B.17. A triple of continuous stochastic process $(X, Y, Z) \in \mathcal{M}[0, T]$ is said to be adapted solution to the problem $B.19$ if, almost surely and $\forall t \in [0, T]$, we have:

$$\begin{cases} X_t = x + \int_0^t b(s, X_s, Y_s, Z_s)ds + \int_0^t \sigma(s, X_s, Y_s, Z_s)dW_s, \\ Y_t = g(X_T) - \int_t^T h(s, X_s, Y_s, Z_s)ds - \int_t^T \widehat{\sigma}(s, X_s, Y_s, Z_s)dW_s. \end{cases} \tag{B.20}$$

There are examples which show that, under this assumption, the solution may not exist.For further results see [6].

Appendix C

Markov process and Blumenthal Zero-One Law

In this part we introduce Markov process and we see that Brownian motion is an example of Markov process, and finally we show the Blumenthal 0-1 law. For simplicity, we can assume that all the filtrations are right-continuous, and this is not a restrictive assumption. Indeed we can construct filtrations with this property from the one considered, for example see [5].

Let (E, d) a metric space, and consider the completion $\overline{\mathcal{B}(E)^{\mu}}$ of the Borel σ-field $\mathcal{B}(E)$ respect to the finite measure μ on $(E, \mathcal{B}(E))$. The universal σ-field $\mathcal{U}(\mathcal{E}) = \bigcap_{\mu} \overline{\mathcal{B}(E)^{\mu}}$, where intersection is over all finite measures μ. A real-valued function is said to be universally measurable if it's $\mathcal{U}(S)$-measurable.

Definition C.1. A d-dimensional Brownian family is a d-dimensional process $W = \{W_t; t \geq 0\}$ on a measurable space $(\Omega, \mathcal{F})$ adapted to the filtration $\mathcal{F}_t$, together with a family of probability measures $\{P^x\}_{x \in \mathbb{R}^d}$ such that:

- for each $F \in \mathcal{F}$, the mapping $x \mapsto P^x(F)$ is universally measurable;
- for each $x \in \mathbb{R}^d$, $P^x([W_0 = x]) = 1$;
- under each P^x, the process W is a d-dimensional Brownian motion starting at x.

Definition C.2. Let d a positive integer and μ a probability measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, and $(\Omega, \mathcal{F}, P^\mu, \mathcal{F}_t)$ a filtered probability space. An adapted, d-dimensional process X is said to be a strong Markov process with initial distribution μ if:

- $P^\mu[X_0 \in \Gamma] = \mu(\Gamma)$, $\forall \Gamma \in \mathcal{B}(\mathbb{R}^d)$;
- for any optional time S of $\{\mathcal{F}_t\}$, $t \geq 0$ and $\Gamma \in \mathcal{B}(\mathbb{R}^d)$, we have $P^\mu[X_{S+t} \in \Gamma \| \mathcal{F}_{S+}] = P^\mu[X_{S+t} \in \Gamma \| X_S]$.

Definition C.3. Let d a positive integer and X_t an adapted, d-dimensional process on $(\Omega, \mathcal{F}, \mathcal{F}_t)$. X_t, together with a family of probability measure $\{P^x\}_{x \in \mathbb{R}^d}$ on $(\Omega, \mathcal{F}, \mathcal{F}_t)$, is said to be a d-dimensional strong Markov family if:

- for each $F \in \mathcal{F}$, the mapping $x \mapsto P^x(F)$ is universally measurable;
- for all $x \in \mathbb{R}^d$, $P^x([X_0 = x]) = 1$;
- for each $x \in \mathbb{R}^d$, $t \geq 0$, $\Gamma \in \mathcal{B}(\mathbb{R}^d)$, and optional time S of $\mathcal{F}_t$;, we have $P^x[X_{S+t} \in \Gamma \| \mathcal{F}_{S+}] = P^x[X_{S+t} \in \Gamma \| X_S]$, $P^x - a.s.$ on $\{S < \infty\}$;
- for each $x \in \mathbb{R}^d$, $t \geq 0$, $\Gamma \in \mathcal{B}(\mathbb{R}^d)$, and optional time S of $\mathcal{F}_t$;, we have $P^x[X_{S+t} \in \Gamma \| X_S = y] = P^y[X_t \in \Gamma]$, $P^x X_S^{-1} - a.e.$ y.

Remark C.4. It's also possible to define a Markov process and a Markov family. For their definitions it's sufficient to substitute the optional time S with an real number $s \in \mathbb{R}$ in the definitions above. In particular we have that a Markov process is a strong Markov process and Markov family is a strong Markov family.

Moreover it can be shown

Theorem C.5. *A d-dimensional Brownian motion is a Markov process, and a Brownian family is a Markov family. The statement is also true for strong Markov process and family.*

Consider a strong Markov process X with initial distribution μ on the probability space $(\Omega, \mathcal{F}^X_\infty, P^\mu)$, where $\mathcal{F}^X_\infty = \sigma\left(\bigcup_{t \geq 0} \mathcal{F}^X_t\right)$ and $\mathcal{F}^X_t = \sigma(X_s; 0 \leq s \leq t)$. The right-continuous filtration which makes this strong Markov process adapted is the augmented filtration $\{\mathcal{F}^\mu_t\}_{t \geq 0}$, where for each t $\mathcal{F}^\mu_t = \sigma\left(\mathcal{F}^X_t \cup \mathcal{N}^\mu\right)$, $\mathcal{N}^\mu = \{F \subseteq$

$\Omega \| \exists t \geq 0, \exists G \in \mathcal{F}_t^X$, with $F \subseteq G$, $P^\mu(G) = 0\}$ is the collection of the sets which have null probability respect to P^μ for some $t \geq 0$. This change of filtrations doesn't affect the property to be a Brownian motion or a strong Markov process indeed:

Theorem C.6. *A d-dimensional Brownian motion W with initial distribution μ on $(\Omega, \mathcal{F}_\infty^W, P^\mu)$ relative to the filtration $\{\mathcal{F}_t^\mu\}_{t\geq 0}$ is still a d-dimensional Brownian motion. A d-dimensional strong Markov process X with initial distribution μ on $(\Omega, \mathcal{F}_\infty^X, P^\mu)$ relative to the filtration $\{\mathcal{F}_t^\mu\}_{t\geq 0}$ is still a d-dimensional strong Markov process.*

The filtration defined above is dependent on μ, so it's inappropriate for Markov and Brownian family, because we have continuum of initial conditions. Motivated by this remark we want to construct a right-continuous filtrations which makes the strong Markov process adapted and independent from the initial distributions. This filtration is called "Universal filtration". Consider μ a probability measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R})^d)$ and a strong Markov family X,$\{P^x\}_{x\in\mathbb{R})^d}$ on $(\Omega, \mathcal{F}_\infty^X)$.Define

$$P^\mu(F) = \int_{\mathbb{R}^d} P^x(F)\mu(dx), \quad \forall F \in \widetilde{\mathcal{F}}_\infty^X$$

Now take the augmented filtration with intersection all over probability measures μ

$$\widetilde{\mathcal{F}}_t = \bigcup_\mu \mathcal{F}_t^\mu, \quad 0 \leq t \leq \infty$$

This filtration is independent of μ as we want and it can be shown that it's right continuous. Moreover we have the chain of inclusions $\mathcal{F}_t^X \subseteq \widetilde{\mathcal{F}_t} \subseteq \mathcal{F}_t^\mu$,therefore it follows that if X is strongly Markovian with both filtrations $\{\mathcal{F}_t^X\}_{t\geq 0}$ and $\{\mathcal{F}_t^\mu\}_{t\geq 0}$ respect to the probability measure P^μ, X is Markovian with the filtration $\widetilde{\mathcal{F}}_t$. Finally we have this fundamental theorem:

Theorem C.7. *If W, $\{P^x\}_{x\in\mathbb{R})^d}$ is a Brownian family on $(\Omega, \mathcal{F}_\infty^W, \mathcal{F}_t^W)$, then it is also a Brownian family on $(\Omega, \widetilde{\mathcal{F}}_\infty, \widetilde{\mathcal{F}}_t)$.*

Now we are ready to proof the Blumenthal 0-1 Law:

Theorem C.8. *Blumenthal Zero-One Law Let $W = \{W_t, \widetilde{\mathcal{F}}_t; t \geq 0\}$,$\{P^x\}_{x\in\mathbb{R}^d}$ a d-dimensional Brownian family on a measurable space $(\Omega, \mathcal{F})$, where $\{\widetilde{\mathcal{F}}_t\}_{t\geq 0}$ is the universal filtration obtained from W_t. If $F \in \widetilde{\mathcal{F}}_0$, then for each $x \in \mathbb{R}^d$ we have either $P^x(F) = 0$ or $P^x(F) = 1$.*

Proof. Let $F \in \widetilde{\mathcal{F}}_0$. For all $x \in \mathbb{R}^d$ there exists $G \in \mathcal{F}_0^W$ such that $P^x(F \triangle G) = 0$. Necessary there exist $\Gamma \in \mathcal{B}(\mathbb{R}^d)$ such that $G = \{W_0 \in \Gamma\}$, so $P^x(G) = \{W_0 \in \Gamma\} = \mathbb{I}_\Gamma(x)$. Since $P^x(F \triangle G) = 0$ imply $P^x(F) = P^x(G)$, we can conclude that $P^x(F) = \mathbb{I}_\Gamma(x)$. $\square$

Appendix D

Viscosity solutions of second order partially differential equations

In this appendix we want to introduce some basics about viscosity solution and exhibit a results on parabolic second order partial differential equations which is useful for our purpose.

Consider a partial differential equation in the form $F(x, u, Du, D^2u) = 0$ where $F : \mathbb{R}^N \times \mathbb{R} \times \mathbb{R}^N \times \mathcal{S}(N) \to \mathbb{R}$ is,unless otherwise said, continuous, and $\mathcal{S}(N)$ is the set of symmetric $N \times N$ matrix. We take u as a real-valued function defined on some subset of $O \subseteq \mathbb{R}^N$. Moreover we require F to satisfy the following two fundamental monotonicity conditions:

$$F(x, r, p, X) \leq F(x, s, p, X) \tag{D.1}$$

whenever $r \leq s$ and

$$F(x, r, p, X) \leq F(x, r, p, Y) \tag{D.2}$$

whenever $Y \leq X$, with $r, s \in \mathbb{R}$, $x, p \in \mathbb{R}^N$, $X, Y \in \mathcal{S}(N)$, and $\mathcal{S}(N)$ equipped with its usual order. The las condition is called "degenerate ellipticity". First of all we need to define viscosity subsolution and supersolution of $F = 0$, then we can give the definition of viscosity solution. Assume that $u : O \subseteq \mathbb{R}$ is a subsolution of $F = 0$ (i.e. $F(x, u, Du, D^2u) \leq 0$)and suppose φ to be a C^2 function and $\widehat{x}$ a local maximum for $u - \varphi$, moreover fix the notations $p = D\varphi(\widehat{x})$, and $X = D^2\varphi(\widehat{x})$.

With these notations it can be shown the following inequality

$$u(x) \leq u(\widehat{x}) + \langle p, x - \widehat{x}\rangle \frac{1}{2}\langle X(x - \widehat{x}), x - \widehat{x}\rangle + o(|x - \widehat{x}|^2) \quad \text{as } x \to \widehat{x} \qquad \text{(D.3)}$$

Fixed u and $\widehat{x}$, we define $J_O^{2,+}u(\widehat{x}) \subseteq \mathbb{R}^N \times \mathcal{S}(N)$ as the set of the couple (p, X) such that verifies inequality D.3 for $O \ni x \to \widehat{x}$. $J_O^{2,+}u(\widehat{x})$ is called the second order "superjet" of u at $\widehat{x}$, and this defines a map $J_O^{2,+}u$ form O to a subset of $\mathbb{R}^N \times \mathcal{S}(N)$. Analogously we can define $J_O^{2,-}u(\widehat{x})$ (second order "subjet" of u at $\widehat{x}$) and the map $J_O^{2,-}u$. Before to define the notions of viscosity subsolution, supersolution and solution we give the following useful notations

$$\text{USC(O)}=\{\text{upper semicontinuous functions } u : O \subseteq \mathbb{R}\}$$
$$\text{LSC(O)}=\{\text{lower semicontinuous functions } u : O \subseteq \mathbb{R}\}$$

Definition D.1. Let F satisfy D.1 and D.2 and $O \subseteq \mathbb{R}^N$. A viscosity subsolution of $F = 0$ on O is a function $u \in USC(O)$ such that

$$F(x, u(x), p, X) \leq 0 \quad \text{for all} \quad x \in O \quad \text{and} \quad (p, X) \in J_O^{2,+}u(x) \qquad \text{(D.4)}$$

Similarly, a viscosity supersolution of $F = 0$ on O is a function $u \in LSC(O)$ such that

$$F(x, u(x), p, X) \geq 0 \quad \text{for all} \quad x \in O \quad \text{and} \quad (p, X) \in J_O^{2,-}u(x) \qquad \text{(D.5)}$$

Finally, u is a viscosity solution of $F = 0$ in O if it is both a viscosity subsolution and a viscosity supersolution of $F = 0$ in O.

Proposition D.2. *Let O be a subset of $\mathbb{R}^N$, $u \in USC(O)$, $v \in LSC(O)$ and*

$$M_\alpha = \sup_{O\times O}(u(x) - v(y) - \frac{\alpha}{2}|x - y|^2)$$

for $\alpha > 0$. Let $M_\alpha < \infty$ for large α and (x_α, y_α) be such that

$$\lim_{\alpha\to\infty}(M_\alpha - (u(x_\alpha) - v(y_\alpha) - \frac{\alpha}{2}|x_\alpha - y_\alpha|^2)) = 0$$

then the following holds:

- $\lim_{\alpha\to\infty} \alpha|x_\alpha - y_\alpha|^2 = 0$

- $\lim_{\alpha\to\infty} M_\alpha = u(\widehat{x} - \widehat{y}) = \sup_O(u(x) - v(x))$ *whenever* $\widehat{x} \in O$ *is a limit point of* x_α *as* $\alpha \to \infty$

Now we to set the second order parabolic problem. In this case the differential equation is in the form

$$u_t + F(t, x, u, Du, D^2u) = 0 \tag{D.6}$$

where $u : O_T \to \mathbb{R}$ is a real-valued function and $O_T = (0, T) \times O$, with $T > 0$ and $O \subseteq \mathbb{R}$ locally compact. We denote by $P_O^{2,+}u(\widehat{x})$ and $P_O^{2,-}u(\widehat{x})$ the parabolic version the semijets $J_O^{2,+}u(\widehat{x})$,$J_O^{2,-}u(\widehat{x})$. In particular they are a subset of $\mathbb{R} \times \mathbb{R}^N \times \mathcal{S}(N)$, where the first component comes from φ_t in a Taylor expansion analog to D.3.

Definition D.3. A viscosity subsolution of D.6 on O_T is a function $u \in USC(O_T)$ such that

$$a + F(t, x, u(t, x), p, X) \leq 0 \quad \text{for all} \quad (t, x) \in O_T \quad \text{and} \quad (a, p, X) \in P_O^{2,+}u(x) \tag{D.7}$$

Similarly, a viscosity supersolution of D.6 on O_T is a function $u \in USC(O_T)$ such that

$$a + F(t, x, u(t, x), p, X) \geq 0 \quad \text{for all} \quad (t, x) \in O_T \quad \text{and} \quad (a, p, X) \in P_O^{2,-}u(x) \tag{D.8}$$

Finally, u is a viscosity solution of D.6 in O_T if it is both a viscosity subsolution and a viscosity supersolution of $F = 0$ in O.

We also consider the case of the Cauchy-Dirichlet problem for the parabolic type, which has the form

$$\begin{cases} \text{(i)}\ , \quad u_t + F(t, x, u, Du, D^2u) = 0 \quad \text{in } (0, T) \times \Omega \\ \text{(ii)}\ , \quad u(t, x) = 0 \quad \text{for } 0 \leq t \leq T \text{ and } x \in \partial\Omega \\ \text{(iii)}, \quad u(0, x) = \psi(x) \text{for } x \in \overline{\Omega} \end{cases} \tag{D.9}$$

where $\Omega \subset \mathbb{R}^N$ is open and $T > 0$ and $\psi(x) \in C(\overline{\Omega})$ are given. by a subsolution of D.9 on $[0, T) \times \overline{\Omega}$, we mean a function $u \in USC([0, T) \times \overline{\Omega})$ such that u is a subsolution of (i),$u(t, x) \leq 0$ for $0 \leq t \leq T$ and $x \in \partial\Omega$ and $u(0, x) \leq \psi(x)$ for

$x \in \overline{\Omega}$. Analogously we can define a supersolution, and we said that u is a solution if it is both a supersolution and a subsolution.

Theorem D.4. *Let O be an open subset of $\mathbb{R}^3$ and $\widetilde{z} = (\widetilde{t}, \widetilde{h}) \in O_T$. If $u : O_T \to \mathbb{R}^3$ and $\varphi \in C^2(O_T,]0, +\infty[)$, then $(a, p, X) \in \overline{P_O}^{2,+} u(\widetilde{z})$(closure of $P_O^{2,+}u(\widetilde{z})$ if and only if*

$$(a\varphi + u\varphi_t, p\varphi + uD_h\varphi, \varphi X + 2p \otimes D_h\varphi + uD_h^2\varphi) \times (\widetilde{z}) \in \overline{P_O}^{2,+} uH(\widetilde{z})$$

where $D_h = (\partial_x, \partial_y)$,$D_h^2$ is the Hessian matrix respect to the spatial variables, and $(p_1, p_2) \otimes (q_1, q_2)$ denotes the matrix

$$\begin{pmatrix} p_1q_1 & \frac{p_1q_2+p_2q_1}{2} \\ \frac{p_1q_2+p_2q_1}{2} & p_2q_2 \end{pmatrix}$$

Remark D.5. An analogous statement holds if $\overline{P_O}^{2,+}$ is replaced by $\overline{P_O}^{2,-}$

Theorem D.6. *Let $u_i \in UCS((0,T)\times \ =_i)$ for $i = 1, \ldots, k$, where O_i is a locally compact subset of $\mathbb{R}^{N_i}$. Let φ be defined on an open neighborhood of $(0, T) \times O_1\times, \cdots, \times O_k$ and such that $(t, x_1, \ldots, x_k) \to \varphi(t, x_1, \ldots, x_k)$ is once continuously differentiable in t and twice continuously differentiable in $(x_1, \ldots, x_k) \in O_1\times, \cdots, \times O_k$. Suppose that $\widehat{t} \in (0, T)$, $\widehat{x}_i \in O_i$, for $i = 1, \ldots, k$ and*

$$\begin{aligned} w(t, x_1, \ldots, x_k) = u_1(t, x_1) + \cdots u_k(t, x_k) - \varphi(t, x_1, \ldots, x_k) \leq \\ \leq w(\widehat{t}, \widehat{x_1}, \ldots, \widehat{x_k}) \end{aligned}$$

for $0 < t < T$ and $x_i \in O$. Assume, moreover, that there is an $r > 0$ such that for every $M > 0$ there is a C such that for $i = 1, \ldots, k$

$$\begin{aligned} &b_i \leq C \text{ whenever } (b_i, q_i, X_i) \in \overline{P_O}^{2,+} u_i(t, x_i), \\ &|x_i - \widehat{x}_i| + |t - \widehat{t}| \leq r \text{ and } |u_i(t, x_i)| + |q_i| + \|X_i\| \leq M. \end{aligned}$$

Then for each $\varepsilon > 0$ there are $X_i \in \mathcal{S}(N_i)$ such that

$$\begin{cases} (b_i, D_{x_i}\varphi(\widehat{t}, \widehat{x}_1, \ldots, \widehat{x}_k)) \in \overline{P_O}^{2,+} u_i(\widehat{t}, \widehat{x}_i) \text{ for } i = 1, \ldots, k, \\ -\left(\frac{1}{\varepsilon} + \|A\|\right) I \leq \begin{pmatrix} X_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & X_k \end{pmatrix} \leq A + \varepsilon A^2, \\ b_1 + \cdots + b_k = \varphi_t(\widehat{t}, \widehat{x}_1, \ldots, \widehat{x}_k). \end{cases} \tag{D.10}$$

where $A = (D_x^2\varphi)(\widehat{t}, \widehat{x}_1, \ldots, \widehat{x}_k)$.

The next theorem gives a comparison principle for viscosity solution

Theorem D.7. *Let*$\Omega \subset \mathbb{R}^N$ *be open and bounded. Let* $F \in C([0,T] \times \overline{\Omega}) \times \mathbb{R} \times \mathbb{R}^N \times \mathcal{S}(N)$ *be continuous, proper and satisfy*

$$\begin{cases} F(y, r, a(x-y), Y) - F(x, r, a(x-y), X) \leq \omega(\alpha|x-y|^2 + |x-y|) \\ \text{whenever } x, y \in \Omega, \quad r \in \mathbb{R}, \quad X, Y \in \mathcal{S}(N), \text{ and for each } \varepsilon > 0 \\ -\frac{3}{\varepsilon}\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \leq \begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \leq \frac{3}{\varepsilon}\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \end{cases}$$

for each fixed $t \in [0, T[$*, with the same function* ω*. If* u *is a subsolution of D.9 and* v *is a supersolution of D.9, then* $u \leq v$ *on* $[0, T[\times\Omega$*.*

For the proof of these theorems and a complete treatment of this topic see [10].

Bibliography

[1] Pascucci Andrea. *PDE and Martingale Methods in Option Pricing*. Milano: Bocconi University Press, Springer, 2001.

[2] Friedman Avner. *Partial Differential Equations of Parabolic Type*. Prentice-Hall, 1983.

[3] Friedman Avner. *Stochastic Differential Equations and Applications, Volume 1*. Academic Press, 1975.

[4] Brezis Haim. *Functional Analysis,Sobolev Spaces and Partial Differential Equations*. Springer, 2011.

[5] Karatzas Iannis and Shreve Steven E. *Brownian Motion and Stochastic calculus*. Springer-Verlag, 1991.

[6] Ma Jin and Yong Jiongmin. *Forward-Backward Stochastic Differential Equations and Their Applications*. Springer-Verlag, 1999.

[7] Forman Julie L. and Srensen Michael. *The Pearson Diffusion: A class of statistically tractable diffusion processes.* Scandivian Journal of Statistics, 2008.

[8] Hörmander Lars. *Hypoelliptic second order differential equations*. Vol. 119. p. 147-171. Milano: Acta Mathematica, 1967.

[9] Tu Loring W. *An Introduction to Manifolds*. Springer, 2008.

[10] Crandall M.G., Ishii H., and Lions P.L. *User's guide to viscosity solutions of second order partial differential equations.* 27 (1) p. 1-67. Bull. Amer. Math. Soc., 1992.

[11] V.A. Solonnikov O.A. Ladyženskaja and Ural'ceva N.N. *Linear and Quasilinear Equations of Parabolic Type*. American Mathematical Society.

[12] Radkevich Olga A Oleinik ans Evgenii V. *Second order differential Equations wiht nonegative Characteristic Form*. American Mathematical Society, 1973.

[13] Protter Philip E. *Stochastic Inegration and Differnetial Equations*. Springer.

[14] LeVeque Randall J. *Numerical methods for conservation laws*. Birkhäuser, 1900.

[15] Carmona René and Dealrue François. *Singular FBDEs and scalr conservation laws driven by diffusion processes*.