Lisbon School of Economics & Management
Universidade de Lisboa

# MASTER

MATHEMATICAL FINANCE

## MASTER'S FINAL WORK

DISSERTATION

# DISCRETE-TIME APPROXIMATION OF A SINGULAR FORWARD-BACKWARD STOCHASTIC DIFFERENTIAL EQUATION AND APPLICATION TO THE CARBON MARKET

MIGUEL ÂNGELO PARREIRA DÂMASO

JUNE 2026

Lisbon School of Economics & Management
Universidade de Lisboa

# MASTER

MATHEMATICAL FINANCE

## MASTER'S FINAL WORK

DISSERTATION

# DISCRETE-TIME APPROXIMATION OF A SINGULAR FORWARD-BACKWARD STOCHASTIC DIFFERENTIAL EQUATION AND APPLICATION TO THE CARBON MARKET

MIGUEL ÂNGELO PARREIRA DÂMASO

**SUPERVISION:**

JOÃO MIGUEL ESPIGUINHA GUERRA

JUNE 2026

# Abstract

In this dissertation, we analyse a discrete time approximation of singular Forward Backward Stochastic Differential Equation. This approximation is based on a splitting scheme in which the transport operator and the diffusion operator are split. We present a practical algorithm to implement this scheme and apply this algorithm to a carbon emissions market model, which incorporates the dynamics of the emission certificates price, accumulated emissions, fuel prices, and electricity demand.

Keywords: Forward Backward Stochastic Differential Equation; Carbon Market; Splitting Scheme; Numerical Algorithms

# Acknowledgments

I am sincerely grateful to Professor João Guerra for his constant guidance and dedication throughout these past months. Whenever I required assistance, he was always available to provide useful advice, and I never felt alone during this process.

To my family and friends, thank you for your love and encouragement. I will never forget the constant companionship and support. Without your care, this journey would have been considerably more challenging.

# Contents

# List of Figures

# List of Tables

# Glossary

**BAU** business as usual. vi, 8, 25

**CDF** Cumulative Distribution Function. iv, vi, 14–16, 20, 27, 28, 30, 33

**EU ETS** European Union Emissions Trading System. vi, 1, 5

**EUA** European Union Allowance. vi, 5

**FBSDE** Forward Backward Stochastic Differential Equation. vi, 1–6, 10, 12, 23, 33

**GBP** Great Britain Pound (Pound Sterling). vi, 23

**MWh** megawatt-hour. vi, 23, 25

**PDE** Partial Differential Equation. vi, 1, 4, 12

**SDE** Stochastic Differential Equation. vi, 6, 13

**SPD** Sticky Particle Dynamics. vi, 14

**UK** United Kingdom. vi, 23

# 1 Introduction

The industrial revolution represents one of the defining events in modern history. The global economy experienced significant changes, and the world population began to grow rapidly. Hence, the emission of greenhouse gases, especially carbon dioxide, increased to undesirable levels, negatively impacting all forms of life on our planet. The release of greenhouse gases naturally contributes to climate change. Climate change is a well established phenomenon, and temperatures are expected to increase further in the coming years, according to scientists. Thus, it becomes clear that effective measures must be taken to substantially reduce its adverse impacts.

This concern has led governments to discuss strategies to mitigate this issue. Profound adjustments are necessary to ensure that the concentration of greenhouse gases in the atmosphere decreases to a level at which the consequences of climate change are no longer excessively severe. Despite the urgency of addressing this problem, reaching an agreement remains a considerable challenge. However, following the Kyoto Protocol, an accord dating from the end of the twentieth century and aimed at mitigating emissions, the European Union adopted a proactive climate strategy and launched the European Union Emissions Trading System (EU ETS), its carbon emissions market. In recent years, carbon markets have experienced significant growth across the world, and, for instance, China, which is the largest emitter of greenhouse gases, has already established such a market.

In our thesis, we analyse a particular discrete time approximation that is applied to a carbon market like EU ETS. This market model is formulated using a singular Forward Backward Stochastic Differential Equation (FBSDE). This type of equation requires careful examination because the terminal condition is discontinuous and a degenerate condition may occur. Furthermore, since the singular FBSDE is fully coupled, the numerical method described in [1] and [2] is not applicable. The numerical approximation method considered in this work, which is suitable for this class of equations, was presented in [11] and [12]. It relies on the splitting of two operators, the transport operator and the diffusion operator, associated with singular FBSDE. This splitting scheme allows to surpass the difficulty of classical FBSDEs methods to capture the solution of the associated quasi-linear Partial Differential Equation (PDE) that involves a diffusion part and a transport part, precisely by considering different numerical approximation methods for the diffusion part and for the non-linear transport part. We will discuss two numerical approximation cases, Case 1 and Case 2. Focusing on the latter, we present an algorithm for practical implementation. Finally, we implement the numerical approximation for a linear

example originally discussed in [8], as well as for a carbon emissions market model with two fuels. As far as we are aware, the discrete approximation method presented in this work has not previously been applied to the carbon emissions market model.

The structure of the dissertation is as follows. In the second chapter, we provide an overview of FBSDEs and discuss the specific carbon emissions market model based on singular FBSDEs and the associated mechanisms. In the third chapter, we define the elements of the splitting scheme, specify two different settings depending on the dynamics of a given process, introduce two distinct numerical approximation cases, and present the decomposition of the error. In the fourth chapter, we develop two numerical algorithms which are implemented to approximate the solutions of two different applications considered in some bibliographic references. Moreover, the numerical results we obtain are shown. In the last chapter, we summarise the main findings of our work and identify further topics for investigation.

# 2 Theoretical Background

## 2.1 Forward Backward Stochastic Differential Equation

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space along with a filtration $(\mathcal{F}_t)_{0\leq t\leq T}$ generated by the $d$-dimensional Brownian motion $W$ (augmented). Let $T > 0$ be the terminal time.

A FBSDE is given by

$$\begin{cases} X_t = X_0 + \int_0^t b\left(s, X_s, Y_s, Z_s\right)\, ds + \int_0^t \sigma\left(s, X_s, Y_s\right)\, dW_s \\ Y_t = g\left(X_T\right) + \int_t^T f\left(s, X_s, Y_s, Z_s\right)\, ds - \int_t^T Z_s\, dW_s, \end{cases} \tag{1}$$

where the functions $b, \sigma, f$ and $g$ are Lipschitz continuous.

The aim is to identify a triple of adapted processes $(X_t, Y_t, Z_t)$ that solves the system, where $X$, $Y$ and $Z$ correspond to the forward, backward, and control components, respectively. The role of $Z$ is essential for obtaining a non-anticipative solution ([1]). In contrast to the forward equation, the solution to the backward, given by $(Y, Z)$, is not specified at the initial time 0, but rather at the terminal time $T$.

FBSDEs have been extensively employed in several fields, such as stochastic control and mathematical finance (see, for example, [14]). They were presented by Bismut [3]. Pardoux and Peng [18] later developed a general study for these equations.

We now introduce the singular FBSDE, representing the most basic non-trivial case of an FBSDE with singular coefficients.

The singular FBSDE is a system with the particular form (in differential notation):

$$\begin{cases} dP_t = b(P_t)\, dt \, + \, \sigma(P_t)\, dW_t \\ dE_t = \mu(Y_t, P_t)\, dt \\ dY_t = Z_t\, dW_t, \end{cases} \tag{2}$$

with solution $(P_t, E_t, Y_t, Z_t)_{0\leq t\leq T}$. $b : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \to \mathcal{M}_d$, where $\mathcal{M}_d$ is the set of $d \times d$ matrices on $\mathbb{R}$, and $\mu : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ are Lipschitz continuous functions. Note that in the particular case of a singular FBSDE, the forward process is written as $X_t = (P_t, E_t)$.

In [8], these equations have been proposed for the modelling of the carbon market. More broadly, they can be used to describe the cap-and-trade scheme implemented by a regulatory authority to restrict the release of a specific pollutant. Within this framework, $Y$ denotes the allowance price, while $E$ represents the accumulated emissions. $P$ consists of a set of state variables affecting emissions, such as demand or energy prices.

In a singular FBSDE, the terminal condition is discontinuous and is typically given by

$$Y_T = \phi(E_T), \tag{3}$$

where $\phi(e) = \mathbf{1}_{\{e \geq \Lambda\}}$, for some $\Lambda > 0$.

The singular FBSDE is fully coupled (the backward process $Y$ appears in the coefficient of the forward component $E$), degenerate in the forward process $E$ (its diffusion term is equal to zero), and the terminal condition is discontinuous (these are the main reasons why the FBSDE is called singular). Hence, establishing the existence and uniqueness of the above singular FBSDE constitutes a complex task.

However, taking into account some structural conditions, Carmona and Delarue in [7] managed to prove the wellposedness of such singular FBSDE.

First, the drift $\mu$ must be strictly increasing and additionally fulfils

$$l_1|y - y'|^2 \leq (y - y')[\mu(y, p) - \mu(y', p)] \leq l_2|y - y'|^2, \tag{4}$$

for constants $l_1, l_2 > 0$ and $y, y' \in \mathbb{R}^2$. Furthermore, at $t = T$, the authors show that $Y_T$ is not determined by $E_T$. As a consequence, the terminal condition can only be imposed in

$$\mathbf{1}_{(\Lambda,\infty)}(E_T) \leq Y_T \leq \mathbf{1}_{[\Lambda,\infty)}(E_T). \tag{5}$$

They also show the classical markovian representation of $Y$, namely

$$Y_t = \mathcal{V}(t, P_t, E_t), \tag{6}$$

for $0 \leq t < T$, where the measurable function $\mathcal{V} : [0, T] \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ is commonly referred to as the decoupling field.

In the case of the singular FBSDE, the decoupling field $\mathcal{V}$ solves the PDE

$$\partial_t u + \partial_e(\mathfrak{M}(p, u)) + \partial_p u b(p) + \frac{1}{2}\mathrm{Tr}\left[\sigma(p)\sigma^{\mathrm{T}}(p)\partial^2_{pp} u\right] = 0, \tag{7}$$

whose terminal condition is $u(T,p,e)=\phi(e)$. $\partial_t$ and $\partial_e$ are derivatives with respect to time and variable $e$, respectively. $\partial_p$ represents the Jacobian relative to $p$ and $\partial^2_{pp}$ the matrix of second derivatives in $p$. Tr represents the matrix trace, T denotes the transpose and the function $\mathfrak{M}(p,u)$ is given by

$$\mathfrak{M}(p,u)=\int_0^u \mu(p,v)dv, \tag{8}$$

with $u\in[0,1]$.

The numerical method proposed in [1] and [2] cannot be used within the framework of singular FBSDEs. This is only feasible if we approximate the discontinuous terminal condition by a continuous one and a diffusion coefficient is incorporated into the process of $E$. However, this is not the simplest approach, and a more tractable method should be considered.

## 2.2 Carbon market

Electricity generation from any combination of fuels produces emissions, so alternative electricity production methods, namely nuclear and renewable, are not considered in the EU ETS.

Each tonne of carbon dioxide released must be matched by the delivery of one European Union Allowance (EUA). The market operates under an annual cap on total emissions, denoted $E^{cap}$, and this restriction on the number of allowance certificates gives them monetary value. These certificates may be auctioned, sold, or simply distributed. Upon the conclusion of the compliance year, electricity producers are required to submit certificates equal to their total emissions. Firms with surplus certificates have two options. They may sell these additional certificates to other producers or retain them for use in the following year. Those with insufficient certificates must pay a fine of $\Lambda$ euros.

The trading of EUAs between producers constitutes the carbon market. Under perfect information, a carbon emissions tax has, economically, the same effect as the carbon market (see [10]). The market includes both suppliers and consumers, and for simplicity, producers are assumed to act as suppliers. We assume the existence of a market administrator who aligns supply and demand. Therefore, consumers have no choice regarding their energy suppliers. The market administrator receive the bids from suppliers (which contains a quantity of electricity together with the price demanded for that quantity), and ranks them in merit order, meaning demand

is allocated to the cheapest supply offer available.

We assume perfect inelastic electricity demand, which means that electricity prices do not influence consumption levels. Electricity production is subject to a capacity constraint. When demand exceeded supply, only the quantity produced can be delivered. Consequently, demand is bounded in a similar way. Subsequently, market equilibrium is achieved when demand equals supply, at any time.

As a result of this equilibrium, the market bid stack can be defined, which characterises the electricity price as a function of fuel prices, emission allowance prices, and supplied quantity of electricity.

In the following, we introduce the mathematical model associated with the carbon market. The assumption below facilitates the modelling of emission allowance price dynamics.

**Assumption 2.1.** *The measure $\mathbb{P}$ is defined as the objective probability measure and admits an equivalent measure $\mathbb{Q}$, known as the risk neutral measure. With respect to $\mathbb{Q}$, the discounted price process of any tradable asset follows a martingale process.*

We now describe the standard structural model for carbon emissions using FBSDEs, as presented in [16] and [10].

Consider the probability space $(\Omega, \mathcal{F}, \mathbb{Q})$. Let $(\mathcal{F}_t)_{t\in[0,T]}$ be the augmented filtration generated by the Brownian motion $(W_t)_{t\in[0,T]}$ with dimension $n+1$.

We define $\xi_{max}$ as the maximum quantity of electricity that can be generated. As mentioned above, we have, for $0 \leq t \leq T$,

$$0 \leq D_t \leq \xi_{max}, \tag{9}$$

and

$$0 \leq \xi_t \leq \xi_{max}, \tag{10}$$

where $D_t$ and $\xi_t$ represent electricity demand and electricity supply at time $t$, respectively. For $t \in [0, T]$, the equilibrium is given by

$$D_t = \xi_t. \tag{11}$$

We derive the dynamics of the process $D_t$, described by the Stochastic Differential Equation (SDE)

$$dD_t = \mu_d(t, D_t)dt + \sigma_d(D_t)dW_t^0, \tag{12}$$

where $(W_t^0)_{t\in[0,T]}$ denotes the first element of $(W_t)_{t\in[0,T]}$. The initial condition is $D_0 = d_0$, with $d_0 \in [0, \xi_{max}]$.

Fuel prices are represented by the process $S_t = (S_t^1, \ldots, S_t^n)$, where $n$ is the number of fuels used in the generation of electricity. $S_t$ is characterised by

$$dS_t = \mu_s(S_t)dt + \sigma_s(S_t)dW_t^n, \tag{13}$$

with initial condition $S_0 = s_0 \in (0, \infty)^n$. $(W_t^n)_{t\in[0,T]}$ represents the last $n$ components of $(W_t)_{t\in[0,T]}$.

We specify $E_t$ as the accumulated emissions until time $t$. Let $E_0 = 0$ and the process evolves as

$$dE_t = \mu_E(D_t, S_t, A_t)dt, \tag{14}$$

where $\mu_E(D_t, S_t, A_t)$ denotes the market emissions rate. Since electricity production is bounded, we introduce $E_{max}$ as the maximum amount of emissions that can be released. Combining this with the previously defined $E^{cap} > 0$, we obtain $0 \leq E_t \leq E^{cap} \leq E_{max}$.

We now present $A_t$, representing the price of the certificate. As stated above, at the final time $T$,

$$A_T = \Lambda \mathbf{1}_{[E^{cap},\infty)}(E_T). \tag{15}$$

The emission allowances are tradable assets. By Assumption 2.1, for $t \in [0, T]$, the discounted price process $e^{-rt}A_t$ is a martingale with respect to the filtration $(\mathcal{F}_t)_{t\in[0,T]}$, with $r$ being the risk free rate. According to the martingale representation theorem, there exists an adapted process $Z_t$ satisfying

$$d[e^{-rt}A_t] = Z_t dW_t, \tag{16}$$

and, applying the Itô formula, we get

$$dA_t = rA_t dt + Z_t dW_t. \tag{17}$$

Consider $I = \{1, \ldots, n\}$ the set of fuels. $\xi_{max}^i > 0$ corresponds to the maximum electricity output that can be produced relying exclusively on the fuel type $i \in I$. Hence, it must hold

$$\xi_{max} = \sum_{i=1}^{n} \xi_{max}^i. \tag{18}$$

We now turn to the definition of the bid rate for each fuel, which is the concept analogous to the bid stack of a market where electricity is generated from a single fuel. First, we analyse the case without any restrictions on total emissions, corresponding to a business as usual (BAU) framework.

**Definition 2.1.** *For fuel $i \in I$,*

$$b_i^{BAU} : [0, \xi_{max}^i] \times \mathbb{R} \to \mathbb{R}_0^+ \tag{19}$$

*is the BAU bid rate and the function $s_i \mapsto b_i^{BAU}(\xi, s_i)$ is strictly increasing, fixing $\xi \in [0, \xi_{max}^i]$.*

The impact of emissions restrictions is reflected in the different levels of emissions released by each fuel.

**Definition 2.2.** *For fuel $i \in I$,*

$$e_i : [0, \xi_{max}^i] \to \mathbb{R}^+ \tag{20}$$

*denotes the marginal emissions stack when electricity is produced using only fuel $i$. Let $e_i$ be a non decreasing function.*

In the presence of emissions restrictions, allowance certificates constitute an additional cost of electricity generation, owing to the similar effects of the carbon markets and the carbon taxes under perfect information. From this, we can formulate the definition of the bid rate in this case.

**Definition 2.3.** *For fuel $i \in I$,*

$$b_i : [0, \xi_{max}^i] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}_0^+ \tag{21}$$

*is the bid rate given by*

$$b_i(\xi, a, s_i) = b_i^{BAU}(\xi, s_i) + a e_i(\xi). \tag{22}$$

**Definition 2.4.** *For fuel $i \in I$, electricity price $p$, certificate price $a$ and fuel prices $s = (s_1, \ldots, s_n)$,*

$$C_i(p, a, s) = \{\xi \in [0, \xi_{max}^i] : b_i(\xi, a, s_i) \leq p\} \tag{23}$$

*is the collection of generators of fuel type $i$ that may be active.*

In other words, $C_i(p, a, s)$ corresponds to the range of production quantities that yield a profit.

Since $b_i^{BAU}$ is strictly increasing, the map $\xi \mapsto b_i(\xi, a, s_i)$ is also strictly increasing, with $a$ and $s_i$ held fixed. The maximum and the minimum are represented by

$$\begin{aligned} \overline{b_i}(a, s_i) &= b_i(\xi^i_{max}, a, s_i), \\ \underline{b_i}(a, s_i) &= b_i(0, a, s_i). \end{aligned} \tag{24}$$

We therefore define the generalised right continuous inverse function as

$$\begin{aligned} b_i^{-1}(p, a, s_i) : \mathbb{R} &\to [0, \xi^i_{max}] \\ p &\mapsto \xi^i_{max} \wedge \inf\{\xi \in [0, \xi^i_{max}] : b_i(\xi, a, s_i) > p\}, \end{aligned} \tag{25}$$

interpreted as the quantity of electricity generated solely from fuel $i$.

Consequently, the electricity produced from all fuel types combined is

$$\begin{aligned} b^{-1} &: \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n \to [0, \xi_{max}] \\ (p, a, s) &\mapsto \sum_{i \in I} b_i^{-1}(p, a, s_i). \end{aligned} \tag{26}$$

Fixing $a$ and $s$, the mapping $p \mapsto b^{-1}(p, a, s)$ is strictly increasing. The market bid stack is thus defined as the generalised left continuous inverse function expressed by

$$b(\xi, a, s) = \min_{i \in I} \underline{b_i}(a, s_i) \vee \sup \left\{ p \in \mathbb{R} : \sum_{i \in I} b_i^{-1}(p, a, s_i) < \xi \right\}. \tag{27}$$

The market emissions rate with capped emissions is characterised by

$$\mu_e(\xi, a, s) = \sum_{i \in I} \left( \int_{C_i(b(\xi, a, s), a, s)} e_i(x) dx \right). \tag{28}$$

Rewriting the expression, we obtain

$$\mu_e(\xi, a, s) = \sum_{i \in I} \left( \int_0^{b_i^{-1}(b(\xi, a, s), a, s)} e_i(x) dx \right), \tag{29}$$

and the cumulative emissions $E_t$ at time $t$ are defined by

$$E_t = \int_0^t \mu_E(D_s, S_s, A_s) ds. \tag{30}$$

# 3 Splitting scheme

The model considered in this section relies on a splitting scheme involving a transport operator and a diffusion operator. On a discrete time grid, the scheme is defined through the iterative composition of these operators.

We first present a class of functions that includes the solution. Let $\mathcal{P}(\mathbb{R})$ be the space of probability measures on $\mathbb{R}$.

**Definition 3.1.** *Let $\mathcal{K}$ denote the class of functions $\Phi : \mathbb{R}^d \times \mathbb{R} \to [0,1]$ such that $\Phi$ satisfies the $L_\Phi$-Lipschitz condition in its first argument for some $L_\Phi > 0$, that is,*

$$|\Phi(p,e) - \Phi(p',e)| \leq L_\Phi |p - p'|, \tag{31}$$

*for all $(p,p',e) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}$, and further verifies, for all $p \in \mathbb{R}^d$,*

$$\Phi(p,e) = \nu(p,(-\infty,e]), \tag{32}$$

*where $\nu(p,\cdot) \in \mathcal{P}(\mathbb{R})$.*

Next, a class of admissible coefficient functions is introduced, which is crucial for establishing the wellposedness of the singular FBSDE and for describing the splitting scheme.

**Definition 3.2.** *We denote by $\mathcal{A}$ the class of L-Lipschitz continuous functions $B : \mathbb{R}^d \to \mathbb{R}^d$, $\Sigma : \mathbb{R}^d \to \mathcal{M}^d$ and $F : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$. $F$ is strictly decreasing in its second argument, and*

$$l_1 |y - y'|^2 \leq (y - y')(F(p,y') - F(p,y)) \leq l_2 |y - y'|^2 \tag{33}$$

*holds for all $p \in \mathbb{R}^d$, for some constants $L, l_1, l_2 > 0$.*

In this framework, the existence and uniqueness theorem for the singular FBSDE is stated below. First, we introduce some notation related to the theorem. $\mathbb{H} = (\mathcal{H}_t)_{t \geq 0}$ represents an arbitrary right continuous filtration and $\mathcal{S}^{2,k}(I,\mathbb{H})$ denotes the space of $\mathbb{R}^k$-valued càdlàg $\mathcal{H}_t$ adapted processes $Y$, for which

$$\|Y\|_{\mathcal{S}^2}^2 = \mathbb{E}\left[\sup_{t \in I} |Y_t|^2\right] < \infty, \tag{34}$$

where $I = [a,b]$ or $I = [a,b)$, with $0 \leq a < b < +\infty$ fixed. The subspace of processes that have continuous trajectories is referred to $\mathcal{S}_c^{2,k}(I,\mathbb{H})$.

Lastly, $\mathcal{H}^{2,k}(I,\mathbb{H})$ is the set of progressively measurable processes $Z$ with values in $\mathbb{R}^k$ that satisfies

$$\|Z\|^2_{\mathcal{H}^2} = \mathbb{E}\left[\int_I |Z_t|^2 dt\right] < \infty, \tag{35}$$

given $I = [a,b]$, with $0 \leq a < b < +\infty$ fixed.

We are now prepared to state the theorem.

**Theorem 3.1.** *(Proposition 2.10 in [7], Proposition 3.2 in [9]) Suppose that $\tau > 0$, $(B,\Sigma,F) \in \mathcal{A}$ and $\Phi \in \mathcal{K}$.*

*For an arbitrary initial condition $(t_0,p,e) \in [0,\tau)\times\mathbb{R}^d\times\mathbb{R}$, a single progressively measurable 4-tuple of processes $(P_t^{t_0,p,e}, E_t^{t_0,p,e}, Y_t^{t_0,p,e}, Z_t^{t_0,p,e})_{t_0\leq t\leq\tau} \in \mathcal{S}_c^{2,d}([t_0,\tau],\mathbb{F})\times \mathcal{S}_c^{2,1}([t_0,\tau],\mathbb{F})\times\mathcal{S}_c^{2,1}([t_0,\tau],\mathbb{F})\times\mathcal{H}^{2,d}([t_0,\tau],\mathbb{F})$ exists and fulfils the dynamics*

$$\begin{aligned} dP_t^{t_0,p,e} &= B(P_t^{t_0,p,e})dt + \Sigma(P_t^{t_0,p,e})dW_t, \\ dE_t^{t_0,p,e} &= F(P_t^{t_0,p,e}, Y_t^{t_0,p,e})dt, \\ dY_t^{t_0,p,e} &= Z_t^{t_0,p,e}dW_t, \end{aligned} \tag{36}$$

*with $P_t^{t_0,p,e} = p \in \mathbb{R}^d$ and $E_t^{t_0,p,e} = e \in \mathbb{R}$. Furthermore,*

$$\mathbb{P}\left[\Phi_-\left(P_\tau^{t_0,p,e}, E_\tau^{t_0,p,e}\right) \leq \lim_{t\uparrow\tau} Y_t^{t_0,p,e} \leq \Phi_+\left(P_\tau^{t_0,p,e}, E_\tau^{t_0,p,e}\right)\right] = 1. \tag{37}$$

*For $(t_0,p,e) \in [0,\tau)\times\mathbb{R}^d\times\mathbb{R}$,*

$$w(t_0,p,e) = Y_{t_0}^{t_0,p,e} \in \mathbb{R} \tag{38}$$

*is the unique decoupling field and is continuous.*

We now describe the theoretical splitting scheme. Initially, we examine the transport step in which the diffusion term is kept constant.

**Definition 3.3.** *Define*

$$\mathcal{T}_h(\psi) = \tilde{v}(0,\cdot) \in \mathcal{K}, \tag{39}$$

*for $(h,\psi) \in (0,\infty)\times\mathcal{K}$, where $\tilde{v}$ denotes the decoupling field introduced in Theorem 3.1 with terminal condition $\Phi = \psi$ and coefficients given by $\tau = h$, $B = 0$, $\Sigma = 0$ and $F = \mu$.*

$\tilde{v}(\cdot)$, whose domain is $[0,h]\times\mathbb{R}^d\times\mathbb{R}$, is the unique entropy solution, see [6] or [15], to

$$\partial_t w + \partial_e(\mathfrak{M}(p,w)) = 0, \tag{40}$$

with $\tilde{v}(h,\cdot) = \psi$.

The entropy solution is a weak solution concept for the degenerate quasilinear PDE associated with the singular FBSDE.

We proceed by introducing the diffusion step, in which the transport component is kept fixed.

**Definition 3.4.** *Define*

$$\mathcal{D}_h(\psi) = \bar{v}(0,\cdot) \in \mathcal{K}, \tag{41}$$

*for* $(h,\psi) \in (0,\infty) \times \mathcal{K}$*, where* $\bar{v}$ *denotes the decoupling field introduced in Theorem 3.1 with terminal condition* $\Phi = \psi$ *and coefficients given by* $\tau = h$*,* $B = b$*,* $\Sigma = \sigma$ *and* $F = 0$*.*

Now, for $N \geq 1$, it is possible to specify the theoretical scheme on the time grid $\pi$, where

$$\pi = \{t_0 = 0 \leq t_1 \leq \cdots \leq t_n \leq \cdots \leq t_N = T\}. \tag{42}$$

We suppose that the time discretization is equidistant, so

$$t_{n+1} - t_n = \frac{T}{N} = h, \tag{43}$$

for $0 \leq n \leq N-1$.

**Definition 3.5.** *Define, for* $(h,\psi) \in (0,\infty) \times \mathcal{K}$*,*

$$\mathcal{S}_h(\psi) = \mathcal{T}_h \circ \mathcal{D}_h(\psi) \in \mathcal{K}. \tag{44}$$

*For* $n \leq N$*, the backward induction on* $\pi$ *described below admits a solution* $u_n^\pi$*:*
*- for* $n = N$*,* $u_N^\pi = \phi$*,*
*- for* $n < N$*,* $u_n^\pi = \mathcal{S}_{t_{n+1}-t_n}\left(u_{n+1}^\pi\right)$*.*

For $t \in \pi$, the decoupling field $\mathcal{V}(t,\cdot)$ is approximated by $(u_n^\pi)_{0\leq n\leq N}$.

We proceed by stating an assumption on the value function in order to allow us to work within a stricter setting. For $I = [0,T)$ or $I = [0,T]$, $\mathscr{C}^{1,2,1}(I \times \mathbb{R}^d \times \mathbb{R})$ denotes the class of functions $f : I \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ which are continuously differentiable in the first and third variables and twice continuously differentiable in the second variable.

**Assumption 3.1.** *The decoupling field* $\mathcal{V} \in \mathscr{C}^{1,2,1}([0,T) \times \mathbb{R}^d \times \mathbb{R})$ *solves*

$$\partial_t \mathcal{V} + \mu(p,\mathcal{V})\partial_e \mathcal{V} + b(p)\partial_p \mathcal{V} + \frac{1}{2}\bar{\sigma}^2 \Delta_p \mathcal{V} = 0. \tag{45}$$

*Recalling* (7), $\sigma(\cdot)$ *is assumed to be a constant function, given by* $\sigma(\cdot) = \bar{\sigma} I_d$, *where* $\bar{\sigma} > 0$ *and* $I_d$ *stands for the* $d \times d$ *identity matrix.* $\Delta_p$ *denotes the Laplacian operator with respect to the variable* $p$. *Additionally, the functions* $\mu$ *and* $b$ *are differentiable (respectively, twice differentiable) and their derivatives are bounded and Lipschitz.*

The approximation of the Wiener measure through a finite collection of discrete paths is the basis of all schemes that we use. In fact, on the time grid $\pi$, we approximate the Brownian motion using a discrete time process $\hat{W}$. $(\Delta\hat{W}_n = \hat{W}_{t_{n+1}} - \hat{W}_{t_n})_{0 \leq n \leq N-1}$ is the discrete approximation of the Brownian increments $(W_{t_{n+1}} - W_{t_n})_{0 \leq n \leq N-1}$.

We write $\mathfrak{S}_n^d \subset \mathbb{R}^d$ for the support of the distribution of $\hat{W}_{t_n}$, at each time $t_n$.

**Definition 3.6.** *The cubature formula we consider employs precisely* $2d$ *points.*

*The canonical basis of* $\mathbb{R}^d$ *is* $\left(e^l\right)_{1 \leq l \leq d}$ *and we define, for* $1 \leq i \leq I = 2d$, $\mathbb{P}\left(\Delta\hat{W}_n = \omega_h^i\right) = \dfrac{1}{2d}$ *and* $\omega_h^i = -\sqrt{dh}e^l$, *if* $i = 2l$ *or* $\omega_h^i = \sqrt{dh}e^l$, *if* $i = 2l - 1$.

## 3.1 General diffusion schemes

We want to explore the generic case where $P$ solves an SDE with Lipschitz coefficients. On the grid $\pi$, the discrete approximation is obtained by the Euler method with $\hat{P}_0 = P_0$ and

$$\hat{P}_{t_{n+1}} = \hat{P}_{t_n} + b\left(\hat{P}_{t_n}\right) h + \sigma\left(\hat{P}_{t_n}\right) \Delta\hat{W}_n. \tag{46}$$

Also, for $p \in \mathbb{R}^d$ and $0 \leq n \leq N-1$,

$$\hat{P}_{t_{n+1}}^{t_n,p} = p + b(p)h + \sigma(p)\Delta\hat{W}_n, \tag{47}$$

and, for $1 \leq i \leq 2d$,

$$\left(\hat{P}_{t_{n+1}}^{t_n,p}\right)^i = p + b(p)h + \sigma(p)\omega_h^i. \tag{48}$$

For $0 \leq n \leq N-1$, $\mathcal{P}_n$ is the discrete finite support of $\hat{P}_{t_n}$. Thanks to the formulation of $\left(\hat{P}_{t_n}\right)$, for $p \in \mathcal{P}_n$,

$$\hat{P}_{t_{n+1}}^{t_n,p} \in \mathcal{P}_{n+1}. \tag{49}$$

Firstly, we define a discrete version of the operator $\mathcal{T}$ that will construct a forward approximation of $\partial_t w + \partial_e(\mathfrak{M}(p, w)) = 0$ and $\tilde{v}(h, \cdot) = \psi$. For that purpose,

we employ Sticky Particle Dynamics (SPD) (see [5] and Section 1.1 of [17]).

**Definition 3.7.** *Consider the polyhedron $D_n \subset \mathbb{R}^n$ given by*

$$D_n = \{X = (x_1, \dots, x_n) : x_1 \leq \cdots \leq x_n\}, \tag{50}$$

*for $n \geq 1$. The vector $X \in D_n$ denotes the initial positions, with the initial velocities represented by $\lambda = (\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$. In the SPD setting, the particle $k \in \{1, \dots, n\}$ is described by the triple $\left(x_k, \lambda_k, \frac{1}{n}\right)$. The last component corresponds to its mass. Before the first collision with another particle, each particle moves along the real line at a uniform velocity. When a collision occurs, the particles stick together, creating a cluster. The new mass is obtained as the ratio of the number of particles involved in the collision to n, whereas the new velocity equals the average velocity of the particles before the collision. In the event of a collision involving multiple clusters, a unique cluster is formed in such a way that total mass and momentum are preserved.*

Since we assume the monotonicity of $(\mu, \psi)$, there is no collision of particles and the implementation of SPD is straightforward.

We introduce $\mathscr{I}$ as the space of CDFs whose functions $\theta$ are of the form

$$\theta(x) = \mu((-\infty, x]) \in [0, 1], \tag{51}$$

where $\mu \in \mathcal{P}(\mathbb{R})$.

Consider $M \geq 1$, $H$ the Heaviside function $x \mapsto \mathbf{1}_{\{x \geq 0\}}$, $*$ the convolution operator, and $\delta_e$ the Dirac mass at $e \in \mathbb{R}$. Define

$$D_M = \{e = (e_1, \dots, e_m, \dots, e_M) \in \mathbb{R}^M \,|\, e_1 \leq \cdots \leq e_m \leq \cdots \leq e_M\}, \tag{52}$$

and

$$\mathcal{I}_M = \left\{ \theta \in \mathscr{I} | \theta(\cdot) = H * \left( \frac{1}{M} \sum_{m=1}^{M} \delta_{e_m} \right), e \in D_M \right\}. \tag{53}$$

Consider, for $M \geq 0$ and $n \in \{0, ..., N\}$, the set of functions

$$\mathcal{K}_n^M = \{\psi : \mathcal{P}_n \times \mathbb{R} \to [0, 1] \,|\, \text{for } p \in \mathcal{P}_n, \psi(p, \cdot) \in \mathcal{I}_M\}. \tag{54}$$

Now, we can introduce $\mathfrak{T}$, the discrete version of $\mathcal{T}$. It operates on empirical

CDF belonging to $\mathscr{I}_M$,

$$\mathfrak{T}_h^M(p,\theta) = H * \left( \frac{1}{M} \sum_{m=1}^{M} \delta_{e_m^p(h)} \right) \in \mathcal{I}_M, \text{ where } (h,p,\theta) \in (0,\infty) \times \mathbb{R}^d \times \mathcal{I}_M. \quad (55)$$

$(e_m^p(h))_{1\leq m\leq M}$ consists of a set of particle positions. Taking into account $M$ particles $(e_m^p)_{1\leq m\leq M}$, the positions at time $t \in [0,h]$ are expressed by

$$e_m^p(t) = e_m(0) + \bar{F}_m(p)t, \quad (56)$$

where $e_m(0) \in D_M$ is the initial position and $(\bar{F}_m(p))_{1\leq m\leq M}$ are the velocities given by $\bar{F}_m(p) = -\int_{(m-1)/M}^{m/M} \mu(p,y)\, dy$.

Concerning the approximation of $\mathcal{D}_h(\psi)$, it is possible to obtain

$$\bar{v}_n(p,e) = H * \left( \frac{1}{2dM} \sum_{i=1}^{I} \sum_{m=1}^{M} \delta_{e_m^i} \right)(e), \quad (57)$$

and we have, for $\psi \in \mathcal{K}_{n+1}^M$,

$$\mathfrak{D}_n^M(\psi) = \bar{v}_n \in \mathcal{K}_n^{2dM}. \quad (58)$$

Returning to the scheme, it will have two main versions. In Case 1, at each step, all particles are retained. The general approach will consist of iterating the two operators $\mathfrak{T}$ and $\mathfrak{D}$.

In Case 2, the particle count remains constant between successive steps. After applying $\mathfrak{T}$, we employ another operator $\mathfrak{R}$ to decrease the number of particles. For $1 \leq m \leq M$ and $\psi \in \mathcal{I}^M$,

$$\mathfrak{R}^{M,m}(\psi) \in \mathcal{I}^m. \quad (59)$$

We describe the schemes corresponding to Case 1 and Case 2.

**Definition 3.8.** *(Scheme corresponding to Case 1) Let $M \geq 1$. We have $M_n = (2d)^{N-n}M$, with $0 \leq n \leq N$.*

1. *If $n = N$, define $e_N = (\Lambda, \ldots, \Lambda) \in D_{M_N}$. The terminal condition $\phi = \mathbf{1}_{\{e\geq\Lambda\}}$ corresponds to the empirical CDF. For every $p \in \mathcal{P}_N$, fixing $u_N^{N,M}(p,\cdot)$, we obtain $u_N^{N,M} \in \mathcal{K}_N^{M_N}$.*

2. *At $n < N$, with $u_{n+1}^{N,M} \in \mathcal{K}_{n+1}^{M_{n+1}}$, we have*

$$\bar{u}_n^{N,M} = \mathfrak{D}_n^{M_{n+1}} \left( u_{n+1}^{N,M} \right) \in \mathcal{K}_n^{M_n}, \tag{60}$$

*and*

$$u_n^{N,M}(p, \cdot) = \mathfrak{T}_h^{M_n} \left( p, \bar{u}_n^{N,M}(p, \cdot) \right), \tag{61}$$

*for every $p \in \mathcal{P}_n$.*

*Hence, $\mathcal{V}(0, P_0, \cdot)$ is approximated by $u_0^{N,M}(P_0, \cdot)$.*

**Definition 3.9.** *(Scheme corresponding to Case 2) Let $M \geq 1$.*

1. *If $n = N$, define $e_N = (\Lambda, \ldots, \Lambda) \in D_M$. The terminal condition $\phi = \mathbf{1}_{\{e \geq \Lambda\}}$ corresponds to the empirical CDF. For every $p \in \mathcal{P}_N$, fixing $v_N^{N,M}(p, \cdot)$, we obtain $v_N^{N,M} \in \mathcal{K}_N^M$.*

2. *At $n < N$, with $v_{n+1}^{N,M} \in \mathcal{K}_{n+1}^M$, we have*

$$\bar{v}_n^{N,M} = \mathfrak{D}_n^M \left( v_{n+1}^{N,M} \right) \in \mathcal{K}_n^{2dM}. \tag{62}$$

*For every $p \in \mathcal{P}_n$, we obtain*

$$\check{v}_n^{N,M}(p, \cdot) = \mathfrak{R}^{2dM,M} \left( \bar{v}_n^{N,M}(p, \cdot) \right), \tag{63}$$

*and*

$$v_n^{N,M}(p, \cdot) = \mathfrak{T}_h^M \left( p, \check{v}_n^{N,M}(p, \cdot) \right), \tag{64}$$

*Hence, $\mathcal{V}(0, P_0, \cdot)$ is approximated by $v_0^{N,M}(P_0, \cdot)$.*

Regarding Case 1, the number of particles at each iteration for a given branch is $M_n = (2d)^{N-n}M$, which is determined by backward induction. The discrete transport operator has to be applied $(2d)^n$ times at step $n$, with the position of $M_{n+1}$ particles being recalculated at each implementation, resulting in a complexity of $O((2d)^{N-1}M)$, for $n < N$. At step 0, the evaluation of (57) costs $O(2dM)$. For all these arguments, the complexity is $O((2d)^N MN)$.

Using Case 2, the unique gain lies in the number of particles at each iteration for a given branch, which is always $M$. The complexity of the discrete transport operator at step $n$ is $(2d)^n M$. Aggregating over all steps leads to a total complexity of order $O((2d)^N M)$. At step 0, the evaluation of (57) costs $O(2dM)$. For all these arguments, the complexity is $O((2d)^N M)$.

**Definition 3.10.** *For $1 \leq m \leq M$ and $\xi \in \mathbb{R}^M$, $S^M$ is the operator that sorts $M$ particles in increasing order,*

$$S^M(\xi) \in \mathcal{D}_M. \tag{65}$$

*At each step $n$, for a collection of particles $(2dM)$,*

$$S^{2dM}(\xi) = \left(\tilde{S}^{2d,1}(\xi), \ldots, \tilde{S}^{2d,M}(\xi)\right), \tag{66}$$

*where, for $1 \leq i \leq M$, $\tilde{S}^{2d,i}(\xi) = \left(S_j^{2dM}(\xi)\right)_{2d(i-1)<j\leq 2di} \in \mathcal{D}_{2d}$. $S_j^{2dM}(\xi)$ represents the $j$-th coordinate of the vector $S^{2dM}(\xi)$. Recall that*

$$\psi^\xi = H * \left(\frac{1}{2dM}\sum_{m=1}^{2dM} \delta_{S_m^{2dM}(\xi)}\right) \in \mathcal{I}^{2dM}, \tag{67}$$

*connected to $\xi \in \mathcal{D}_{2dM}$. Hence, we define*

$$\mathfrak{R}^{2dM,M}(\psi^\xi) = H * \left(\frac{1}{M}\sum_{m=1}^{M} \delta_{\tilde{e}_m}\right) \in \mathcal{I}^M, \tag{68}$$

*and, for $1 \leq m \leq M$, the optimal position (optim) is $\tilde{e}_m = S_{d(2m-1)}^{2dM}(\xi)$.*

## 3.2 Simplified functional brownian framework

We focus on a specific instance of the process $P$ leading to a considerable reduction in the computational cost of the numerical algorithm.

In this particular case, $P$ is represented by a Brownian motion function,

$$P_t = \mathfrak{P}\left(t, W_t\right), \tag{69}$$

with $\mathfrak{P} : [0,T] \times \mathbb{R}^d \mapsto \mathbb{R}^d$.

This framework is especially relevant when $P$ is either a Brownian motion or a Geometric Brownian Motion.

$P$ is approximated by

$$\hat{P}_t = \mathfrak{P}\left(t, \hat{W}_t\right), t \in \pi. \tag{70}$$

For $0 \leq n \leq N$, the discrete support of $\hat{P}_{t_n}$ is

$$\mathcal{P}_n = \{p = \mathfrak{P}\left(t_n, w\right), w \in \mathfrak{S}_n\}. \tag{71}$$

For $p = \mathfrak{P}(t_n, w)$,

$$\hat{P}_{t_{n+1}}^{t_n,p} = \mathfrak{P}\left(t_{n+1}, w + \Delta\hat{W}_n\right), \tag{72}$$

and, for $1 \leq i \leq 2d$,

$$\left(\hat{P}_{t_{n+1}}^{t_n,p}\right)^i = \mathfrak{P}\left(t_{n+1}, w + \omega_h^i\right). \tag{73}$$

In this setting, using Case 1, $M_{n+1} = O(M(N-n+1)^d)$ and, at each step $n$, the number of branches is $O(n^d)$. Hence, for large $N$, the complexity is given by $O\left(M\sum_{n=0}^{N-1}(N-n)^d n^d\right) = O\left(MN^{2d+1}\right)$.

For Case 2, $M_{n+1} = M$ and, at each step $n$, the number of branches is $O(n^d)$. Hence, for large $N$, the complexity is given by $O\left(M\sum_{n=0}^{N-1} n^d\right) = O\left(MN^{d+1}\right)$.

In terms of complexity, both cases suffer from the curse of dimensionality. Nevertheless, in the brownian setting, models of dimension roughly 5 can be handled, specifically using Case 2.

## 3.3 Error decomposition

In this context, we consider the error decomposition presented in [12], which is established only for Case 1.

The error is given by

$$\mathrm{err}(N, M) = \int_{\mathbb{R}} |\gamma_0(P_0, e) - \mathcal{V}(0, P_0, e)| de, \tag{74}$$

with $\gamma_n = u_n^{N,M}$.

**Theorem 3.2.** *Suppose that Assumption 3.1 holds and let $\epsilon \in (0,1)$. Considering Case 1, we obtain the error bound*

$$err(N, M) \leq C\left(\frac{1}{M} + \frac{\sqrt{h}}{\epsilon^2} + \epsilon\right), \tag{75}$$

*where $C$ is a positive constant. If $\epsilon = h^{\frac{1}{6}}$ and $M = \frac{1}{\epsilon}$,*

$$\int_{\mathbb{R}} \left|u_0^{N,M}(P_0, e) - \mathcal{V}(0, P_0, e)\right| de \leq 3Ch^{\frac{1}{6}}. \tag{76}$$

The proof of this theorem can also be found in [12].

# 4 Numerics

## 4.1 Numerical algorithm for Case 2

In this section, we implement one model to evaluate the scheme corresponding to Case 2 presented above. The model under consideration is the Linear model, presented in [11] and [12]. In addition, we include a numerical proxy against which we compare our results.

We proceed with the Linear model

$$dP_t = \sigma\, dW_t,\ dE_t = \left(\frac{1}{\sqrt{d}}\sum_{l=1}^{d} P_t^l - Y_t\right) dt, \tag{77}$$

with terminal function $(p, e) \mapsto \phi(p, e) = \mathbf{1}_{\{e\geq 0\}}$. $W$ denotes a Brownian motion with dimension $d$ and $\sigma > 0$.

Regarding the numerical proxy, we apply a change of variables to reduce the model to one dimension, and then we use the method proposed in [4]. It is possible to prove

$$\mathcal{V}(t, p, e) = u\left(t, e + \frac{1}{\sqrt{d}}\sum_{l=1}^{d} p^l\right), \tag{78}$$

where $u$ is a solution to

$$\partial_t u - u\partial_\xi u + \frac{\sigma^2}{2}\partial^2_{\xi\xi} u = 0 \,\text{and}\, u(T, \xi) = \phi(\xi). \tag{79}$$

The model allows us to proceed within a simplified functional Brownian framework, thanks to the equation governing $P$.

The algorithm takes the following form:

Consider the uniform time grid

$$\pi = \{t_0 = 0 \leq t_1 \leq \cdots \leq t_n \leq \cdots \leq t_N = T\}. \tag{80}$$

We have

$$T = 1,\ N = 20 \text{ and } h = t_{n+1} - t_n = \frac{T}{N} = 0.05, \tag{81}$$

for $n = 0, \ldots, N - 1$.

$M = 3500$ is the number of particles, the volatility is $\sigma = 1$ and the dimension of our model is $d = 4$.

The terminal condition is given by

$$\Phi(e) = \mathbf{1}_{\{e \geq 0\}}. \tag{82}$$

The driver of the Linear model is

$$\mu(p, y) = \frac{1}{d} \sum_{l=1}^{d} p^l - y. \tag{83}$$

At each time step, the Brownian motion in $\mathbb{R}^d$ is approximated using a cubature rule with $2d$ points.

For $l = 1, \ldots, d$, $e^l$ is the canonical basis of $\mathbb{R}^d$ and we define $\omega_h^i = -\sqrt{dh}e^l$, if $i = 2l$ or $\omega_h^i = \sqrt{dh}e^l$, if $i = 2l - 1$, $i = 1, \ldots, 2d$.

Afterwards, the resulting increments are collected in an array of size $2d$.

We construct a tree, where $Z_0 = \{0\}$.

Set $z^i = -e^l$, if $i = 2l$ or $z^i = e^l$, if $i = 2l - 1$, $l = 1, \ldots, d$ and $i = 1, \ldots, 2d$, and arrange them in an array with $2d$ elements.

For $n = 0, \ldots, N-1$ and $z \in Z_n$, introduce $Z_{n+1}$ as $Z_{n+1} = \{z + z^i, i = 1, \ldots, 2d\}$, merging duplicates.

For every node $z \in Z_n$, represent its children at level $n + 1$ as pairs where the first element is the index assigned to the node we obtain in $Z_{n+1}$ and the second is the index associated with the respective Brownian increment.

We build the Brownian grid $P_n$, for $n = 0, \ldots, N$. The initial condition is given by $P_0 = 0 \in \mathbb{R}^d$.

$P_n$ is the array that contains the parent positions at $n = 0, \ldots, N - 1$. Define $P_{n+1}$ as a zero array, whose length is $(K_{n+1}, d)$, $K_{n+1}$ being the number of nodes at time $t_{n+1}$.

Define $P_{n+1}^{child} = P_n^{parent} + \sigma \Delta W_n$, for $n = 0, \ldots, N - 1$.

The transport step for a particle vector $e_m, m = 1, \ldots, M$, is given by

1. $e_{(m)}$ is the sorted representation of $e_m$.

2. $w_m = \frac{m+0.5}{M}, m = 0, \ldots, M - 1$ are the empirical CDF levels.

3. $v_m = \mu(p, w_m) = \frac{1}{\sqrt{d}} \sum_{l=1}^{d} p^l - w_m$ are the drift terms.

4. Set $e_m^{new} = e_{(m)} - v_m h$.

5. Define $\bar{e}_m^{new} = max\{e_1^{new}, \ldots, e_m^{new}\}, m = 1, \ldots, M$ and obtain the new particle vector.

Concerning the diffusion part, consider $e_{n+1}^j \in \mathbb{R}^M$ the vector defined above with $j$ being a child node. Construct the concatenated vector

$$e_{2dM} = (e_{n+1}^{j_1}, \ldots, e_{n+1}^{j_{2d}}) \in \mathbb{R}^{2dM}, \tag{84}$$

where the children follow the order determined by the tree.

We apply the reduction operator to decrease the particles from $2dM$ to $M$.

1. $e_{(2dM)}$ is the concatenated vector arranged in increasing order.

2. For $m = 1, \ldots, M$, compute $k_m = d(2m-1)$. When required, project it onto the admissible range $\{1, \ldots, 2dM\}$.

3. Then $e_m^{red} = e_{k_m}, m = 1, \ldots, M$.

At each time level, apply the splitting method consisting of a diffusion step with the reduction operator, followed by a transport step.

At $t_N = T$, the particle vector is defined by $e_N^i = 0 \in \mathbb{R}^M$, $i = 1, \ldots, K_N$.

For $n = N-1, \ldots, 0$ and $i = 1, \ldots, K_n$, $p_n^i$ are the nodes at time $t_n$.

For every node $i$,

1. Apply the diffusion step by aggregating the particles from all child nodes $j$ at time $t_{n+1}$.

2. Perform the reduction operator to get $e_m^{red}, m = 1, \ldots, M$.

3. Apply the transport step, $e_n^i = T_h^M(e^{red}, p_n^i)$.

Retain $e_n^i$ associated with all nodes at level $n$.

At $t_0 = 0$, $p_0 = 0$ is the unique node. $e_0^1$ represents the vector of particles. At the initial points $(t_0, p_0, e_0)$, the decoupling field is approximated by

$$\mathcal{V}(t_0, p_0, e_0) \approx P(e_0^1 \leq e_0) = \frac{1}{M} \sum_{m=1}^{M} \mathbf{1}_{\{e_{0,m}^1 \leq e_0\}}. \tag{85}$$

With respect to the numerical proxy, we follow the Bossy-Talay discretisation

$$X_{t+\Delta t}^i = X_t^i - \frac{i - 0.5}{M} \Delta t + \sigma \Delta W_t^i, \tag{86}$$

where $r^i = \frac{i-0.5}{M}$ is the empirical rank associated with particle $X_t^i$.

The decoupling field is approximately estimated by

$$\mathcal{V}(t_0, p_0, e_0) = \frac{1}{M} \sum_{i=1}^{M} \mathbf{1}_{\{X^i_{T-t_0} \geq 0\}}. \tag{87}$$

For complete details about the numerical algorithm for the proxy and the Bossy-Talay discretisation, please see [4].

## 4.2 Numerical results for Case 2

We begin by analysing the Linear model, plotting the decoupling field $\mathcal{V}$ under Case 2 and contrasting it with the numerical proxy.

As shown in Figure 1, both lines exhibit the same behaviour. Consequently, Case 2 successfully captured the proxy solution. This result is crucial for the empirical validation of our numerical algorithm.

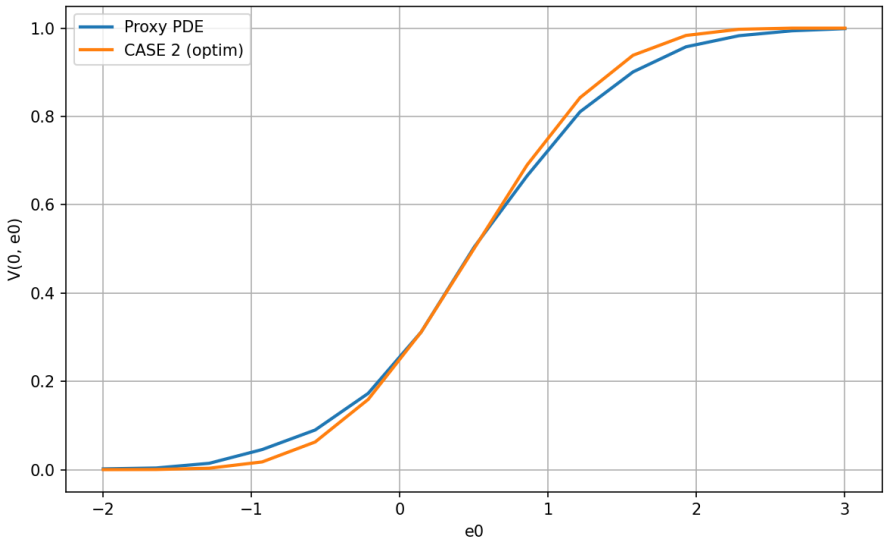


Figure 1: Linear model with Case 2 with M = 3500 and N = 20

However, the implementation of the numerical algorithm results in an error. The L1-error associated with Case 2 is computed with respect to the numerical proxy solution. Let $N = \{2, 4, 8, 16, 18, 20\}$ and, as before, set $h = \frac{T}{N}$. As illustrated in Figure 2, for the time step, we obtain an empirical convergence rate of 0.516 using Case 2, which compared to the value given for Case 1 in Theorem 3.2 represents an improvement.

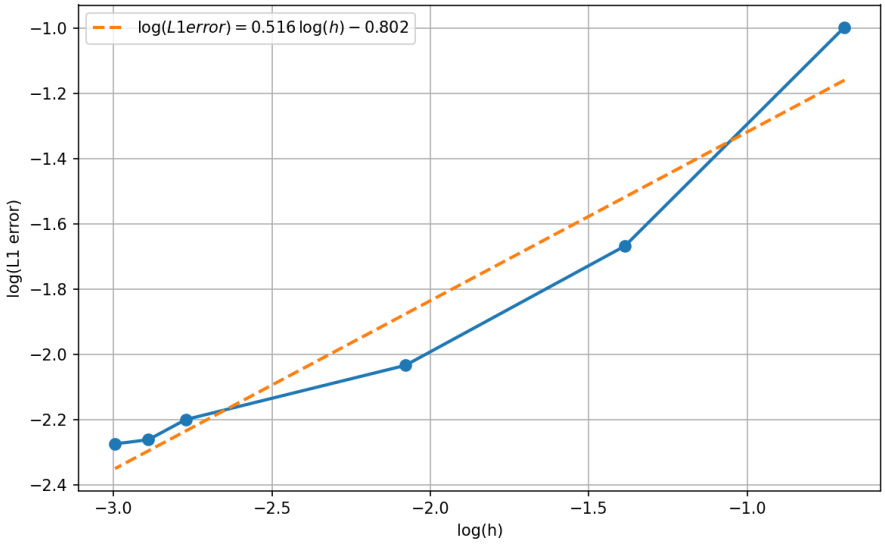


Figure 2: Time step convergence rate for the Linear model with Case 2

## 4.3 Numerical algorithm for Case 2 applied to the carbon market

Consider an alternative model whose aim is to apply Case 2 to the carbon market (following [10] and Chapter 6 of [13]).

We analyse the FBSDE

$$\begin{cases} dE_t = \mu_E \left(D_t, S_t^C, S_t^G\right) dt \\ dY_t = Z_t \, dW_t \\ E_0 = 0 \\ Y_T = \phi(E_T) = \Lambda \mathbf{1}_{[E^{cap},\infty)}(E_T), \end{cases} \tag{88}$$

where $S^C$ and $S^G$ denote the coal and gas prices, respectively.

This model is implemented using parameters estimated from real data, which are shown in Table 1. We use data corresponding to the United Kingdom (UK) energy market for the period from 2012 to 2014. In the UK, excluding renewable electricity generation, most electricity is produced using coal and gas, so we uniquely include these two fuels in our model. Their prices ($S^C$ and $S^G$) are valued at Great Britain Pound (Pound Sterling) (GBP) per megawatt-hour (MWh) of electricity, that is, they express the fuel cost involved in producing 1MWh of electricity. $D$ corresponds to the electricity demanded in MWh during a 24-hour interval, which is generated

exclusively by the two fuels in our model. The method employed to estimate these three processes is non-linear regression. $Y$ corresponds to the futures contract price of a certificate maturing at terminal time $T$.

First, we present the seasonality functions for the demand, coal price, and gas price, respectively:

$$\begin{aligned} h^D(t) &= a^D + b^D t + c_1^D \cos\left(\frac{2\pi t}{63/252}\right) + d_1^D \sin\left(\frac{2\pi t}{63/252}\right) \\ &\quad + c_2^D \cos(2\pi t) + d_2^D \sin(2\pi t), \\ h^C(t) &= a^C + b_1^C t + b_2^C t^2 + c^C \cos(2\pi t) + d^C \sin(2\pi t), \\ h^G(t) &= a^G + b_1^G t + b_2^G t^2 + b_3^G t^3 + c_1^G \cos\left(\frac{2\pi t}{63/252}\right) \\ &\quad + d_1^G \sin\left(\frac{2\pi t}{63/252}\right) + c_2^G \cos(2\pi t) + d_2^G \sin(2\pi t). \end{aligned} \tag{89}$$

Although $h^C$ uses annual data, both $h^D$ and $h^G$ have quarterly and yearly seasonality. In one year, the number of trading days is 252, due to weekends and holidays. Consequently, quarterly seasonality is represented by 63/252. These functions are used in the demand, coal price and gas price processes

$$\begin{aligned} D_t &= \kappa^D \exp\{h^D(t) + X_t^D\}, \\ S_t^C &= \kappa^C \exp\{h^C(t) + X_t^C\}, \\ S_t^G &= \kappa^G \exp\{h^G(t) + X_t^G\}, \end{aligned} \tag{90}$$

with $X^i, i = D, C, G$ being diffusion processes and $\kappa^i, i = D, C, G$ constants.

The estimation of $X_t^D$ and $X_t^C$ relies on an Ornstein-Uhlenbeck process. Hence,

$$\begin{aligned} dX_t^D &= -\lambda^D X_t^D dt + \sigma^D dW_t^D, X_0^D = 0, \\ dX_t^C &= -\lambda^C X_t^C dt + \sigma^C dW_t^C, X_0^C = 0. \end{aligned} \tag{91}$$

The parameters $\lambda^D$, $\sigma^D$, $\lambda^C$ and $\sigma^C$ are estimated using the maximum likelihood estimation method.

The Ornstein-Uhlenbeck process is not appropriate for $X^G$ since the differenced gas series is not normally distributed. Consequently, we implement a process with non-constant volatility, given by

$$dX_t^G = -\lambda^G (X_t^G - \mu^G) dt + \sqrt{\upsilon^G(X_t^G)} dW_t^G. \tag{92}$$

The function $\upsilon^G$ can be written as

$$\upsilon^G(x) = \frac{2\lambda^G\sqrt{\delta^2+x^2}\left(x - \frac{\beta\delta}{\sqrt{\alpha^2-\beta^2}}\right)}{\alpha x - \beta\sqrt{\delta^2+x^2}}. \tag{93}$$

The parameter $\lambda^G$ is obtained through standard non-linear regression, while the parameters $\alpha$, $\beta$ and $\delta$ are determined using the maximum likelihood estimation method.

For $i \in \{C, G\}$ and $x \in [0, \xi^i_{max}]$, $e_i(x) = \alpha_i\eta_i$ are the marginal emissions stack and the BAU bid rates are $b_i^{BAU}(x, s) = \gamma_i + \delta_i\eta_i x$. The role of these functions is to derive the bid stack. They have a parametric representation, being $e_i$ a constant function and $b_i^{BAU}$ a linear function, in order to simplify the problem.

The parameters $\alpha_C$ and $\alpha_G$ represent the emissions intensity factors employed to determine the value of clean dark spread and clean spark spread. These indicators quantify the profitability of producing electricity using only coal and gas, respectively. $\eta_C$ and $\eta_G$ denote the quantities of coal and gas measured in MWh-equivalent units, respectively, needed to generate 1MWh of electricity. Finally, $\gamma_C$, $\delta_C$, $\gamma_G$ and $\delta_G$ are estimated through the standard non-linear least squares method.

The bid curve is, for $i \in \{C, G\}$ and $x \in [0, \xi^i_{max}]$,

$$b_i(x, a, s_i) = a\alpha_i\eta_i + s_i(\gamma_i + \delta_i\eta_i x), \tag{94}$$

and it is possible to obtain the inverse function

$$b_i^{-1}(p, a, s_i) = \xi^i_{max} \wedge \left(\frac{p - a\alpha_i\eta_i - s_i\gamma_i}{s_i\delta_i\eta_i}\right) \vee 0, \tag{95}$$

with $p \in \left[\underline{b_i}(a, s_i), \overline{b_i}(a, s_i)\right]$.

Consequently, invert $p \mapsto \sum_i b_i^{-1}(p, a, s_i), i = \{C, G\}$, represents the bid stack $b(x, a, s)$. With this result, the formula for the emissions rate is

$$\mu_E(x, a, s) = k\sum_i \left[\alpha_i\eta_i b_i^{-1}(b(x, a, s), a, s_i)\right], \tag{96}$$

recall (29). $k$ represents the annual trading days.

Finally, the cap on total emissions is $E^{cap} = T\eta\mu_E(D_0, 0, S_0^C, S_0^G)$. We define $E^{average} = T\mu_E(D_0, 0, S_0^C, S_0^G)$ as the accumulated amount of emissions when the allowance price is zero and demand, coal price, and gas price are equal to their

initial mean values. This quantity corresponds to an initial value of the expected cumulative emissions during the compliance period in the absence of any emissions restrictions. Then, according to [16], $E^{cap}$ is given by $E^{cap} = \eta E^{average}$.

Table 1: Parameters estimates

| Parameter | Estimate | Parameter | Estimate | Parameter | Estimate |
|---|---|---|---|---|---|
| $a^D$ | 3.32 | $b^D$ | -0.0867 | $c_1^D$ | -0.0119 |
| $d_1^D$ | -0.0212 | $c_2^D$ | 0.145 | $d_2^D$ | 0.0301 |
| $a^C$ | 4.16 | $b_1^C$ | -0.159 | $b_2^C$ | 0.0103 |
| $c^C$ | 0.031 | $d^C$ | -0.00615 | $a^G$ | 3.88 |
| $b_1^G$ | 0.698 | $b_2^G$ | -0.419 | $b_3^G$ | 0.0587 |
| $c_1^G$ | -0.00613 | $d_1^G$ | 0.0124 | $c_2^G$ | 0.0817 |
| $d_2^G$ | 0.1 | $\kappa^D$ | 24000 | $\kappa^C$ | 0.1228 |
| $\kappa^G$ | 0.3412 | $\lambda^D$ | 80.3 | $\sigma^D$ | 1.32 |
| $\lambda^C$ | 10.1 | $\sigma^C$ | 0.16 | $\lambda^G$ | 8.08 |
| $\mu^G$ | 0.0 | $\alpha$ | 14.7 | $\beta$ | 0.524 |
| $\delta$ | 0.126 | $\xi_{max}^C$ | $5.9013 \times 10^5$ | $\xi_{max}^G$ | $3.997 \times 10^5$ |
| $\alpha_C$ | 0.411 | $\eta_C$ | 2.04 | $\alpha_G$ | 0.973 |
| $\eta_G$ | 2.63 | $\gamma_C$ | 0 | $\delta_C$ | $6.05 \times 10^{-6}$ |
| $\gamma_G$ | 0.178 | $\delta_G$ | $3.73 \times 10^{-6}$ | $k$ | 252 |

We now adapt Case 2 to make it compatible with the carbon market model. Contrary to the Linear model, we are no longer in a functional Brownian setting, due to the dynamics of $P$. Therefore, the analysis is conducted under a generic diffusion framework.

We conduct several experiments using this model. Given the complexity associated with the definition of the diffusion component in this particular setting, various combinations of $M$ and $N$ are considered, resulting in different values of $h = \frac{T}{N}$.

Furthermore, we consider the same time grid $\pi$ with $T = 8$. In addition, we have $d = 3$, $\Lambda = 83.47$, and $\eta = 0.2$.

The terminal condition of this model is

$$\Phi(e) = \Lambda \mathbf{1}_{\{e \geq E^{cap}\}}. \tag{97}$$

The cubature rule with $2d$ points is used again to generate Brownian increments, although we build a purely combining tree rather than a recombining one.

The initial point is $X_0 = (0, 0, 0, 0)$. Each node produces $2d$ children resulting from an Euler discretization of the diffusion processes. Hence, if $K_n$ is the number of nodes at level $n$, we get $K_{n+1} = 2dK_n$ at $n + 1$, for $n = 0, \ldots, N - 1$.

Consequently, we have

$$X_{n+1}^{child} = X_n^{parent} + b\left(X_n^{parent}\right) h + \sigma\left(X_n^{parent}\right) \Delta W, \tag{98}$$

with the drift coefficient $b$ and the volatility coefficient $\sigma$ introduced above.

The diffusion part and the reduction operator are the same as before. Regarding the transport step, $e_m$ is the particle vector for $m = 1, \ldots, M$ and $e_{(m)}$ is the vector obtained after sorting its elements.

We define

$$y_m = \frac{m + 0.5}{M} \tag{99}$$

as the empirical CDF levels, and

$$a_m = \Lambda y_m \tag{100}$$

are the equivalent allowance levels.

For all $a_m$, the drift terms are derived from the emissions rate as

$$v_m = \mu_E\left(X_n^i, a_m, t\right), \tag{101}$$

and the transport operator is given by

$$e_m^{new} = e_{(m)} + h v_m. \tag{102}$$

Finally, we get the new particle vector,

$$\bar{e}_m^{new} = max\{e_1^{new}, \ldots, e_m^{new}\}, \tag{103}$$

for $m = 1, \ldots, M$.

At $t = T$ and for $i = 1, \ldots, K_N$ let

$$e_N^i = E^{cap}. \tag{104}$$

After that, the backward recursion follows the same methodology as in the Linear model.

Given the initial condition $X_0$, we approximate the decoupling field by

$$\mathcal{V}(0, X_0, e) = E^{cap} \frac{1}{M} \sum_{m=1}^{M} \mathbf{1}_{\{e^1_{0,m} \leq e\}}. \tag{105}$$

Next, our goal is to derive forward paths of both the emissions process and the allowance prices. Hence, at each time level of the pure cubature tree, we keep the particle vectors

$$\{e^i_n\}_{i=1}^{K_n}, n = 0, \ldots, N, \tag{106}$$

to compute the decoupling field throughout the forward paths.

When the backward recursion finishes, forward trajectories are simulated by random sampling of the branches of the tree. A branch $(k_0, \ldots, k_{N-1})$ is generated by sampling $k_n \in \{1, \ldots, 2d\}$ uniformly for each time index $t_n$.

The initial point is $E_0 = 0$. Recursively updated, the current node, represented by $e^{node}_n$, is propagated along the pure tree structure.

The empirical CDF is given by

$$u_n(E_n) = \frac{1}{M} \sum_{m=1}^{M} \mathbf{1}_{\{e^{node}_{n,m} \leq E_n\}}. \tag{107}$$

Multiplying by $\Lambda$, we get the allowance prices $Y_n$.

Accordingly, the sample trajectories of $(E_t)_{t\in[0,T]}$ and $(Y_t)_{t\in[0,T]}$ are constructed using an Euler discretization of the emission process

$$E_{n+1} = E_n + h\mu_E(X_n, Y_n, t_n). \tag{108}$$

Additionally, we conduct another experiment, evaluating the decoupling field at $e_{ref} = \Lambda$. After backward recursion is performed, for every node $i$ and time index $t_n$, assess the decoupling field using empirical CDF

$$u_n(P^i_n, e_{ref}) = \frac{1}{M} \sum_{m=1}^{M} \mathbf{1}_{\{e^i_{n,m} \leq e_{ref}\}}, \tag{109}$$

where $e^i_{n,m}, m = 1, \ldots, M$ is the vector of particles.

As before, multiplying by $\Lambda$, we obtain the allowance prices $Y_n$. We select an illustrative branch by recursively choosing the first child at each branching step.

This allows us to observe one possible progression of the allowance price over time.

The final test examines how different values of $\eta$ affect the trajectories of $Y$ and $E$. Here, the only modification is the inclusion of different values of $\eta$ in the numerical algorithm, producing a single trajectory for each $\eta$, instead of several trajectories for the same value of $\eta$.

## 4.4 Numerical results for Case 2 applied to the carbon market

In the context of Case 2 applied to the carbon market, we first analyse the evolution of three sample paths of the processes $Y$ and $E$ with $M = 2000$ and $N = 8$. As illustrated in Figure 3, all trajectories of $Y$ exhibit similar dynamics, remaining at 0 for an initial period, then rising abruptly before stabilising at $\Lambda$.

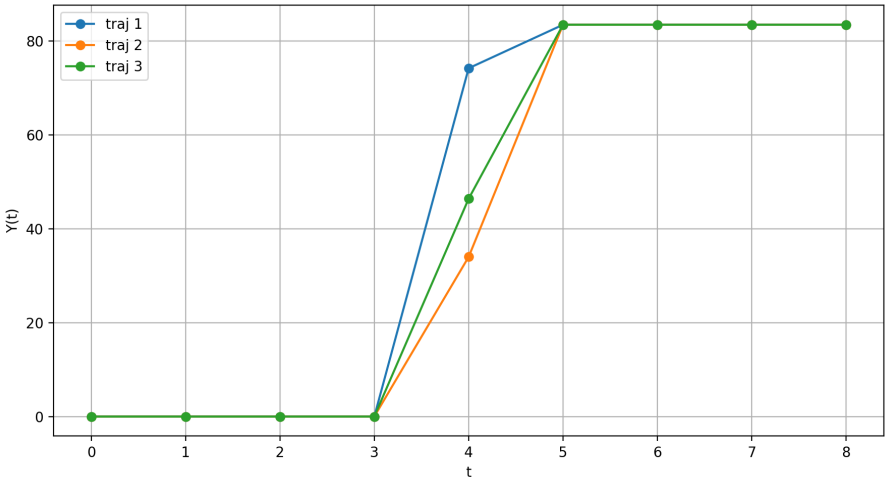


Figure 3: Three sample paths of Y with $M = 2000$ and $N = 8$

Regarding the emissions process $E$, the paths follow a common pattern, increasing continuously over time and exceeding the threshold $E^{cap}$, as shown in Figure 4.

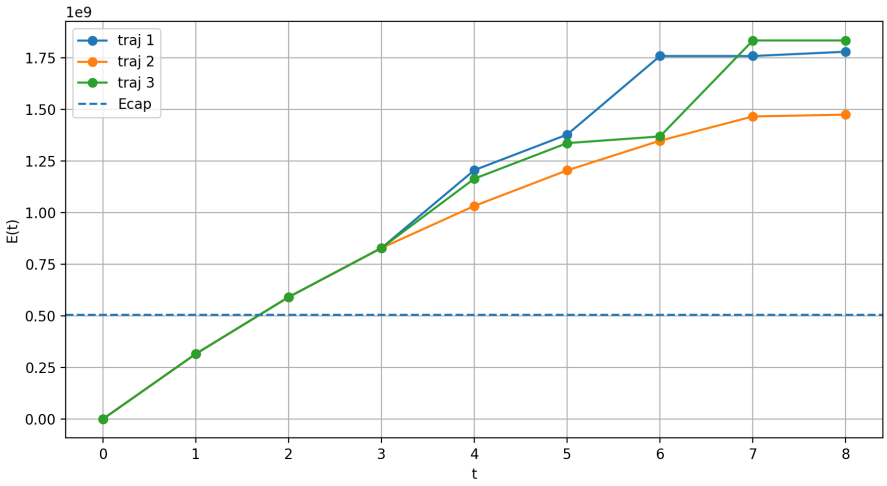


Figure 4: Three sample paths of E with $M = 2000$ and $N = 8$

Next, at $t = 0$, the graphs of the empirical CDF and the decoupling field are connected through the relation $\mathcal{V}(0, P_0, e) = \Lambda u(0, P_0, e)$. We then examine how they change as $e$ varies, with $M = 2500$ and $N = 4$.

In Figure 5, the empirical CDF progressively increases from 0 to 1, while, in Figure 6, the decoupling field grows until $\Lambda$. Both approximations are monotone step functions as a consequence of the CDF.

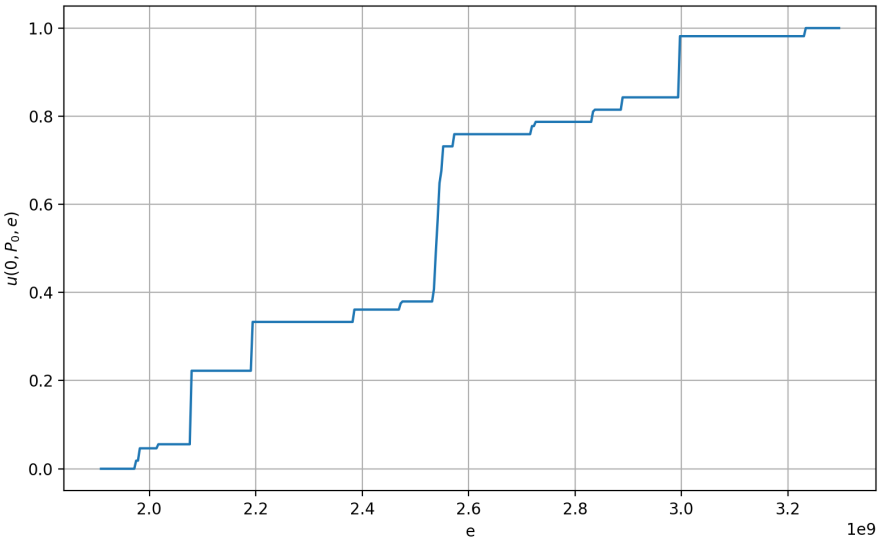


Figure 5: Evolution of the empirical CDF with $M = 2500$ and $N = 4$

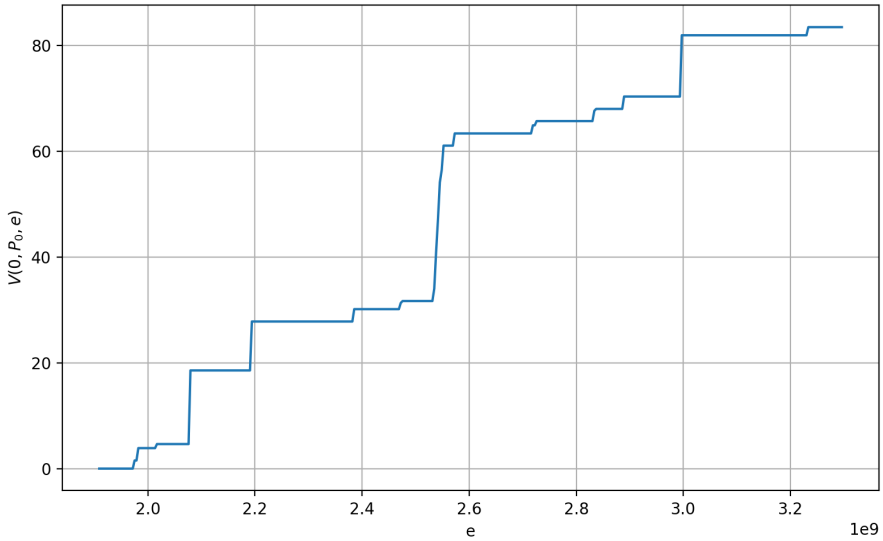


Figure 6: Evolution of the decoupling field with $M = 2500$ and $N = 4$

Afterwards, we investigate $\mathcal{V}(t, P_t, e_{ref})$ over time for a fixed $e_{ref}$ with $M = 2500$ and $N = 4$. In Figure 7, for $e_{ref} = \Lambda$, the process $Y$ remains at 0 until $t = 6$, after which it increases sharply towards $\Lambda$.

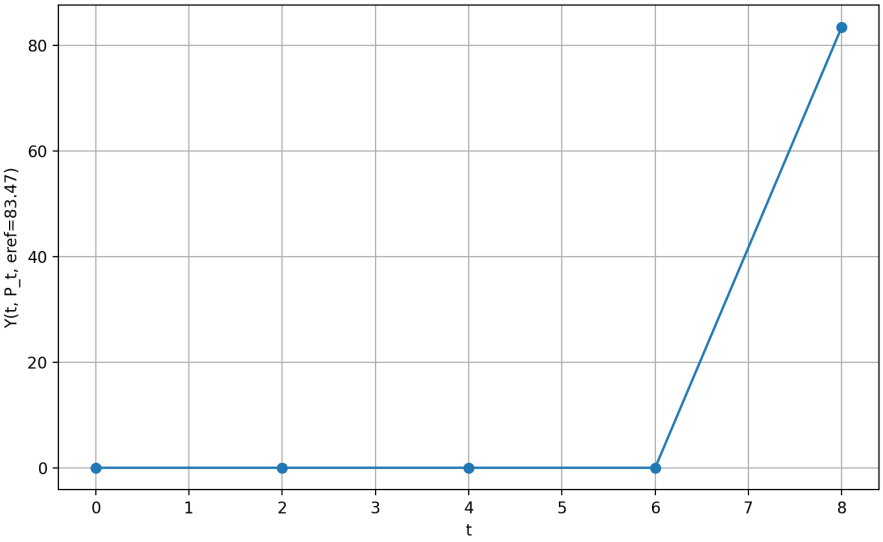


Figure 7: Evolution of the decoupling field with $e_{ref} = \Lambda$ with $M = 2500$ and $N = 4$

Finally, we explore the influence of $\eta$ on the value of $E^{cap}$ and, subsequently, on the sample paths of $Y$ and $E$. In this case, $M = 3000$ and $N = 6$. For $\eta = 0.2$, we have already described the behaviour of the trajectory of $Y$ above. As presented in Figure 8, when $\eta = 1$, $Y_t = 0$, for all $t$. With $\eta = 0.6$, the sample path of $Y$ is equal

to 0 for most of the time horizon before beginning to increase, reaching $\Lambda$ at the terminal time $T = 8$. Consequently, an increase in $E^{cap}$ leads to a decrease in the price of emission allowances.

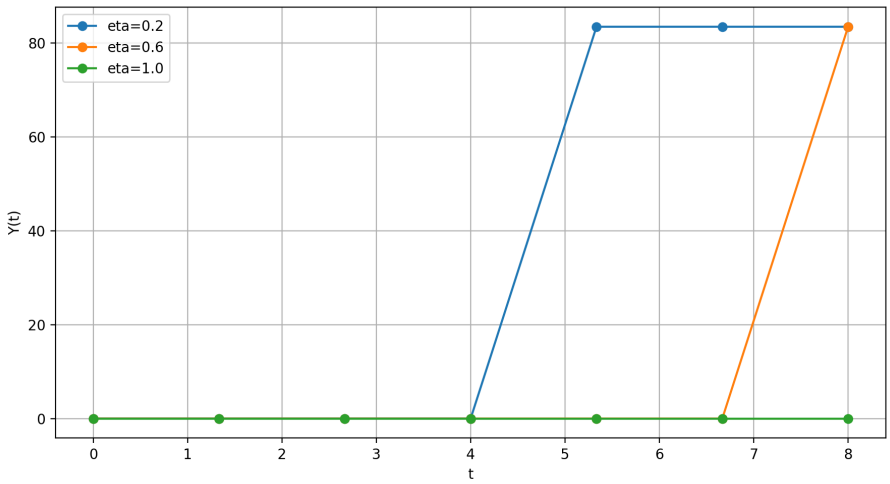


Figure 8: Impact of $\eta$ on the trajectories of $Y$ with $M = 3000$ and $N = 6$

In Figure 9, the trajectories of $E$ evolve similarly for all three values of $\eta$, increasing over time as expected, with only the sample path corresponding to $\eta = 0.2$ deviating slightly near the end of the time interval.

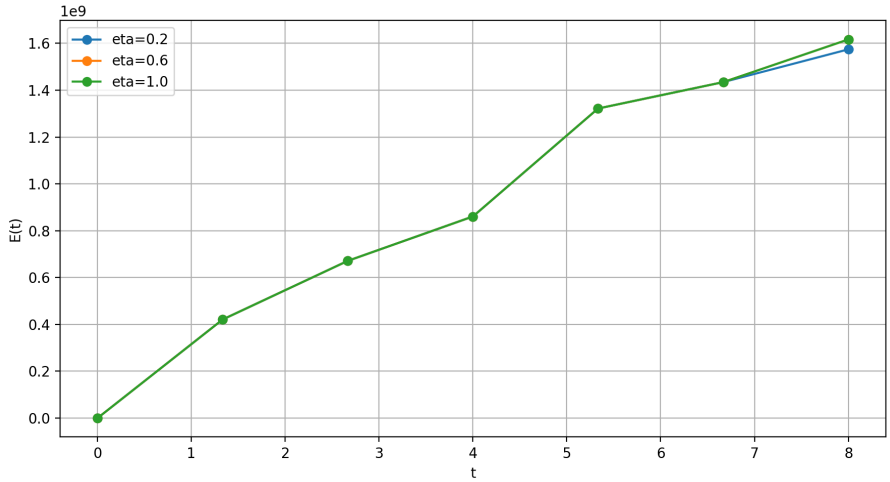


Figure 9: Impact of $\eta$ on the trajectories of $E$ with $M = 3000$ and $N = 6$

# 5 Conclusion

This dissertation focuses on studying a discrete time approximation applied to a carbon emissions market model, where the model is described by a singular FBSDE. The market is characterised by a process representing the certificate price, another related to emissions, and several state variables, namely electricity demand and the prices of coal and gas. All processes evolve forward in time, with the exception of the certificate price, for which only the terminal value is specified.

The numerical approximation employs a splitting scheme, discussed in [12], based on a diffusion operator and a transport operator, together with a variant that preserves the number of particles across two consecutive steps. Moreover, we apply a reduction operator to work with a smaller number of particles. We test the scheme using the Linear model and a proxy solution.

Subsequently, we implement the splitting scheme to the carbon market model. We obtain two graphs illustrating several trajectories of the allowance price process and the accumulated emissions process. Furthermore, we present the evolution of the empirical CDF and the decoupling field at the initial time, computed over a range of values of released emissions. In addition, we analyse the behaviour of the decoupling field for a given level of cumulative emissions as well as the changes observed in the sample paths of both the certificate price process and the total emissions process resulting from distinct values of $\eta$.

However, this topic offers several opportunities for future research. In the splitting scheme, the method is subject to the curse of dimensionality. Therefore, further developments in the diffusion component would be advantageous. It should also be noted that, due to the complexity of the model, certain numerical experiments were conducted using relatively small values of $M$ and $N$. The investigation of more efficient approaches that enable the use of larger values for these parameters may prove beneficial. The incorporation of additional fuels would also merit further study. Such an extension increases the dimensionality of the model and renders the simulations more computationally demanding. Consequently, exploring new methodologies for handling this challenge could be of practical value. Finally, the inclusion of jump processes and Lévy processes in carbon emissions market model constitutes a relevant area for future analysis.

# References


[1] Christian Bender and Jianfeng Zhang. Time discretization and markovian iteration for coupled fbsdes. *The Annals of Applied Probability*, 18(1):143–177, 2008.

[2] André Bento. Forward backward stochastic differential equations and pricing in emissions markets. Master's thesis, ISEG, University of Lisbon, 2022.

[3] Jean-Michel Bismut. Conjugate convex functions in optimal stochastic control. *Journal of Mathematical Analysis and Applications*, 44(2):384–404, 1973.

[4] Mireille Bossy and Denis Talay. Convergence rate for the approximation of the limit law of weakly interacting particles: application to the burgers equation. *The Annals of Applied Probability*, 6(3):818–861, 1996.

[5] Yann Brenier and Emmanuel Grenier. Sticky particles and scalar conservation laws. *SIAM Journal on Numerical Analysis*, 35(6):2317–2328, 1998.

[6] Alberto Bressan. Hyperbolic systems of conservation laws. *Revista Matemática Complutense*, 12(1):135–200, 1999.

[7] René Carmona and François Delarue. Singular fbsdes and scalar conservation laws driven by diffusion processes. *Probability Theory and Related Fields*, 157(1):333–388, 2013.

[8] René Carmona, François Delarue, Gilles-Edouard Espinosa, and Nizar Touzi. Singular forward-backward stochastic differential equations and emissions derivatives. *The Annals of Applied Probability*, 23(3):1086–1128, 2013.

[9] Jean-François Chassagneux, Hinesh Chotai, and Dan Crisan. Modelling multi-period carbon markets using singular forward backward sdes. *Mathematics of Operations Research*, 48(1):463–497, 2023.

[10] Jean-François Chassagneux, Hinesh Chotai, and Mirabelle Muûls. *A Forward-Backward SDEs Approach to Pricing in Carbon Markets*. Springer Briefs in Mathematics of Planet Earth, Springer, 2017.

[11] Jean-François Chassagneux and Mohan Yang. Numerical approximation of singular forward-backward sdes. *Journal of Computacional Physics*, 468, 2022.

[12] Jean-François Chassagneux and Mohan Yang. Convergence of particles and tree based scheme for singular fbsdes. *ESAIM: Mathematical Modelling and Numerical Analysis*, 59(3):1333–1363, 2025.

[13] Hinesh Chotai. *Forward-backward stochastic differential equations and applications to carbon emissions markets*. PhD thesis, Imperial College, 2019.

[14] Nicole El Karoui, Shige Peng, and Marie-Clarie Quenez. Backward stochastic differential equations in finance. *Mathematical Finance*, 7(1):1–71, 1997.

[15] Lawrence Craig Evans. A survey of entropy methods for partial differential equations. *Bulletin of the American Mathematical Society*, 41(4):409–438, 2004.

[16] Sam Howison and Daniel Schwarz. Risk-neutral pricing of financial instruments in emission markets: a structural approach. *SIAM review*, 57(1):95–127, 2015.

[17] Benjamin Jourdain and Julien Reygner. Optimal convergence rate of the multiple sticky particle approximation of one-dimensional diagonal hyperbolic systems with monotonic initial data. *Discrete and Continuous Dynamical Systems*, 36(9):4963–4996, 2016.

[18] Étienne Pardoux and Shige Peng. Adapted solution of a backward stochastic differential equation. *Systems  Control Letters*, 14(1):55–61, 1990.