Singular Forward-Backward Stochastic Differential Equations and Emissions Derivatives

René Carmona, François Delarue, Gilles-Edouard Espinosa, Nizar Touzi

SINGULAR FORWARD-BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS AND EMISSIONS DERIVATIVES

RENÉ CARMONA, FRANÇOIS DELARUE, GILLES-EDOUARD ESPINOSA, AND NIZAR TOUZI

ABSTRACT. We introduce two simple models of forward-backward stochastic differential equations with a singular terminal condition and we explain how and why they appear naturally as models for the valuation of CO_2 emission allowances. Single phase cap-and-trade schemes lead readily to terminal conditions given by indicator functions of the forward component, and using fine partial differential equations estimates, we show that the existence theory of these equations, as well as the properties of the candidates for solution, depend strongly upon the characteristics of the forward dynamics. Finally, we give a first order Taylor expansion and show how to numerically calibrate some of these models for the purpose of CO_2 option pricing.

1. INTRODUCTION

This paper is motivated by the mathematical analysis of the emissions markets, as implemented for example in the European Union (EU) Emissions Trading Scheme (ETS). These market mechanisms have been hailed by some as the most cost efficient way to control Green House Gas (GHG) emissions They have been criticized by others for being a tax in disguise and adding to the burden of industries covered by the regulation. Implementation of cap-and-trade schemes is not limited to the implementation of the Kyoto protocol. The successful US acid rain program is a case in point. However, a widespread lack of understanding of their properties, and misinformation campaigns by advocacy groups more interested in pushing their political agendas than using the results of objective scientific studies have muddied the water and add to the confusion. More mathematical studies are needed to increase the understanding of these market mechanisms and raise the level of awareness of their advantages as well as their shortcomings. This paper was prepared in this spirit.

In a first part, we introduce simple single-firm models inspired by the workings of the electricity markets (remember that electric power generation is responsible for most of the CO_2 emissions worldwide). Despite the specificity of some assumptions, our treatment is quite general in the sense that individual risk averse power producers choose their own utility functions. Moreover, the financial markets in which they can trade emission allowances are not assumed to be complete.

While market incompleteness prevents us from identifying the optimal trading strategy of each producer, we show that, independently of the choice of the utility function, the optimal production or abatement strategy is what we expect by proving mathematically, and in full generality (i.e. without assuming completeness of the markets), a *folk theorem* in environmental economics: the equilibrium

---

Date: July 22, 2010.
2000 *Mathematics Subject Classification.* Primary .
Partially supported by NSF: DMS-0806591.

allowance price equals the marginal abatement cost, and market participants implement all the abatement measures whose costs are not greater than the cost of compliance (i.e. the equilibrium price of an allowance).

The next section puts together the economic activities of a large number of producers and search for the existence of an equilibrium price for the emissions allowances. Such a problem leads naturally to a *forward* stochastic differential equation (SDE) for the aggregate emissions in the economy, and a *backward* stochastic differential equation (BSDE) for the allowance price. However, these equations are "coupled" since a nonlinear function of the *price of carbon* (i.e. the price of an emission allowance) appears in the forward equation giving the dynamics of the aggregate emissions. This feedback of the emission price in the dynamics of the emissions is quite natural. For the purpose of option pricing, this approach was described in[4] where it was called *detailed risk neutral approach*.

Forward backward stochastic differential equations (FBSDEs) of the type considered in this section have been studied for a long time. See for example [12], or [16]. However, the FBSDEs we need to consider for the purpose of emission prices have an unusual pecularity: the terminal condition of the backward equation is given by a discontinuous function of the terminal value of the state driven by the forward equation. We use our first model to prove that this lack of continuity is not an issue when the forward dynamics are *strongly elliptic*, in other words when the volatility of the forward SDE is bounded from below. However, using our second equilibrium model, we also show that when the forward dynamics are degenerate (even if they are hypoelliptic), discontinuities in the terminal condition and lack of uniform ellipticity in the forward dynamics can conspire to produce point masses in the terminal distribution of the forward component, at the locations of the discontinuities. This implies that the terminal value of the backward component is not given by a deterministic function of the forward component, for the forward scenarios ending at the locations of jumps in the terminal condition, and justifies relaxing the definition of a solution of the FBSDE.

Even though we only present a detailed proof for a very specific model for the sake of definiteness, we believe that our result is representative of a large class of models. Since from the point of view of the definition of "aggregate emissions", the degeneracy of the forward dynamics is expected, and this seemingly pathological result should not be overlooked. Indeed, it sheds new light on an absolute continuity assumption made repeatedly in equilibrium analyses, even in discrete time models. See for example [3] and [2]. This assumption was regarded as an annoying technicality, but in the light of the results of this paper, it looks more intrinsic to these types of models. In any case, it fully justifies the need to relax the definition of a solution of a FBSDE when the terminal condition of the forward part jumps.

A vibrant market for options written on allowance futures/forward contracts has recently developed and increased in liquidity. See for example [4] for details on these markets. Reduced formed models have been proposed to price these options. See [4] or [5]. Several attempts have been made at matching the smile (or lack thereof) contained in the quotes published daily by the exchanges. Section 5 develops the technology needed to price these options in the context of the equilibrium framework developed in the present paper. We identify the option prices in terms of solutions of nonlinear partial differential equations and we prove when the dynamics of the aggregate emissions are given by a geometric Brownian motion, a Taylor expansion formula when the nonlinear abatement feedback is small. We derive an explicit integral form for the first order Taylor expansion coefficient which can easily be computed by Monte Carlo methods. We believe that the present paper is the first rigorous attempt to include the nonlinear feedback term in the dynamics of aggregate emissions for the purpose of emissions option pricing.

The final section 6 illustrates numerically how the option prices computed from our equilibrium model differ from the *linear* prices computed in [5],[18] and [4]. Furthermore, we show how the first order Taylor approximation result of Section 5 can be used to compute numerically option prices and efficiently fit the implied volatility smile present in recent option price quotes.

Acknowledgements: We would like to thank two anonymous referees for pointing out inconsistencies in the original proofs of Theorem 1 and Proposition 3, and for comments which led to improvements in the presentation of the results of the paper.

2. TWO SIMPLE MODELS OF GREEN HOUSE GAS EMISSION CONTROL

We first describe the optimization problem of a single power producer facing a carbon cap-and-trade regulation. We assume that this producer is a small player in the market in the sense that his actions have no impact on prices and that a liquid market for pollution permits exists. In particular, we assume that the price of an allowance is given exogenously, and we use the notation $Y = (Y_t)_{0 \le t \le T}$ for the (stochastic) time evolution of the price of an emission allowance. For the sake of simplicity we assume that $[0, T]$ is a single phase of the regulation and that no banking or borrowing of the certificates is possible at the end of the phase. For illustration purposes, we analyze two simple models. Strangely enough, the first steps of these analyses, namely the identifications of the optimal abatement and production strategies, do not require the full force of the sophisticated techniques of optimal stochastic control.

2.1. **Modeling First the Emissions Dynamics.** We assume that the source of randomness in the model is given by $W = (W_t)_{0 \le t \le T}$, a (possibly infinite) sequence of independent one-dimensional Wiener processes $W^j = (W^j_t)_{0 \le t \le T}$. In other words, $W_t = (W^0_t, W^1_t, \cdots, W^i_t, \cdots)$ for each fixed $t \in [0, T]$. All these Wiener processes are assumed to be defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and we denote by $\mathbb{F} = \{\mathcal{F}_t, t \ge 0\}$ the Brownian filtration they generate. Here, $T > 0$ is a fixed time horizon representing the end of the regulation period.

We will eventually extend the model to include N firms, but for the time being, we consider only the problem of one single firm whose production of electricity generates emissions of carbon dioxyde, and we denote by E_t the cumulative emissions up to time t of the firm. We also denote by $\tilde{E}_t$ the perception at time t (for example the conditional expectation) of what the total cumulative emission E_T will be at the end of the time horizon. Clearly, E and $\tilde{E}$ can be different stochastic processes, but they have the same terminal values at time T, i.e. $E_T = \tilde{E}_T$. We will assume that dynamics of the proxy $\tilde{E}$ for the cumulative emissions of the firm are given by an Itô process of the form:

$$\tilde{E}_t = \int_0^t (b_s - \xi_s) ds + \int_0^t \sigma_s dW_s, \tag{1}$$

(so $\tilde{E}_0 = 0$) where b represents the (conditional) expectation of what the rate of emission would be in a world without carbon regulation, in other words in what is usually called *Business As Usual*, while ξ is the instantaneous rate of abatement chosen by the firm. In mathematical terms, ξ represents the control on emission reduction implemented by the firm. Clearly, in such a model, the firm only acts on the drift of its *perceived* emissions. For the sake of simplicity we assume that the processes b and

σ are adapted and bounded. Because of the vector nature of the Brownian motion W, the volatility process σ is in fact a sequence of scalar volatility processes $(\sigma^j)_{j\geq 0}$. For the purpose of this section, we could use one single scalar Wiener process and one single scalar volatility process as long as we allow the filtration $\mathbb{F}$ to be larger than the filtration generated by this single Wiener process. This fact will be needed when we study a model with different firms.

Continuing on with the description of the model, we assume that the abatement decision is based on a cost function $c : \mathbb{R} \to \mathbb{R}$ which is assumed to be continuously differentiable (C^1 in notation), strictly convex and satisfy Inada-like conditions:

$$c'(-\infty) = -\infty \qquad \text{and} \qquad c'(+\infty) = +\infty. \tag{2}$$

Note that $(c')^{-1}$ exists because of the assumption of strict convexity. Since $c(x)$ can be interpreted as the cost to the firm for an abatement rate of level x, without any loss of generality we will also assume $c(0) = \min c = 0$. Notice that (2) implies that $\lim_{x\to\pm\infty} c(x) = +\infty$.

Remark 1. *A typical example of abatement cost function is given by the quadratic cost function $c(x) = \alpha x^2$ for some $\alpha > 0$ used in* [18]*, or more generally the power cost function $c(x) = \alpha|x|^{1+\beta}$ for some $\alpha > 0$ and $\beta > 0$.*

The firm controls its destiny by choosing its own abatement schedule ξ as well as the quantity θ of pollution permits it holds through trading in the allowance market. For these controls to be admissible, ξ and θ need only to be progressively measurable processes satisfying the integrability condition

$$\mathbb{E}\int_0^T [\theta_t^2 + \xi_t^2]dt < \infty. \tag{3}$$

We denote by $\mathcal{A}$ the set of admissible controls (ξ, θ). Given its initial wealth x, the terminal wealth X_T of the firm is given by:

$$X_T = X_T^{\xi,\theta} = x + \int_0^T \theta_t dY_t - \int_0^T c(\xi_t)dt - E_T Y_T. \tag{4}$$

Recall that we use the notation Y_t for the price of an emission allowance at time t. Recall also that at this stage, we are not interested in the existence or the formation of this price. We merely assume the existence of a liquid and frictionless market for emission allowances, and that Y_t is the price at which each firm can buy or sell one allowance at time t. The risk preferences of the firm are given by a utility function $U : \mathbb{R} \to \mathbb{R}$, which is assumed to be C^1, increasing, strictly concave and satisfying the Inada conditions:

$$(U)'(-\infty) = +\infty \qquad \text{and} \qquad (U)'(+\infty) = 0. \tag{5}$$

The optimization problem of the firm can be written as the computation of:

$$V(x) = \sup_{(\xi,\theta)\in\mathcal{A}} \mathbb{E}U(X_T^{\xi,\theta}) \tag{6}$$

where $\mathbb{E}$ denotes the expectation under the *historical* measure $\mathbb{P}$, and $\mathcal{A}$ is the set of abatement and trading strategies (ξ, θ) admissible to the firm. The following simple result holds.

Proposition 1. *The optimal abatement strategy of the firm is given by:*

$$\xi_t^* = [c']^{-1}(Y_t).$$

Remark 2. *Notice that the optimal abatement schedule is independent of the utility function. The beauty of this simple result is its powerful intuitive meaning: given a price Y_t for an emission allowance, the firm implements all the abatement measures which make sense economically, namely all those costing less than the current market price of one allowance (i.e. one unit of emission).*

Proof. If we rewrite the last term in the expression (4) of the terminal wealth by replacing E_T by $\tilde{E}_T$, a simple integration by parts gives:

$$\begin{aligned} E_T Y_T &= Y_T \left(\int_0^T b_t dt + \int_0^T \sigma_t dW_t \right) - Y_T \int_0^T \xi_t dt \\ &= Y_T \left(\int_0^T b_t dt + \int_0^T \sigma_t dW_t \right) - \int_0^T Y_t \xi_t dt - \int_0^T \left(\int_0^t \xi_s ds \right) dY_t \end{aligned}$$

so that $X_T = A_T^{\tilde{\theta}} + B_T^{\xi}$ with

$$A_T^{\tilde{\theta}} = \int_0^T \tilde{\theta}_t dY_t - Y_T \left(\int_0^T b_t dt + \int_0^T \sigma_t dW_t \right)$$

where the modified control $\tilde{\theta}$ is defined by $\tilde{\theta}_t = \theta_t + \int_0^t \xi_s ds$, and

$$B_T^{\xi} = x - \int_0^T [c(\xi_t) - Y_t \xi_t] dt.$$

Notice that B^{ξ} depends only upon ξ without depending upon $\tilde{\theta}$ while $A^{\tilde{\theta}}$ depends only upon $\tilde{\theta}$ without depending upon ξ. The set $\mathcal{A}$ of admissible controls is equivalently described by varying the couples (θ, ξ) or $(\tilde{\theta}, \xi)$, so when computing the maximum

$$\sup_{(\theta,\xi)\in\mathcal{A}} \mathbb{E}\{U(X_T)\} = \sup_{(\tilde{\theta},\xi)\in\mathcal{A}} \mathbb{E}\{U(A_T^{\tilde{\theta}} + B_T^{\xi})\}$$

one can perform the optimizations over $\tilde{\theta}$ and ξ separately, for example by fixing $\tilde{\theta}$ and optimizing with respect to ξ before maximizing the result with respect to $\tilde{\theta}$. The proof is complete once we notice that U is increasing and that for each $t \in [0, T]$ and each $\omega \in \Omega$, the quantity B_T^{ξ} is maximized by the choice $\xi_t^* = (c')^{-1}(Y_t)$. □

Remark 3. *The above result argues neither existence nor uniqueness of an optimal admissible set (ξ^*, θ^*) of controls. We believe that once the optimal rate of abatement ξ^* is implemented, the optimal investment strategy θ^* should hedge the financial risk created by the implementation of the abatement strategy. This fact can be proved using the classical tools of portfolio optimization in the case of complete market models. Indeed, if we introduce the convex dual $\tilde{U}$ of U defined by:*

$$\tilde{U}(y) := \sup_x \{U(x) - xy\}$$

and the function I by $I = (U')^{-1}$ so that $\tilde{U}(y) = U \circ I(y) - yI(y)$ and if we denote by $\mathbb{E}$ and $\mathbb{E}^{\mathbb{Q}}$ respectively the expectations with respect to $\mathbb{P}$ and the unique equivalent measure $\mathbb{Q}$ under which Y is a martingale (we write Z_t its volatility given by the martingale representation theorem), then from *the a.s. inequality*

$$U(X_T^{\xi,\theta}) - y\frac{d\mathbb{Q}}{d\mathbb{P}}X_T^{\xi,\theta} \leq U \circ I\left(y\frac{d\mathbb{Q}}{d\mathbb{P}}\right) - y\frac{d\mathbb{Q}}{d\mathbb{P}}I\left(y\frac{d\mathbb{Q}}{d\mathbb{P}}\right),$$

valid for any admissible (ξ, θ)*, and* $y \in \mathbb{R}$*, we get*

$$\mathbb{E}U(X_T^{\xi,\theta}) \leq \mathbb{E}U \circ I\left(y\frac{d\mathbb{Q}}{d\mathbb{P}}\right) + y\mathbb{E}^{\mathbb{Q}}\left[X_T^{\xi,\theta} - I\left(y\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right]$$

after taking expectations under $\mathbb{P}$*. Computing* $\mathbb{E}^{\mathbb{Q}}X_T^{\xi,\theta}$ *by integration by parts we get:*

$$\mathbb{E}U(X_T^{\xi,\theta}) \leq \mathbb{E}U \circ I\left(y\frac{d\mathbb{Q}}{d\mathbb{P}}\right) + y\left[x - \mathbb{E}^{\mathbb{Q}}\int_0^T [c \circ (c')^{-1}(Y_t) + Y_t(b_t - (c')^{-1}(Y_t)) + \sigma_t Z_t dt] \right.$$
$$\left. -\mathbb{E}^{\mathbb{Q}} I\left(y\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right]$$

if we use the optimal rate of abatement. So if we choose $y = \hat{y} \in \mathbb{R}$ *as the unique solution of:*

$$\mathbb{E}^{\mathbb{Q}}I\left(\hat{y}\frac{d\mathbb{Q}}{d\mathbb{P}}\right) = x - \mathbb{E}^{\mathbb{Q}}\int_0^T c \circ (c')^{-1}(Y_t) + Y_t(b_t - (c')^{-1}(Y_t)) + \sigma_t Z_t dt.$$

it follows that

$$\mathbb{E}^{\mathbb{Q}}X_T^{\hat{\xi},\theta} = \mathbb{E}^{\mathbb{Q}}I\left(\hat{y}\frac{d\mathbb{Q}}{d\mathbb{P}}\right),$$

and finally, if the market is complete, the claim $I\left(\hat{y}^j\frac{d\mathbb{Q}}{d\mathbb{P}}\right)$ *is attainable by a certain* θ^**. This completes the proof.*

2.2. **Modeling the Electricity Price First.** We consider a second model for which again, part of the global stochastic optimization problem reduces to a mere path-by-path optimization. As before, the model is simplistic, especially in the case of a single firm in a regulatory environment with a liquid frictionless market for emission allowances. However, this model will become very informative later on when we consider N firms interacting on the same market, and we try to construct the allowance price Y_t by solving a Forward-Backward Stochastic Differential Equation (FBSDE). The model concerns an economy with one production good (say electricity) whose production is the source of a negative externality (say GHG emissions). Its price $(P_t)_{0\leq t\leq T}$ evolves according to the following Itô stochastic differential equation:

$$dP_t = \mu(P_t)dt + \sigma(P_t)dW_t \tag{7}$$

where the deterministic functions μ and σ are assumed to be C^1 with bounded derivatives. At each time $t \in [0, T]$, the firm chooses its instantaneous rate of production q_t and its production costs are $c(q_t)$ where c is a function $c : \mathbb{R}_+ \hookrightarrow \mathbb{R}$ which is assumed to be C^1 and strictly convex. With these notations, the profits and losses from the production at the end of the period $[0, T]$, are given by the integral:

$$\int_0^T [P_t q_t - c(q_t)]dt.$$

The emission regulation mandates that at the end of the period $[0, T]$, the cumulative emissions of each firm be measured, and that one emission permit be redeemed per unit of emission. As before, we denote by $(Y_t)_{0\le t\le T}$ the process giving the price of one emission allowance. For the sake of simplicity, we assume that the cumulative emissions E_t up to time t are proportional to the production in the sense that $E_t = \epsilon Q_t$ where the positive number ϵ represents the rate of emission of the production technology used by the firm, and Q_t denotes the cumulative production up to and including time t:

$$Q_t = \int_0^t q_s ds.$$

At the end of the time horizon, the cost incurred by the firm because of the regulation is given by $E_T Y_T = \epsilon Q_T Y_T$. The firm may purchase allowances: we denote by θ_t the amount of allowances held by the firm at time t. Under these conditions, the terminal wealth of the firm is given by:

$$X_T = X_T^{q,\theta} = x + \int_0^T \theta_t dY_t + \int_0^T [P_t q_t - c(q_t)]dt - \epsilon Q_T Y_T \tag{8}$$

where as before, we used the notation x for the initial wealth of the firm. The first integral in the right hand side of the above equation gives the proceeds from trading in the allowance market, the next term gives the profits from the production and the sale of electricity, and the last term gives the costs of the emission regulation. We assume that the risk preferences of the firm are given by a utility function $U : \mathbb{R} \to \mathbb{R}$, which is assumed to be C^1, increasing, strictly concave and satisfying the Inada conditions (5) stated earlier. The optimization problem of the firm can be written as:

$$V(x) = \sup_{(q,\theta)\in\mathcal{A}} \mathbb{E}U(X_T^{q,\theta}) \tag{9}$$

where $\mathbb{E}$ denotes the expectation under the *historical* measure $\mathbb{P}$, and $\mathcal{A}$ is the set of admissible production and trading strategies (q, θ). As before, for these controls to be admissible, q and θ need only be adapted processes satisfying the integrability condition

$$\mathbb{E}\int_0^T [\theta_t^2 + q_t^2]dt < \infty. \tag{10}$$

Proposition 2. *The optimal production strategy of the firm is given by:*

$$q_t^* = (c')^{-1}(P_t - \epsilon Y_t).$$

Remark 4. *As before, the optimal production strategy q^* is independent of the risk aversion (i.e. the utility function) of the firm. The intuitive interpretation of this result is clear: once a firm observes both prices P_t and Y_t, it computes the price for which it can sell the good minus the price it will have to pay because of the emission regulation, and the firm uses this corrected price to choose its optimal rate of production in the usual way.*

Proof. A simple integration by part (notice that E_t is of bounded variations) gives:

$$Q_T Y_T = \int_0^T Y_t dQ_t + \int_0^T Q_t dY_t = \int_0^T Y_t q_t dt + \int_0^T Q_t dY_t, \tag{11}$$

so that $X_T = A_T^{\tilde\theta} + B_T^q$ with

$$A_T^{\tilde\theta} = \int_0^T \tilde\theta_t dY_t \qquad \text{with} \qquad \tilde\theta_t = \theta_t - \epsilon\int_0^t q_s ds$$

which depends only upon $\tilde{\theta}$ and

$$B_T^q = x + \int_0^T [(P_t - \epsilon Y_t)q_t - c(q_t)]dt,$$

which depends only upon q without depending upon $\tilde{\theta}$. Since the set $\mathcal{A}$ of admissible controls is equivalently described by varying the couples $(q, \tilde{\theta})$ or $(q, \tilde{\theta})$, when computing the maximum

$$\sup_{(q,\theta)\in\mathcal{A}} \mathbb{E}\{U(X_T)\} = \sup_{(\tilde{\theta},\xi)\in\mathcal{A}} \mathbb{E}\{U(A_T^{\tilde{\theta}} + B_T^q)\}$$

one can perform the optimizations over q and $\tilde{\theta}$ separately, for example by fixing $\tilde{\theta}$ and optimizing with respect to q before maximizing the result with respect to $\tilde{\theta}$. The proof is complete once we notice that U is increasing and that for each $t \in [0, T]$ and each $\omega \in \Omega$, the quantity B_T^q is maximized by the choice $q_t^* = (c')^{-1}(P_t - \epsilon Y_t)$. □

3. Allowance Equilibrium Price and a First Singular FBSDE

The goal of this section is to extend the first model introduced in section 2 to an economy with N firms, and solve for the allowance price.

3.1. **Switching to a Risk Neutral Framework.** As before, we assume that $Y = (Y_t)_{t\in[0,T]}$ is the price of one allowance in a one-compliance period cap-and-trade model, and that the market for allowances is frictionless and liquid. Y is a martingale for a measure $\mathbb{Q}$ equivalent to the historical measure $\mathbb{P}$. Because we are in a Brownian filtration,

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left[\int_0^T \alpha_t dW_t \; - \; \frac{1}{2}\int_0^T |\alpha_t|^2 dt\right]$$

for some sequence $\alpha = (\alpha_t)_{t\in[0,T]}$ of adapted processes. By Girsanov's theorem, the process $\tilde{W} = (\tilde{W}_t)_{t\in[0,T]}$ defined by

$$\tilde{W}_t = W_t - \int_0^t \alpha_s ds$$

is a Wiener process for $\mathbb{Q}$ so that equation (1) giving the dynamics of the perceived emissions of a firm now reads:

$$d\tilde{E}_t = (\tilde{b}_t - \xi_t)dt \; + \sigma_t d\tilde{W}_t$$

under $\mathbb{Q}$, where the new drift $\tilde{b}$ is defined by $\tilde{b}_t = b_t + \sigma_t\alpha_t$ for all $t \in [0, T]$.

3.2. **Market Model with N Firms.** We now consider an economy comprising N firms labelled by $\{1, \cdots, N\}$, and we work in the risk neutral framework for allowance trading discussed above. When a specific quantity such as cost function, utility, cumulative emission, trading strategy, ... depends upon a firm, we use a superscript ^{i} to emphasize the dependence upon the i-th firm. So in equilibrium (i.e. whenever each firm implements its optimal abatement strategy), for each firm $i \in \{1, \cdots, N\}$ we have

$$d\tilde{E}_t^i = \{\tilde{b}_t^i - [(c^i)']^{-1}(Y_t)\}dt + \sigma_t^i d\tilde{W}_t$$

with $\tilde{E}_0^i = 0$. Consequently, the aggregate perceived emission $\tilde{E}$ defined by

$$\tilde{E} = \sum_{i=1}^{N} \tilde{E}_t^i$$

satisfies

$$d\tilde{E}_t = \big(b_t - f(Y_t)\big)dt + \sigma_t d\tilde{W}_t,$$

where

$$b_t = \sum_{i=1}^{N} \tilde{b}_t^i, \qquad \sigma_t = \sum_{i=1}^{N} \sigma_t^i \qquad \text{and} \qquad f(x) = \sum_{i=1}^{N} [(c^i)']^{-1}(x).$$

Again, since we are in a Brownian filtration, it follows from the martingale representation theorem that there exists a progressively measurable process $Z = (Z_t)_{t\in[0,T]}$ such that

$$dY_t = Z_t d\tilde{W}_t \quad \text{and} \quad \mathbb{E}\Big[\int_0^T |Z_t|^2 dt\Big] < \infty.$$

Furthermore, we assume the existence of deterministic continuous functions $[0,T]\times\mathbb{R} \ni (t,e) \hookrightarrow b(t,e)$ and $[0,T] \ni t \hookrightarrow \sigma(t)$ such that $b_t = b(t,E_t)$ and $\sigma_t = \sigma(t)$, for all $t \in [0,T]$, $\mathbb{Q}$-a.s.

Consequently, the processes $\tilde{E}$, Y, and Z satisfy a system of Forward Backward Stochastic Differential Equations (FBSDEs for short) which we restate for the sake of later reference:

$$\begin{cases} d\tilde{E}_t = \big(b(t,\tilde{E}_t) - f(Y_t)\big)dt + \sigma(t)d\tilde{W}_t, & \tilde{E}_0 = 0 \\ dY_t = Z_t dW_t, \quad Y_T = \lambda \mathbf{1}_{[\Lambda,+\infty)}(\tilde{E}_T). \end{cases} \tag{12}$$

Notice that since all the cost functions c^i are strictly convex, f is strictly increasing. We shall make the following additional assumptions:

$$b(t,e) \text{ is Lipschitz in } e \text{ uniformly in } t, \tag{13}$$

$$\sigma \in \mathbb{L}^2([0,T],dt) \text{ and } \inf_{[0,T]} \sigma^2 > 0, \tag{14}$$

$$f \text{ is Lipschitz continuous (and strictly increasing).} \tag{15}$$

We denote by $\mathbb{H}^0$ the collection of all progressively measurable processes on $[0,T]\times\mathbb{R}$, and we introduce the subsets:

$$\mathbb{H}^2 := \Big\{Z \in \mathbb{H}^0;\ \mathbb{E}\int_0^T |Z_s|^2 ds < \infty\Big\} \quad \text{and} \quad \mathbb{S}^2 := \Big\{Y \in \mathbb{H}^0;\ \mathbb{E}[\sup_{0\le t\le T} |Y_s|^2] < \infty\Big\}.$$

3.3. **Solving the Singular Equilibrium FBSDE.** The purpose of this subsection is to prove existence and uniqueness of a solution to FBSDE (12).

Theorem 1. *Under assumptions* (13) *to* (15)*, for any* $\lambda > 0$ *and* $\Lambda \in \mathbb{R}$*, FBSDE* (12) *admits a unique solution* $(\tilde{E}, Y, Z) \in \mathbb{S}^2 \times \mathbb{S}^2 \times \mathbb{H}^2$*. Moreover,* $\tilde{E}_t$ *is non-increasing with respect to* λ *and non-decreasing with respect to* Λ.

Proof. For any function φ, we write FBSDE(φ) for the FBSDE (12) when the indicator function appearing in the terminal condition in the backward component of (12) is replaced by φ.

(i) We first prove uniqueness. Let $(\tilde{E}, Y, Z)$ and $(\tilde{E}', Y', Z')$ be two solutions of FBSDE (12). Clearly it is sufficient to prove that $Y = Y'$. Let us set:

$$\delta E_t := \tilde{E}_t - \tilde{E}'_t, \quad \delta Y_t := Y_t - Y'_t, \quad \delta Z_t := Z_t - Z'_t, \quad \text{and} \quad \beta_t := \frac{b(t,\tilde{E}_t) - b(t,\tilde{E}'_t)}{\delta \tilde{E}_t} \mathbf{1}_{\{\delta \tilde{E}_t \neq 0\}}.$$

Notice that $(\beta_t)_{0 \le t \le T}$ is a bounded process. By direct calculation, we see that

$$d(B_t \delta E_t \delta Y_t) = -B_t \delta Y_t \big(f(Y_t) - f(Y'_t)\big) dt + B_t \delta E_t \delta Z_t d\tilde{W}_t \quad \text{where} \quad B_t := e^{-\int_0^t \beta_s ds}.$$

Since $\delta E_0 = 0$ and $\delta E_T \delta Y_T = (\tilde{E}_T - \tilde{E}'_T)\big(g(\tilde{E}_T) - g(\tilde{E}'_T)\big) \ge 0$, because g is nondecreasing, this implies that

$$\mathbb{E}\Big[\int_0^T B_t \delta Y_t \big(f(Y_t) - f(Y'_t)\big) dt\Big] \ \le\ 0.$$

Since $B_t > 0$ and f is (strictly) increasing, this implies that $\delta Y = 0$ $dt \otimes d\mathbb{Q}-$a.e. and therefore $Y = Y'$ by continuity.

(ii) We next prove existence. Let $(g^n)_{n \ge 1}$ be an increasing sequence of smooth non-decreasing functions with $g^n \in [0,1]$ and $g^n \longrightarrow g := \mathbf{1}_{(\Lambda,\infty)}$.

(ii-1) Since g^n is smooth, the existence of a solution $(\tilde{E}^n, Y^n, Z^n) \in \mathbb{S}^2 \times \mathbb{S}^2 \times \mathbb{H}^2$ of FBSDE(g^n) follows from classical results (see [12] or [6]). Moreover $Y^n_t = u^n(t, \tilde{E}^n_t)$ for some deterministic function u^n that takes values in $[0, \lambda]$. Since the sequence $(g^n)_{n \ge 1}$ is increasing, it follows from Theorem 8.1 of [13] that the sequence of functions $(u^n)_{n \ge 1}$ is non-decreasing. We may then define:

$$u(t,e) := \lim_{n \to \infty} \uparrow u^n(t,e), \ t \in [0,T], \ e \in \mathbb{R}_+.$$

From Theorem 8.1 of [13], it also follows that, for each fixed n, u^n is non-decreasing in λ and non-increasing in Λ. Therefore the same monotonicity properties hold for u.

By Theorem 1.3 in [7], for any $\varepsilon > 0$, u^n is locally Lipschitz in e uniformly in t on $[0, T-\varepsilon]$, uniformly in n. Then, the limit function u is locally Lipschitz in e, uniformly in t on any $[0, T-\varepsilon]$.

Notice that the process $\tilde{E}^n$ solves the (forward) stochastic differential equation

$$d\tilde{E}^n_t = \big(b(t, \tilde{E}^n_t) - f \circ u^n(t, \tilde{E}^n_t)\big) dt + \sigma(t) d\tilde{W}_t,$$

where here and in the following, we use the notation $f \circ u$ for the composition of the functions f and u. Since f is increasing and the sequence $(u^n)_{n \ge 1}$ is non-decreasing, it follows from the comparison theorem for (forward) stochastic differential equations that the sequence of processes $(\tilde{E}^n)_{n \ge 1}$ is non-increasing. We may then define:

$$\hat{E}_t := \lim_{n \to \infty} \downarrow \tilde{E}^n_t \qquad \text{for all } t \in [0,T).$$

Notice also that, for the same reason, for each n, $\tilde{E}^n$ is non-increasing in λ and non-decreasing in Λ, so that the same holds true for $\hat{E}$.

(ii-2) To identify the dynamics of the limiting process $\hat{E}$, we introduce the process $\tilde{E}$ defined on $[0,T)$ as the unique strong solution of the stochastic differential equation

$$d\tilde{E}_t = (b - f \circ u)(t, \tilde{E}_t) dt + \sigma(t) d\tilde{W}_t.$$

The fact that the function u is bounded and locally Lipschitz, together with our assumptions on b and f guarantee the existence of such a strong solution. Moreover, the Girsanov change of measure argument given below implies that this strong solution does not explode. Since $u^n(t,e) = \lambda\mathbb{E}[g^n(\tilde{E}^n_T)|\tilde{E}^n_t = e]$, and g^n is non-decreasing, $u^n(t,.)$ is non-decreasing, for any n, and the same holds for u. We then use the fact that $\tilde{E}^n \geq \tilde{E}$ (from the classical comparison result for SDEs) together with the increase of $u^n(t,.)$ to compute, using Itô's formula, that:

$$
\begin{aligned}
(\tilde{E}^n_t - \tilde{E}_t)^2 &= 2\int_0^t (\tilde{E}^n_s - \tilde{E}_s)\big((b - f\circ u^n)(s,\tilde{E}^n_s) - (b - f\circ u)(s,\tilde{E}_s)\big)ds \\
&\leq C\int_0^t \big|\tilde{E}^n_s - \tilde{E}_s\big|^2 ds + 2\int_0^t (\tilde{E}^n_s - \tilde{E}_s)(f\circ u - f\circ u^n)(s,\tilde{E}_s)ds \\
&\leq (C+1)\int_0^t \big|\tilde{E}^n_s - \tilde{E}_s\big|^2 ds + \int_0^t \big|(f\circ u - f\circ u^n)(s,\tilde{E}_s)\big|^2 ds
\end{aligned}
$$

by the Lipschitz property of the coefficient b. Then

$$
\sup_{u\leq t}(\tilde{E}^n_u - \tilde{E}_u)^2 \leq (C+1)\int_0^t \sup_{u\leq s}\big|\tilde{E}^n_u - \tilde{E}_u\big|^2 ds + \int_0^t \big|(f\circ u - f\circ u^n)(s,\tilde{E}_s)\big|^2 ds,
$$

and taking expectations, we see that

$$
\mathbb{E}\Big[\sup_{u\leq t}(\tilde{E}^n_u - \tilde{E}_u)^2\Big] \leq (C+1)\int_0^t \mathbb{E}\Big[\sup_{u\leq s}\big(\tilde{E}^n_u - \tilde{E}_u\big)^2\Big]ds + \varepsilon^n
$$

where $\varepsilon^n := \mathbb{E}\big[\int_0^T \big|(f\circ u - f\circ u^n)(s,\tilde{E}_s)\big|^2 ds\big] \longrightarrow 0$, by the dominated convergence theorem. Therefore it follows from the Gronwall's inequality that $\tilde{E}^n \longrightarrow \tilde{E}$ in $\mathbb{S}^2$, and as a consequence, $\hat{E} = \tilde{E}$.

(ii-3) Since f is bounded on $[0,\Lambda]$, we may introduce an equivalent measure $\tilde{\mathbb{Q}} \sim \mathbb{Q}$ under which the process $B_t := \tilde{W}_t - \sigma^{-1} f\circ u(t,\tilde{E}_t)$, $t\in[0,T]$ is a Brownian motion. Then $\tilde{E}$ solves the stochastic differential equation

$$
d\tilde{E}_t = b(t,\tilde{E}_t)dt + \sigma(t)dB_t. \tag{16}
$$

By (14), the function $\sigma(.)^2$ is bounded away from zero on $[0,T]$. Then, the law of $\tilde{E}_T$ under $\tilde{\mathbb{Q}}$ has a density with respect to the Lebesgue measure, and the same holds true under the equivalent measure $\mathbb{Q}$. Consequently, it follows from (16) that

$$
\mathbb{Q}[\tilde{E}_T = \Lambda] = 0, \tag{17}
$$

which implies that we can use $g = \mathbb{1}_{(\Lambda,\infty)}$ instead of $\mathbb{1}_{[\Lambda,\infty)}$ in (12). Moreover, we also have:

$$
\lim_{n\to\infty}\mathbb{Q}[\tilde{E}^n_T > \Lambda|\mathcal{F}_t] = \mathbb{Q}[\tilde{E}_T > \Lambda|\mathcal{F}_t] \tag{18}
$$

for each $t < T$. The fact that $g^n \leq g$ implies:

$$
Y^n_t = \lambda\mathbb{E}_t[g^n(\tilde{E}^n_T)] \leq \lambda\mathbb{E}_t[g(\tilde{E}^n_T)] \longrightarrow \lambda\mathbb{E}_t[g(\tilde{E}_T)]
$$

as $n\to\infty$ by (18). On the other hand, since $\tilde{E}^n_T \geq \tilde{E}_T$, it follows from the non-decrease of g_n, the dominated convergence theorem, and (18) that

$$
Y^n_t = \mathbb{E}_t[g^n(\tilde{E}^n_T)] \geq \mathbb{E}_t[g^n(\tilde{E}_T)] \longrightarrow \mathbb{E}_t[g(\tilde{E}_T)].
$$

Hence $Y_t^n \longrightarrow Y_t := \lambda \mathbb{E}_t[g(\tilde{E}_T)]$. Now, let $Z \in \mathbb{H}^2$ be such that

$$Y_t = \lambda g(\tilde{E}_T) - \int_t^T Z_s dW_s, \qquad t \in [0,T].$$

Notice that Y takes values in $[0, \lambda]$, and therefore $Y \in \mathbb{S}^2$. Similarly, using the increase and the decrease of the sequences $(u^n)_n$ and $(E^n)_n$ respectively, together with the increase of $u^n(t,.)$ and $u(t,.)$, we see that for $t \in [0,T]$:

$$u(t, \tilde{E}_t) = \lim_{n\to\infty} u^n(t, \tilde{E}_t) \le \liminf_{n\to\infty} u^n(t, \tilde{E}_t^n) \le \limsup_{n\to\infty} u^n(t, \tilde{E}_t^n) \le \lim_{n\to\infty} u(t, \tilde{E}_t^n) = u(t, \tilde{E}_t).$$

Since $Y_t^n = u^n(t, \tilde{E}_t^n)$, this shows that $Y_t = u(t, \tilde{E}_t)$ on $[0,T)$, and the proof is complete. □

Impact on the model for emission control. As expected, the previous result implies that the tougher the regulation (i.e. the larger λ and/or the smaller Λ), the higher the emission reductions (the lower $\tilde{E}_t$). In particular, in the absence of regulation which corresponds to $\lambda = 0$, the aggregate level of emissions is at its highest.

4. Enlightening Example of a Singular FBSDE

We saw in the previous section that when the forward dynamics are non-degenerate, the terminal condition of the backward equation can be a discontinuous function of the terminal value of the forward component without threatening existence or uniqueness of a solution to the FBSDE. In this section, we show that this is not the case when the forward dynamics are degenerate, even if they are hypoelliptic and the solution of the forward equation has a density. We explained in the introduction why this seemingly pathological mathematical property should not come as a surprise in the context of equilibrium models for cap-and-trade schemes.

Motivated by the second model given in subsection 2.2, we consider the FBSDE:

$$\left\{ \begin{array}{l} dP_t = dW_t, \\ dE_t = \big(P_t - Y_t\big) dt, \\ dY_t = Z_t dW_t, \quad 0 \le t \le T, \end{array} \right. \tag{19}$$

with the terminal condition

$$Y_T = \mathbf{1}_{[\Lambda,\infty)}(E_T), \tag{20}$$

for some real number Λ. Here, $(W_t)_{t\in[0,T]}$ is a one-dimensional Wiener process. This unrealistic model corresponds to quadratic costs of production, and choosing appropriate units for the penalty λ and the emission rate ϵ to be 1. Our interest in this model is the outcome of its mathematical analysis, not its realism! We prove the following unexpected result.

Theorem 2. *Given $(p, e) \in \mathbb{R}^2$, there exists a unique progressively measurable triple $(P_t, E_t, Y_t)_{0\le t\le T}$ satisfying (19) together with the initial conditions $P_0 = p$ and $E_0 = e$, and*

$$\mathbf{1}_{(\Lambda,\infty)}(E_T) \le Y_T \le \mathbf{1}_{[\Lambda,\infty)}(E_T). \tag{21}$$

Moreover, the marginal distribution of E_t is absolutely continuous with respect to the Lebesgue measure for any $0 \le t < T$, but has a Dirac mass at Λ when $t = T$. In other words:

$$\mathbb{P}\{E_T = \Lambda\} > 0.$$

In particular, $(P_t, E_t, Y_t)_{0\le t\le T}$ may not satisfy the terminal condition $\mathbb{P}\{Y_T = \mathbf{1}_{[\Lambda,\infty)}(E_T)\} = 1$. However, the weaker form (21) of terminal condition is sufficient to guarantee uniqueness.

Before we engage in the technicalities of the proof we notice that the transformation

$$(P_t, E_t)_{0\le t\le T} \hookrightarrow (\bar{E}_t = E_t + (T-t)P_t)_{0\le t\le T} \tag{22}$$

maps the original FBSDE (19) into the simpler one

$$\begin{cases} d\bar{E}_t = -Y_t dt + (T-t)dW_t, \\ dY_t = Z_t dW_t, \end{cases} \tag{23}$$

with the same terminal condition $Y_T = \mathbf{1}_{[\Lambda,\infty)}(\bar{E}_T)$. Moreover, the dynamics of $(E_t)_{0\le t\le T}$ can be recovered from those of $(\bar{E}_t)_{0\le t\le T}$ since $(P_t)_{0\le t\le T}$ in (19) is purely autonomous. In particular, except for the proof of the absolute continuity of E_t for $t < T$, we restrict our analysis to the proof of Theorem 2, for $\bar{E}$ solution of (23) since E and $\bar{E}$ have the same terminal values at time T.

We emphasize that system (23) is doubly singular at maturity time T: the diffusion coefficient of the forward equation vanishes as t tends to T and the boundary condition of the backward equation is discontinuous at point Λ. Together, both singularities make the emission process accumulate a non-zero mass at time T and at point Λ. This phenomenon must be seen as a *stochastic residual* of the shock wave observed in the inviscid Burgers equation

$$\partial_t v(t,e) - v(t,e)\partial_e v(t,e) = 0, \quad t \in [0,T),\ e \in \mathbb{R}, \tag{24}$$

with $v(T,e) = \mathbf{1}_{[\Lambda,+\infty)}(e)$ as boundary condition. (As explained below, equation (24) is the first-order version of the second-order equation associated with (23).)

Indeed, it is well-known that the characteristics of (24) may meet at time T and at point Λ. By analogy, the trajectories of the forward process in (23) may hit Λ at time T with a non-zero probability, then producing a Dirac mass. In other words, the shock phenomenon behaves like a trap into which the process $(E_t)_{0\le t\le T}$ (or equivalently the process $(\bar{E}_t)_{0\le t\le T}$) may fall with a non-zero probability. It is then well-understood that the noise plugged into the forward process $(\bar{E}_t)_{0\le t\le T}$ may help it to escape the trap. For example, we saw in Section 3 that the emission process did not see the trap in the uniformly elliptic setting. In the current framework, the diffusion coefficient vanishes in a linear way as time tends to the maturity: it decays too fast to prevent almost every realization of the process from falling into the trap.

As before, we prove existence of a solution to (23) by first smoothing the singularity in the terminal condition, solving the problem for a smooth terminal condition, and obtaining a solution to the original problem by a limiting argument. However, in order to prove the existence of a limit, we will use PDE a priori estimates and compactness arguments instead of comparison and monotonicity arguments. We call *mollified equation* the system (23) with a terminal condition

$$Y_T = \phi(\bar{E}_T), \tag{25}$$

given by a Lipschitz non-decreasing function ϕ from $\mathbb{R}$ to $[0,1]$ which we view as an approximation of the indicator function appearing in the terminal condition (20).

4.1. Lipschitz Regularity in Space.

Proposition 3. *Assume that the terminal condition in* (23) *is given by* (25) *with a Lipschitz non-decreasing function ϕ with values in $[0,1]$. Then, for each $(t_0,e) \in [0,T] \times \mathbb{R}$,* (23) *admits a unique solution $(\bar{E}_t^{t_0,e}, Y_t^{t_0,e}, Z_t^{t_0,e})_{t_0 \le t \le T}$ satisfying $\bar{E}_{t_0}^{t_0,e} = e$ and $Y_T^{t_0,e} = \phi(\bar{E}_T^{t_0,e})$. Moreover, the mapping*

$$(t,e) \hookrightarrow v(t,e) = Y_t^{t,e}$$

is $[0,1]$-valued, is of class $\mathcal{C}^{1,2}$ on $[0,T) \times \mathbb{R}$ and has Hölder continuous first-order derivative in time and first and second-order derivatives in space.

Moreover, the Hölder norms of v, $\partial_e v$, $\partial^2_{e,e} v$ and $\partial_t v$ on a given compact subset of $[0,T) \times \mathbb{R}$ do not depend upon the smoothness of ϕ provided ϕ is $[0,1]$-valued and non-decreasing. Specifically, the first-order derivative in space satisfies

$$0 \le \partial_e v(t,e) \le \frac{1}{T-t}, \quad t \in [0,T). \tag{26}$$

In particular, $e \hookrightarrow v(t,e)$ is non-decreasing for any $t \in [0,T)$.

Finally, for a given initial condition (t_0,e), the processes $(Y_t^{t_0,e})_{t_0 \le t \le T}$ and $(Z_t^{t_0,e})_{t_0 \le t < T}$, solution to the backward equation in (23) *(with ϕ as boundary condition), are given by:*

$$Y_t^{t_0,e} = v(t, \bar{E}_t^{t_0,e}),\ t_0 \le t \le T\ ; \quad Z_t^{t_0,e} = (T-t)\partial_e v(t, \bar{E}_t^{t_0,e}),\ t_0 \le t < T. \tag{27}$$

Proof. The problem is to solve the system

$$\begin{cases} d\bar{E}_t = -Y_t dt + (T-t) dW_t, \\ dY_t = Z_t dW_t, \end{cases} \tag{28}$$

with $\xi = \phi(\bar{E}_T)$ as terminal condition and (t_0,e) as initial condition. The drift in the first equation, i.e. $(t,y) \in [0,T] \times \mathbb{R} \hookrightarrow -y$, is decreasing in y, and Lipschitz continuous, uniformly in t. By Theorem 2.2 in Peng and Wu [15] (with $G=1$, $\beta_1 = 0$ and $\beta_2 = 1$ therein), we know that equation (28) admits at most one solution. Unfortunately, Theorem 2.6 in Peng and Wu (see also Remark 2.8 therein) does not apply to prove existence directly.

To prove existence, we use a variation of the induction method in Delarue [6]. By Theorem 1.1 in [6], existence and uniqueness are known to hold in small time. Specifically, we can find some small positive real δ, possibly depending on the Lipschitz constant of ϕ, such that (28) admits a unique solution for $T - t_0 \le \delta$. Remember that the initial condition is $\bar{E}_{t_0} = e$. As a consequence, we can define the *value function* $v : (t_0,e) \in [T-\delta,T] \times \mathbb{R} \hookrightarrow Y_{t_0}^{t_0,e}$. By Corollary 1.5 in [6], it is known to be Lipschitz in space uniformly in time as long as the time parameter remains in $[T-\delta,T]$. The diffusion coefficient $T-t$ in (28) being uniformly bounded away from 0 on the interval $[0,T-\delta]$, by Theorem 2.6 in [6], (28) admits a unique solution on $[t_0,T-\delta]$. Therefore, we can construct a solution to (28) in two steps: we first solve (28) on $[t_0,T-\delta]$ with $\bar{E}_{t_0} = e$ as initial condition and $v(T-\delta,\cdot)$ as terminal condition, the solution being denoted by $(\bar{E}_t, Y_t, Z_t)_{t_0 \le t \le T-\delta}$; then, we solve (28) on $[T-\delta,T]$ with the previous $\bar{E}_{T-\delta}$ as initial condition and with ϕ as terminal condition, the solution being denoted by $(\bar{E}'_t, Y'_t, Z'_t)_{T-\delta \le t \le T}$. We already know that $\bar{E}'_{T-\delta}$ matches $\bar{E}_{T-\delta}$. To patch $(\bar{E}_t, Y_t, Z_t)_{t_0 \le t \le T-\delta}$ and $(\bar{E}'_t, Y'_t, Z'_t)_{T-\delta \le t \le T}$ into a single solution over the whole time interval $[t_0,T]$, it is sufficient to check the continuity property $Y_{T-\delta} = Y'_{T-\delta}$ as done in Delarue [6]. This continuity property is a straightforward consequence of Corollary 1.5 in [6]: on $[T-\delta,T]$, $(Y'_t)_{T-\delta \le t \le T}$ has the form

$Y'_t = v(t, \bar{E}'_t)$. In particular, $Y'_{T-\delta} = v(T-\delta, \bar{E}'_{T-\delta}) = v(T-\delta, \bar{E}_{T-\delta}) = Y_{T-\delta}$. This proves the existence of a solution to (28) with $\bar{E}_{t_0} = e$ as initial condition.

We conclude that, for any (t_0, e), (28) admits a unique solution $(\bar{E}_t^{t_0,e}, Y_t^{t_0,e}, Z_t^{t_0,e})_{t_0 \le t \le T}$ satisfying $\bar{E}_{t_0}^{t_0,e} = e$ and $Y_T^{t_0,e} = \phi(\bar{E}_T^{t_0,e})$. In particular, the value function $v : (t_0, e) \hookrightarrow Y_{t_0}^{t_0,e}$ (i.e. the value at time t_0 of the solution $(Y_t)_{t_0 \le t \le T}$ under the initial condition $\bar{E}_{t_0} = e$) can be defined on the whole $[0,T] \times \mathbb{R}$.

From Corollary 1.5 in [6] and the discussion above, we know that the mapping $e \hookrightarrow v(t,e)$ is Lipschitz continuous when $T-t$ is less than δ and that, for any $t_0 \in [0,T]$, $Y_t^{t_0,e}$ has the form $Y_t^{t_0,e} = v(t, \bar{E}_t^{t_0,e})$ when $T-t$ is less than δ. In particular, on any $[0, T-\delta']$, δ' being less than δ, (28) may be seen as a uniformly elliptic FBSDE with a Lipschitz boundary condition. By Theorem 2.1 in Delarue and Guatteri [8] (together with the discussion in Section 8 therein), we deduce that v belongs to $\mathcal{C}^0([0,T] \times \mathbb{R}, \mathbb{R}) \cap \mathcal{C}^{1,2}([0,T) \times \mathbb{R}, \mathbb{R})$, that $t \hookrightarrow \|\partial_e v(t,\cdot)\|_\infty$ is bounded on the whole $[0,T]$ and that $t \hookrightarrow \|\partial^2_{ee} v(t,\cdot)\|_\infty$ is bounded on every compact subset of $[0,T)$[1]. Moreover, (27) holds.

By the martingale property of $(Y_t^{t_0,e})_{t_0 \le t \le T}$, it is well-seen that v is $[0,1]$-valued. To prove that it is non-decreasing (with respect to e), we follow the proof of Theorem 1. We notice that $(\bar{E}_t^{t_0,e})_{t_0 \le t \le T}$ satisfies the SDE:

$$d\bar{E}_t^{t_0,e} = -v(t, \bar{E}_t^{t_0,e})dt + (T-t)dW_t, \quad t_0 \le t \le T,$$

which has a Lipschitz drift with respect to the space variable. In particular, for $e \le e'$, $\bar{E}_T^{t_0,e} \le \bar{E}_T^{t_0,e'}$, so that $v(t_0,e) = \mathbb{E}[\phi(\bar{E}_T^{t_0,e})] \le \mathbb{E}[\phi(\bar{E}_T^{t_0,e'})] = v(t_0,e')$.

We now establish (26). For $t_0 \le t \le T$, the forward equation in (28) has the form

$$\bar{E}_t^{t_0,e} = e - \int_{t_0}^t v(s, \bar{E}_s^{t_0,e})ds + \int_{t_0}^t (T-s)dW_s. \tag{29}$$

Since v is $\mathcal{C}^1$ in space on $[0,T) \times \mathbb{R}$ with bounded Lipschitz first-order derivative, we can apply standard results on the differentiability of stochastic flows (see for example Kunita's monograph [10]). We deduce that, for almost every realization of the randomness and for any $t \in [t_0, T)$, the mapping $e \hookrightarrow \bar{E}_t^{t_0,e}$ is differentiable and

$$\partial_e \bar{E}_t^{t_0,e} = 1 - \int_{t_0}^t \partial_e v(s, \bar{E}_s^{t_0,e}) \partial_e \bar{E}_s^{t_0,e} ds. \tag{30}$$

In particular,

$$\partial_e \bar{E}_t^{t_0,e} = \exp\left(- \int_{t_0}^t \partial_e v(s, \bar{E}_s^{t_0,e}) ds \right). \tag{31}$$

Since v is non-decreasing, we know that $\partial_e v \ge 0$ on $[0,T) \times \mathbb{R}$ so that $\partial_e \bar{E}_t^{t_0,e}$ belongs to $[0,1]$. Since $\partial_e v$ is also bounded on the whole $[0,T) \times \mathbb{R}$, we deduce by differentiating the right-hand side in (29)

[1] Specifically, Theorem 2.1 in [8] says that v belongs to $\mathcal{C}^0([0,T) \times \mathbb{R}, \mathbb{R})$ and that $t \hookrightarrow \|\partial_e v(t,\cdot)\|_\infty$ is bounded on every compact subset of $[0,T)$. In fact, by Corollary 1.5 in Delarue [6], we know that v belongs to $\mathcal{C}^0([T-\delta,T] \times \mathbb{R}, \mathbb{R})$ and that $t \hookrightarrow \|\partial_e v(t,\cdot)\|_\infty$ is bounded on $[T-\delta, T]$ for δ small enough.

with $t = T$ that $\partial_e \bar{E}_T^{t_0,e}$ exists as well and that $\partial_e \bar{E}_T^{t_0,e} = \lim_{t\to T} \partial_e \bar{E}_t^{t_0,e} \in [0,1]$. To complete the proof of (26), we then notice that for any $t \in [t_0, T]$,

$$d\big[(T-t)Y_t^{t_0,e} - \bar{E}_t^{t_0,e}\big] = (T-t)dY_t^{t_0,e} - (T-t)dW_t = (T-t)[Z_t^{t_0,e} - 1]dW_t,$$

so that taking the expectations we get:

$$(T-t_0)v(t_0,e) - e = -\mathbb{E}\big[\bar{E}_T^{t_0,e}\big].$$

Now, differentiating with respect to e, we have:

$$(T-t_0)\partial_e v(t_0,e) = 1 - \mathbb{E}\big[\partial_e \bar{E}_T^{t_0,e}\big] \le 1,$$

which concludes the proof of (26).

It now remains to investigate the Hölder norms (both in time and space) of v, $\partial_e v$, $\partial^2_{ee} v$ and $\partial_t v$.

We first deal with v itself. For $0 < t < s < T$,

$$\begin{aligned} v(s,e) - v(t,e) &= v(s,e) - v(s,\bar{E}_s^{t,e}) + v(s,\bar{E}_s^{t,e}) - v(t,e) \\ &= v(s,e) - v(s,\bar{E}_s^{t,e}) + Y_s^{t,e} - Y_t^{t,e} \\ &= v(s,e) - v(s,\bar{E}_s^{t,e}) + \int_t^s Z_r^{t,e} dB_r. \end{aligned}$$

From (26), we deduce

$$\begin{aligned} |v(s,e) - v(t,e)| &\le \frac{1}{T-s}\mathbb{E}\big|\bar{E}_s^{t,e} - e\big| + \mathbb{E}\left[\left|\int_t^s Z_r^{t,e} dB_r\right|\right] \\ &\le \frac{1}{T-s}\left[s - t + \left(\int_t^s (T-r)^2 dr\right)^{1/2}\right] + \mathbb{E}\left[\int_t^s |Z_r^{t,e}|^2 dr\right]^{1/2} \\ &\le \frac{1}{T-s}\left[s - t + \left(\int_t^s (T-r)^2 dr\right)^{1/2}\right] + (s-t)^{1/2}, \end{aligned}$$

since $Z_r^{t,e} = (T-r)\partial_e v(r, \bar{E}_r^{t,e}) \in [0,1]$. So for $\epsilon > 0$, v is 1/2-Hölder continuous in time $t \in [0, T-\epsilon]$, uniformly in space and in the smoothness of ϕ.

Now, by Theorem 2.1 in Delarue and Guatteri [8], we know that v satisfies the PDE

$$\partial_t v(t,e) + \frac{(T-t)^2}{2}\partial^2_{ee} v(t,e) - v(t,e)\partial_e v(t,e) = 0, \quad t \in [0,T),\ e \in \mathbb{R}, \tag{32}$$

with ϕ as boundary condition. On $[0, T-\epsilon] \times \mathbb{R}$, $\epsilon > 0$, equation (32) is a non-degenerate second-order PDE of dimension 1 with $-v$ as drift, this drift being $\mathcal{C}^{1/2,1}$-continuous independently of the smoothness of ϕ. By well-known results in PDEs (so called Schauder estimates, see for example Theorem 8.11.1 in Krylov [9]), for any small $\eta > 0$, the $\mathcal{C}^{(3-\eta)/2, 3-\eta}$-norm of v on $[0, T-\epsilon] \times \mathbb{R}$ is independent of the smoothness of ϕ. □

Remark 5. *As announced, equation* (32) *is of Burgers type. In particular, it has the same first-order part as equation* (24).

4.2. **Boundary Behavior.** Still in the framework of a terminal condition given by a smooth (i.e. non-decreasing Lipschitz) function with values in $[0,1]$, we investigate the shape of the solution as t approaches T.

Proposition 4. *Assume that there exists some real Λ^+ such that $\phi(e)=1$ on $[\Lambda^+,+\infty)$. Then, there exists a universal constant $c>0$ such that for any $\delta>0$*

$$v\big(t,\Lambda^+ + T - t + \delta\big) \geq 1 - \exp\big(-c\frac{\delta^2}{(T-t)^3}\big), \quad 0 \leq t < T. \tag{33}$$

In particular, $v(t,e)\to 1$ as $t\nearrow T$ uniformly in e in compact subsets of $(\Lambda^+,+\infty)$.

Similarly, assume that there exists an interval $(-\infty,\Lambda^-]$ such that $\phi(e)=0$ on $(-\infty,\Lambda^-]$. Then, for any $\delta>0$,

$$v(t,\Lambda^- - \delta) \leq \exp\big(-c\frac{\delta^2}{(T-t)^3}\big). \tag{34}$$

In particular, $v(t,e)\to 0$ as $t\nearrow T$ uniformly in e in compact subsets of $(-\infty,\Lambda^-)$.

Proof. We only prove (33), the proof of (34) being similar. To do so, we fix $(t_0,e)\in[0,T)\times\mathbb{R}$ and consider the following system

$$\begin{cases} dE_t^- = -dt + (T-t)dW_t \\ dY_t^- = Z_t^- dW_t, \quad t_0 \leq t \leq T, \end{cases}$$

with $E_{t_0}^- = e$ as initial condition for the forward equation and $Y_T^- = \phi(E_T^-)$ as terminal condition for the backward part. The solution $(\bar{E}_t^{t_0,e}, Y_t^{t_0,e}, Z_t^{t_0,e})_{t_0\leq t\leq T}$ given by Proposition 3 with $\bar{E}_{t_0}^{t_0,e}=e$ and $Y_T^{t_0,e}=\phi(\bar{E}_T^{t_0,e})$ satisfies $Y_t^{t_0,e}\in[0,1]$ for any $t\in[t_0,T]$ so that $E_t^-\leq \bar{E}_t^{t_0,e}$ almost surely for $t\in[t_0,T]$. Now, since ϕ is non-decreasing, $\phi(E_T^-)\leq\phi(\bar{E}_T^{t_0,e})$ almost surely, namely $Y_{t_0}^-\leq Y_{t_0}^{t_0,e}$. Setting $v^-(t_0,e)=Y_{t_0}^-$, recall that $Y_{t_0}^-$ is deterministic, we see that:

$$v^-(t_0,e) \leq v(t_0,e) \leq 1. \tag{35}$$

Now, since

$$v^-(t_0,e) = \mathbb{E}\big[\phi(E_T^-)\big] = \mathbb{E}\bigg[\phi\bigg(e-(T-t_0)+\int_{t_0}^T (T-s)dW_s\bigg)\bigg]$$

with $\phi\geq \mathbf{1}_{[\Lambda^+,+\infty)}$, by choosing $e=\Lambda^+ + (T-t_0)+\delta$ as in the statement of Proposition 4 we get:

$$\begin{aligned}
\mathbb{E}\big[\phi(E_T^-)\big] &= \mathbb{E}\bigg[\phi\bigg(\Lambda^+ + \delta + \int_{t_0}^T (T-s)dW_s\bigg)\bigg] \\
&\geq \mathbb{P}\bigg\{\Lambda^+ + \delta + \int_{t_0}^T (T-s)dW_s \geq \Lambda^+\bigg\} \\
&= \mathbb{P}\bigg\{\int_{t_0}^T (T-s)dW_s \geq -\delta\bigg\} = 1 - \mathbb{P}\bigg\{\int_{t_0}^T (T-s)dW_s \leq -\delta\bigg\}
\end{aligned}$$

and we complete the proof by applying standard estimates for the decay of the cumulative distribution function of a Gaussian random variable. Note indeed that $\mathrm{var}(\int_{t_0}^T (T-s)dW_s) = (T-t_0)^3/3$ if we use the notation $\mathrm{var}(\xi)$ for the variance of a random variable ξ. □

The following corollary elucidates the boundary behavior between Λ^- and $\Lambda^+ + (T-t)$ with Λ^- and Λ^+ as above.

Corollary 1. *Choose ϕ as in Proposition 4. If there exists an interval $[\Lambda^+, +\infty)$ on which $\phi(e) = 1$, then for $\alpha > 0$ and $e < \Lambda^+ + (T-t) + (T-t)^{1+\alpha}$ we have:*

$$v(t,e) \geq \frac{e-\Lambda^+}{T-t} - \exp\big(-\frac{c}{(T-t)^{1-2\alpha}}\big) - (T-t)^\alpha, \tag{36}$$

for the same c as in the statement of Proposition 4.

Similarly, if there exists an interval $(-\infty, \Lambda^-]$ on which $\phi(e) = 0$, then for $\alpha > 0$ and $e > \Lambda^- - (T-t)^{1+\alpha}$ we have:

$$v(t,e) \leq \frac{e-\Lambda^-}{T-t} + \exp\big(-\frac{c}{(T-t)^{1-2\alpha}}\big) + (T-t)^\alpha. \tag{37}$$

Proof. We first prove (36). Since $v(t,\cdot)$ is $1/(T-t)$ Lipschitz continuous, we have:

$$\begin{aligned} v\big(t, \Lambda^+ + (T-t) + (T-t)^{1+\alpha}\big) - v(t,e) &\leq \frac{\Lambda^+ - e + (T-t) + (T-t)^{1+\alpha}}{T-t} \\ &= \frac{\Lambda^+ - e}{T-t} + 1 + (T-t)^\alpha. \end{aligned}$$

Therefore,

$$v(t,e) \geq v\big(t, \Lambda^+ + (T-t) + (T-t)^{1+\alpha}\big) - 1 - (T-t)^\alpha - \frac{\Lambda^+ - e}{T-t},$$

and applying (33)

$$v(t,e) \geq \frac{e-\Lambda^+}{T-t} - \exp\big(-c(T-t)^{2\alpha-1}\big) - (T-t)^\alpha.$$

For the upper bound, we use the same strategy. We start from

$$v(t,e) - v\big(t, \Lambda^- - (T-t)^{1+\alpha}\big) \leq \frac{e-\Lambda^-}{T-t} + (T-t)^\alpha,$$

so that

$$v(t,e) \leq \frac{e-\Lambda^-}{T-t} + \exp\big(-c(T-t)^{2\alpha-1}\big) + (T-t)^\alpha.$$

□

4.3. Existence of a Solution.

We now establish the existence of a solution to (23) with the original terminal condition. We use a compactness argument giving the existence of a *value function* for the problem.

Proposition 5. *There exists a continuous function $v : [0,T) \times \mathbb{R} \hookrightarrow [0,1]$ satisfying*

(1) *v belongs to $\mathcal{C}^{1,2}([0,T) \times \mathbb{R}, \mathbb{R})$ and solves the PDE* (32)*,*
(2) *$v(t,\cdot)$ is non-decreasing and $1/(T-t)$-Lipschitz continuous for any $t \in [0,T)$,*
(3) *v satisfies* (33) *and* (34) *with $\Lambda^- = \Lambda^+ = \Lambda$,*
(4) *v satisfies* (36) *and* (37) *with $\Lambda^- = \Lambda^+ = \Lambda$,*

and for any initial condition $(t_0, e) \in [0, T) \times \mathbb{R}$*, the strong solution* $(\bar{E}_t^{t_0,e})_{t_0 \le t < T}$ *of*

$$\bar{E}_t = e - \int_{t_0}^t v(s, \bar{E}_s) ds + \int_{t_0}^t (T - s) dW_s, \quad t_0 \le t < T, \tag{38}$$

is such that $(v(t, \bar{E}_t^{t_0,e}))_{t_0 \le t < T}$ *is a martingale with respect to the filtration generated by* W*.*

Proof. Choose a sequence of $[0, 1]$-valued smooth non-decreasing functions $(\phi^n)_{n \ge 1}$ such that $\phi^n(e) = 0$ for $e \le \Lambda - 1/n$ and $\phi^n(e) = 1$ for $e \ge \Lambda + 1/n$, $n \ge 1$, and denote by $(v^n)_{n \ge 1}$ the corresponding sequence of functions given by Proposition 3. By Proposition 3, we can extract a subsequence, which we will still index by n, converging uniformly on compact subsets of $[0, T) \times \mathbb{R}$. We denote by v such a limit. Clearly, v satisfies (1) in the statement of Proposition 5. Moreover, it also satisfies (2) because of Proposition 3, (3) by Proposition 4, and (4) by Corollary 1. Having Lipschitz coefficients, the stochastic differential equation (38) has a unique strong solution on $[t_0, T)$ for any initial condition $\bar{E}_{t_0} = e$. If we denote the solution by $(\bar{E}_t^{t_0,e})_{t_0 \le t < T}$, Itô's formula and the PDE (32), imply that the process $(v(t, \bar{E}_t^{t_0,e}))_{t_0 \le t < T}$ is a local martingale. Since it is bounded, it is a bona fide martingale. □

We finally obtain the desired solution to the FBSDE in the sense of Theorem 2.

Proposition 6. v *and* $(\bar{E}_t^{t_0,e})_{t_0 \le t < T}$ *being as above and setting*

$$Y_t^{t_0,e} = v(t, \bar{E}_t^{t_0,e}), \; Z_t^{t_0,e} = (T - t) \partial_e v(t, \bar{E}_t^{t_0,e}), \quad t_0 \le t < T,$$

the process $(\bar{E}_t^{t_0,e})_{t_0 \le t < T}$ *has an a.s. limit* $\bar{E}_T^{t_0,e}$ *as t tends to* T*. Similarly, the process* $(Y_t^{t_0,e})_{t_0 \le t < T}$ *has an a.s. limit* $Y_T^{t_0,e}$ *as* t *tends to* T *and the extended process* $(Y_t^{t_0,e})_{t_0 \le t \le T}$ *is a martingale with respect to the filtration generated by* W*. Morever,* $\mathbb{P}$*-a.s., we have:*

$$\mathbf{1}_{(\Lambda,\infty)}(\bar{E}_T^{t_0,e}) \le Y_T^{t_0,e} \le \mathbf{1}_{[\Lambda,\infty)}(\bar{E}_T^{t_0,e}). \tag{39}$$

and

$$Y_T^{t_0,e} = Y_{t_0}^{t_0,e} + \int_{t_0}^T Z_t^{t_0,e} dW_t, \tag{40}$$

Notice that $Z_t^{t_0,e}$ is not defined for $t = T$.

Proof. The proof is straightforward now that we have collected all the necessary ingredients. We start with the extension of $(\bar{E}_t^{t_0,e})_{t_0 \le t < T}$ up to time T. The only problem is to extend the drift part in (38), but since v is non-negative and bounded, it is clear that the process

$$\left(\int_{t_0}^t v(s, \bar{E}_s^{t_0,e}) ds \right)_{t_0 \le t < T}$$

is almost-surely increasing in t, so that the limit exists. The extension of $(Y_t^{t_0,e})_{t_0 \le t < T}$ up to time T follows from the almost-sure convergence theorem for positive martingales.

To prove (39), we apply (3) in the statement of Proposition 5. If $\bar{E}_T^{t_0,e} = \lim_{t \to T} \bar{E}_t^{t_0,e} > \Lambda$, then we can find some $\delta > 0$ such that $\bar{E}_t^{t_0,e} > \Lambda + (T - t) + \delta$ for t close to T, so that $Y_t^{t_0,e} = v(t, \bar{E}_t^{t_0,e}) \ge 1 - \exp[-c\delta^2/(T - t)^3]$ for t close to T, i.e. $Y_T^{t_0,e} \ge 1$. Since $Y_T^{t_0,e} \le 1$, we deduce that

$$\bar{E}_T^{t_0,e} > \Lambda \Rightarrow Y_T^{t_0,e} = 1.$$

In the same way,

$$\bar{E}_T^{t_0,e} < \Lambda \Rightarrow Y_T^{t_0,e} = 0.$$

This proves (39). Finally (40) follows from Itô's formula. Indeed, by Itô's formula and (32),

$$Y_t^{t_0,e} = Y_{t_0}^{t_0,e} + \int_{t_0}^{t} Z_s^{t_0,e} dW_s, \quad t_0 \le t < T.$$

By definition, $Z_s^{t_0,e} = (T-s)\partial_e v(s, \bar{E}_s^{t_0,e})$, $t_0 \le s < T$. By Point (2) in the statement of Proposition 5, it is in $[0,1]$. Therefore, the Itô integral

$$\int_{t_0}^{T} Z_s^{t_0,e} dW_s$$

makes sense as an element of $L^2(\Omega, \mathbb{P})$. This proves (40). □

4.4. **Improved Gradient Estimates.** Using again standard results on the differentiability of stochastic flows (see again Kunita's monograph [10]) we see that formulae (30) and (31) still hold in the present situation of a discontinuous terminal condition. We also prove a representation for the gradient of v of Malliavin-Bismut type.

Proposition 7. *For $t_0 \in [0,T)$, $\partial_e v(t_0,e)$ admits the representation*

$$\partial_e v(t_0,e) = 2(T-t_0)^{-2}\mathbb{E}\left[\lim_{\delta\to 0} v\big(T-\delta, \bar{E}_{T-\delta}^{t_0,e}\big) \int_{t_0}^{T} \partial_e \bar{E}_t^{t_0,e} dW_t\right]. \tag{41}$$

In particular, there exists some constant $A > 0$ such that

$$\sup_{|e|>A} \sup_{0\le t\le T} \partial_e v(t,e) < +\infty. \tag{42}$$

Proof. For $\delta > 0$, Proposition 6 yields

$$\begin{aligned}\mathbb{E}\left[v\big(T-\delta, \bar{E}_{T-\delta}^{t_0,e}\big) \int_{t_0}^{T} \partial_e \bar{E}_t^{t_0,e} dW_t\right] &= \mathbb{E}\left[\int_{t_0}^{T-\delta} Z_t^{t_0,e} dW_t \int_{t_0}^{T} \partial_e \bar{E}_t^{t_0,e} dW_t\right] \\ &= \mathbb{E}\left[\int_{t_0}^{T-\delta} (T-t)\partial_e v\big(t, \bar{E}_t^{t_0,e}\big)\partial_e \bar{E}_t^{t_0,e} dt\right].\end{aligned}$$

By the bounds we have on $\partial_e v$ and $(\partial_e \bar{E}_t^{t_0,e})_{t_0\le t<T}$, we can exchange the symbols $\mathbb{E}$ and $\int$. We obtain:

$$\mathbb{E}\left[v\big(T-\delta, \bar{E}_{T-\delta}^{t_0,e}\big) \int_{t_0}^{T} \partial_e \bar{E}_t^{t_0,e} dW_t\right] = \int_{t_0}^{T-\delta} (T-t)\mathbb{E}\big[\partial_e\big[v\big(t, \bar{E}_t^{t_0,e}\big)\big]\big] dt.$$

Similarly, we can exchange the symbols $\mathbb{E}$ and ∂_e, so that

$$\mathbb{E}\left[v\big(T-\delta, \bar{E}_{T-\delta}^{t_0,e}\big) \int_{t_0}^{T} \partial_e \bar{E}_t^{t_0,e} dW_t\right] = \int_{t_0}^{T-\delta} (T-t)\partial_e\big[\mathbb{E}\big[v\big(t, \bar{E}_t^{t_0,e}\big)\big]\big] dt.$$

Since $(v(t,\bar{E}_t^{t_0,e}))_{t_0 \le t \le T-\delta}$ is a martingale, we deduce:

$$\begin{aligned}\mathbb{E}\Big[v\big(T-\delta,\bar{E}_{T-\delta}^{t_0,e}\big)\int_{t_0}^{T}\partial_e \bar{E}_t^{t_0,e}dW_t\Big] &= \partial_e v(t_0,e)\int_{t_0}^{T-\delta}(T-t)dt \\ &= \frac{1}{2}(T-\delta-t_0)(T+\delta-t_0)\partial_e v(t_0,e).\end{aligned}$$

Letting δ tend to zero and applying dominated convergence, we complete the proof of the representation formula of the gradient.

To derive the bound (42), we emphasize that, for e away from Λ (say for example $e \ll \Lambda$), the probability that $(\bar{E}_t^{t_0,e})_{t_0 \le t \le T}$ hits Λ is very small and decays exponentially fast as $T-t_0$ tends to 0. On the complement, i.e. for $\sup_{t_0 \le t \le T}\bar{E}_t^{t_0,e} < \Lambda$, we know that $v(t,\bar{E}_t^{t_0,e})$ tends to 0 as t tends to T. Specifically, following the proof of Proposition 4, there exists a universal constant $c' > 0$ such that for any $e \le \Lambda - 1$ and $t_0 \in [0,T)$

$$\begin{aligned}(T-t_0)^2\partial_e v(t_0,e) &\le 2(T-t_0)^{1/2}\mathbb{P}^{1/2}\big\{\sup_{t_0 \le t \le T}\bar{E}_t^{t_0,e} \ge \Lambda\big\} \\ &\le 2(T-t_0)^{1/2}\mathbb{P}^{1/2}\big\{\Lambda - 1 + \sup_{t_0 \le t \le T}\int_{t_0}^{t}(T-s)dW_s \ge \Lambda\big\} \\ &\le 2(T-t_0)^{1/2}\mathbb{P}^{1/2}\big\{\sup_{t_0 \le t \le T}\int_{t_0}^{t}(T-s)dW_s \ge 1\big\} \\ &\le 2(T-t_0)^{1/2}\exp\big(-\frac{c'}{(T-t_0)^3}\big),\end{aligned}$$

the last line following from maximal inequality (IV.37.12) in Rogers and Williams [17].

The same argument holds for $e > \Lambda + 2$ by noting that (41) also holds for $v - 1$. □

Remark 6. *The stochastic integral in the Malliavin-Bismut formula* (41) *is at most of order* $(T-t_0)^{1/2}$. *Therefore, the typical resulting bound we obtain for* $\partial_e v(t,e)$ *in the neighborhood of* (T,Λ) *is* $(T-t)^{-3/2}$. *Obviously, it is less accurate than the bound given by Propositions 3 and 5. This says that the Lipschitz smoothing of the singularity of the boundary condition obtained in Propositions 3 and 5, namely* $\partial_e v(t,e) \le (T-t)^{-1}$, *follows from the first-order Burgers structure of the PDE* (32) *and that the diffusion term plays no role in it. This a clue to understand why the diffusion process* $\bar{E}$ *feels the trap made by the boundary condition. On the opposite, the typical bound for* $\partial_e v(t,e)$ *we would obtain in the uniformly elliptic by applying a Malliavin-Bismut formula (see Exercice 2.3.5 in Nualart* [14]*) is of order* $(T-t)^{-1/2}$, *which is much better than* $(T-t)^{-1}$.

Nevertheless, the following proposition shows that the diffusion term permits to improve the bound obtained in Propositions 3 and 5. Because of the noise plugged into $\bar{E}$, *the bound* $(T-t)^{-1}$ *cannot be achieved. This makes a real difference with the inviscid Burgers equation* (24) *which admits*

$$(t,e) \in [0,T) \times \mathbb{R} \hookrightarrow \psi\big(\frac{e-\Lambda}{T-t}\big),$$

as solution, with $\psi(e) = 1 \wedge e^+$ *for* $e \in \mathbb{R}$. *(See for example (10.12') in Lax* [11]*.)*

We thus prove the following stronger version of Propositions 3 and 5:

Proposition 8. *For any* $(t_0,e) \in [0,T) \times \mathbb{R}$, *it holds* $(T-t_0)\partial_e v(t_0,e) < 1$.

Proof. Given $(t_0, e) \in [0, T) \times \mathbb{R}$, we consider $(\bar{E}_t^{t_0,e}, Y_t^{t_0,e}, Z_t^{t_0,e})_{t_0 \le t \le T}$ as in the statement of Proposition 6. As in the proof of Proposition 3, we start from

$$d\big[(T-t)Y_t^{t_0,e} - \bar{E}_t^{t_0,e}\big] = (T-t)dY_t^{t_0,e} - (T-t)dW_t = (T-t)[Z_t^{t_0,e} - 1]dW_t, \quad t_0 \le t < T.$$

Therefore, for any initial condition (t_0, e),

$$(T-t_0)v(t_0,e) - e = -\mathbb{E}\big[\bar{E}_T^{t_0,e}\big].$$

Unfortunately, we do not know whether $\bar{E}_T^{t_0,e}$ is differentiable with respect to e. Anyhow,

$$\begin{aligned}(T-t_0)\partial_e v(t_0,e) &= 1 - \lim_{h\to 0} h^{-1}\mathbb{E}\big[\bar{E}_T^{t_0,e+h} - \bar{E}_T^{t_0,e}\big] \\ &= 1 - \lim_{h\to 0} h^{-1} \lim_{t\nearrow T} \mathbb{E}\big[\bar{E}_t^{t_0,e+h} - \bar{E}_t^{t_0,e}\big] \le 1 - \lim_{h\to 0}\lim_{t\nearrow T} \inf_{|u|\le h} \mathbb{E}\big[\partial_e \bar{E}_t^{t_0,e+u}\big]\end{aligned}$$

Using (31), the non-negativity of $\partial_e v$ and Fatou's lemma,

$$\begin{aligned}(T-t_0)\partial_e v(t_0,e) &\le 1 - \lim_{h\to 0}\lim_{t\nearrow T}\inf_{|u|\le h}\mathbb{E}\Big[\exp\Big(-\int_{t_0}^t \partial_e v(s,\bar{E}_s^{t_0,e+u})ds\Big)\Big] \\ &\le 1 - \lim_{h\to 0}\inf_{|u|\le h}\mathbb{E}\Big[\exp\Big(-\int_{t_0}^T \partial_e v(s,\bar{E}_s^{t_0,e+u})ds\Big)\Big] \\ &\le 1 - \mathbb{E}\Big[\exp\Big(-\lim_{h\to 0}\sup_{|u|\le h}\int_{t_0}^T \partial_e v(s,\bar{E}_s^{t_0,e+u})ds\Big)\Big].\end{aligned}$$

Consequently, in order to prove that $(T-t_0)\partial_e v(t_0,e) < 1$, it is enough to prove that the limit superior

$$\lim_{h\to 0}\sup_{|u|\le h}\int_{t_0}^T \partial_e v(t,\bar{E}_t^{t_0,e+u})dt \tag{43}$$

is finite with a non-zero probability. To do so, the Lipschitz bound given by Proposition 3 is not sufficient since the integral of the bound is divergent. To overcome this difficulty, we use (42): with non-zero probability, the values of the process $(\bar{E}_t)_{t_0\le t\le T}$ at the neighborhood of T may be made as large as desired. Precisely, for A as in Proposition 7, it is sufficient to prove that there exists $\delta > 0$ small enough such that $\mathbb{P}\{\inf_{|h|\le 1}\inf_{T-\delta\le t\le T}\bar{E}_t^{t_0,e+h} > A\} > 0$. For $\delta > 0$, we deduce from the boundedness of the drift in (38) that

$$\mathbb{P}\big\{\inf_{|h|\le 1}\inf_{T-\delta\le t\le T}\bar{E}_t^{t_0,e+h} > A\big\} \ge \mathbb{P}\Big\{e - 1 - (T-t_0) + \inf_{T-\delta\le t\le T}\int_{t_0}^t (T-s)dW_s > A\Big\}.$$

By independence of the increments of the Wiener integral, we get

$$\begin{aligned}&\mathbb{P}\big\{\inf_{|h|\le 1}\inf_{T-\delta\le t\le T}\bar{E}_t^{t_0,e+h} > A\big\} \\ &\ge \mathbb{P}\Big\{e - 1 - (T-t_0) + \int_{t_0}^{T-\delta}(T-s)dW_s > 2A\Big\}\mathbb{P}\Big\{\inf_{T-\delta\le t\le T}\int_{T-\delta}^t (T-s)dW_s > -A\Big\}.\end{aligned}$$

The first probability in the above right-hand side is clearly positive for $T-\delta > t_0$. The second one is equal to

$$\mathbb{P}\Big\{\inf_{T-\delta\le t\le T}\int_{T-\delta}^{t}(T-s)dW_s > -A\Big\} = 1-\mathbb{P}\Big\{\sup_{T-\delta\le t\le T}\int_{T-\delta}^{t}(T-s)dW_s \ge A\Big\}.$$

Using maximal inequality (IV.37.12) in Rogers and Williams [17], the above right hand-side is always positive. By (42), we deduce that, with non-zero probability, the limit superior in (43) is finite. □

4.5. **Distribution of $\bar{E}_t$ for $t_0 \le t \le T$.** We finally claim:

Proposition 9. *Keep the notation of Propositions 5 and 6 and choose some starting point $(t_0, e) \in [0,T) \times \mathbb{R}$ and some $p \in \mathbb{R}$. Then, for every $t \in [t_0, T)$, the law of the variable*

$$E_t^{t_0,e,p} = \bar{E}_t^{t_0,e} - (T-t)P_t^p = \bar{E}_t^{t_0,e} - (T-t)\big[p + W_t\big],$$

obtained by transformation (22)*, is absolutely continuous with respect to the Lebesgue measure. At time $t = T$, it has a Dirac mass at point Λ.*

Proof. Obviously, we can assume $p = 0$, so that $P_t = W_t$. (For simplicity, we will write $E_t^{t_0,e}$ for $E_t^{t_0,e,p}$.) We start with the absolute continuity of $E_t^{t_0,e}$ at time $t < T$. Since v is smooth away from T, we can compute the Malliavin derivative of $E_t^{t_0,e}$. (See Theorem 2.2.1 in Nualart [14].) It satisfies

$$D_s E_t^{t_0,e} = t - s - \int_s^t \partial_e v\big(r, E_r^{t_0,e} + (T-r)W_r\big) D_s E_r^{t_0,e} dr - \int_s^t (T-r)\partial_e v\big(r, E_r^{t_0,e} + (T-r)W_r\big) dr,$$

for $t_0 \le s \le t$. In particular,

$$\begin{aligned} D_s E_t^{t_0,e} = \int_s^t \Big[&\big[1 - (T-r)\partial_e v\big(r, E_r^{t_0,e} + (T-r)W_r\big)\big] \\ &\times \exp\Big(-\int_r^t \partial_e v\big(u, E_u^{t_0,e} + (T-u)W_u\big)du\Big)\Big] dr. \end{aligned} \tag{44}$$

By Proposition 8, we deduce that $D_s E_t^{t_0,e} > 0$ for any $t_0 \le s \le t$. By Theorem 2.1.3 in Nualart [14], we deduce that the law of $E_t^{t_0,e}$ has a density with respect to the Lebesgue measure.

To prove the existence of a point mass at time T, it is enough to focus on $\bar{E}_T^{t_0,e}$ since the latter is equal to $E_T^{t_0,e}$. We prove the desired result by comparing the stochastic dynamics of $\bar{E}_T^{t_0,e}$ to the time evolution of solutions of simpler stochastic differential equations. With the notation used so far, $\bar{E}_t^{t_0,e}$ is a solution of the SDE

$$d\bar{E}_t = -v(t, \bar{E}_t)dt + (T-t)dW_t \tag{45}$$

so it is natural to compare the solution of this SDE to solutions of SDEs with similar drifts. Following Remark 5, we are going to do so by comparing v with the solution of the inviscid Burgers equation (24). To this effect we use once more the function ψ defined by $\psi(e) = 1 \wedge e^+$ introduced earlier. As said in Remark 6, the function $\psi((e-\Lambda)/(T-t))$ is a solution of the Burgers equation (24) which, up to the diffusion term (which decreases to 0 like $(T-t)^2$ when $t \nearrow T$), is the same as the PDE satisfied by v. Using (33) and (34) with $\Lambda^- = \Lambda^+ = \Lambda$ and $\delta = (T-t)^{5/4}$, and (36) and (37) with $\Lambda^- = \Lambda^+ = \Lambda$ and $\alpha = 1/4$, we have

$$\big|v(t,e) - \psi\big(\frac{e-\Lambda}{T-t}\big)\big| \leq C(T-t)^{1/4}, \tag{46}$$

for some universal constant C. We now compare (45) with

$$dX_t^{\pm} = -\psi\big(\frac{X_t^{\pm}-\Lambda}{T-t}\big)dt \pm C(T-t)^{1/4}dt + (T-t)dW_t, \quad t_0 \leq t < T, \tag{47}$$

with $X_{t_0}^{\pm} = e$ as initial conditions. Clearly,

$$X_t^- \leq \bar{E}_t^{e,t_0} \leq X_t^+, \quad t_0 \leq t < T. \tag{48}$$

Knowing that $\psi(x) = x$ when $0 \leq x \leq 1$, we anticipate that scenarios satisfying $0 \leq X_t^{\pm} - \Lambda \leq T-t$ can be viewed as solving the SDEs

$$dZ_t^{\pm} = -\frac{Z_t^{\pm}-\Lambda}{T-t}dt \pm C(T-t)^{1/4}dt + (T-t)dW_t,$$

with $Z_{t_0}^{\pm} = e$ as initial conditions. This remark is useful because these SDEs have explicit solutions:

$$Z_t^{\pm} = \Lambda + (T-t)\big[W_t - W_{t_0} \mp 4C(T-t)^{1/4} \pm 4C(T-t_0)^{1/4} + \frac{e-\Lambda}{T-t_0}\big], \qquad t_0 \leq t \leq T. \tag{49}$$

We define the event F by:

$$F = \left\{ \sup_{t_0 \leq t \leq T} |W_t - W_{t_0}| \leq \frac{1}{8} \right\}$$

and we introduce the quantities $\underline{e}(t_0)$ and $\bar{e}(t_0)$ defined by

$$\underline{e}(t_0) = \Lambda + \frac{1}{4}(T-t_0) \qquad \text{and} \qquad \bar{e}(t_0) = \Lambda + \frac{3}{4}(T-t_0)$$

so that

$$\frac{1}{4} \leq \frac{e-\Lambda}{T-t_0} \leq \frac{3}{4}$$

whenever $\underline{e}(t_0) \leq e \leq \bar{e}(t_0)$. For such a choice of e, since

$$\frac{Z^{\pm}-\Lambda}{T-t} = W_t - W_{t_0} \mp 4C(T-t)^{1/4} \pm 4C(T-t_0)^{1/4} + \frac{e-\Lambda}{T-t_0},$$

it is easy to see that if we choose t_0 such that $T-t_0$ is small enough for $32C(T-t_0)^{1/4} < 1$ to hold, then

$$\forall t \in [t_0, T], \quad 0 \leq \frac{Z_t^- - \Lambda}{T-t} \leq \frac{Z_t^+ - \Lambda}{T-t} \leq 1.$$

on the event F. This implies that $(X_t^{\pm})_{t_0 \leq t < T}$ and $(Z_t^{\pm})_{t_0 \leq t < T}$ coincide on F, and consequently that $X_T^+ = X_T^- = \Lambda$ and hence $\bar{E}_T^{t_0,e} = \Lambda$ on F by (48). This completes the proof for these particular choices of t_0 and e. In fact, the result holds for any e and any $t_0 \in [0,T)$. Indeed, since $\bar{E}_t^{t_0,e}$

has a strictly positive density at any time $t \in (t_0, T)$, so that, if we choose $t_1 \in (t_0, T)$ so that $32C(T-t_1)^{1/4} < 1$, then using the Markov property we get

$$\mathbb{P}\big\{\bar{E}_T^{t_0,e} = \Lambda\big\} \geq \int_{\underline{e}(t_1)}^{\bar{e}(t_1)} \mathbb{P}\big\{\bar{E}_T^{t_1,e'} = \Lambda\big\}\mathbb{P}\big\{\bar{E}_{t_1}^{t_0,e} \in de'\big\} > 0$$

which completes the proof in the general case. □

Remark 7. *We emphasize that the expression for $D_s E_t^{t_0,e}$ given in* (44) *can vanish with a non-zero probability when replacing t by T. Indeed, the integral*

$$\int_r^T \partial_e v\big(u, E_u^{t_0,e} + (T-u)W_u\big)du$$

may explode with a non-zero probability since the derivative $\partial_e v(u,e)$ is expected to behave like $(T-u)^{-1}$ as u tends to T and e to Λ. Indeed, v is known to behave like the solution of the Burgers equation when close to the boundary, see (46)*. As a consequence, we expect $\partial_e v$ to behave like the gradient of the solution of the Burgers equation. The latter is singular in the neighborhood of the final discontinuity and explodes like $(T-u)^{-1}$ in the cone formed by the characteristics of the equation.*

However, in the uniformly elliptic setting, the integral above is always bounded since $\partial_e v(u,\cdot)$ is at most of order $(T-u)^{-1/2}$ as explained in Remark 6.

4.6. **Uniqueness.** Our proof of uniqueness is based on a couple of comparison lemmas.

Lemma 1. *Let ϕ be a non-decreasing smooth function with values in $[0,1]$ greater than $\mathbf{1}_{[\Lambda,+\infty)}$, and w be the solution of the PDE* (32) *with ϕ as terminal condition. Then, any solution $(\bar{E}'_t, Y'_t, Z'_t)_{t_0 \le t \le T}$ of* (23) *starting from $\bar{E}'_{t_0} = e$ and satisfying $\mathbf{1}_{(\Lambda,+\infty)}(\bar{E}'_T) \le Y'_T \le \mathbf{1}_{[\Lambda,+\infty)}(\bar{E}'_T)$ also satisfies*

$$w(t, \bar{E}'_t) \ge Y'_t, \quad t_0 \le t \le T.$$

Similarly, if ϕ is less than $\mathbf{1}_{(\Lambda,+\infty)}$, then

$$w(t, \bar{E}'_t) \le Y'_t, \quad t_0 \le t \le T.$$

Proof. Applying Itô's formula to $(w(t, \bar{E}'_t))_{t_0 \le t \le T}$, we obtain

$$d\big[w(t,\bar{E}'_t) - Y'_t\big] = \big(w(t,\bar{E}'_t) - Y'_t\big)\partial_e w(t,\bar{E}'_t)dt + \big[(T-t)\partial_e w(t,\bar{E}'_t) - Z'_t\big]dW_t.$$

Therefore,

$$d\bigg\{\big[w(t,\bar{E}'_t) - Y'_t\big]\exp\bigg(-\int_{t_0}^t \partial_e w(s,\bar{E}'_s)ds\bigg)\bigg\} = \exp\bigg(-\int_{t_0}^t \partial_e w(s,\bar{E}'_s)ds\bigg)\big[(T-t)\partial_e w(t,\bar{E}'_t) - Z'_t\big]dW_t.$$

In particular,

$$w(t,\bar{E}'_t) - Y'_t = \exp\bigg(\int_{t_0}^t \partial_e w(s,\bar{E}'_s)ds\bigg)\mathbb{E}\bigg[\exp\bigg(-\int_{t_0}^T \partial_e w(s,\bar{E}'_s)ds\bigg)\big[w(T,\bar{E}'_T) - Y'_T\big]|\mathcal{F}_t\bigg],$$

which completes the proof. □

The next lemma can be viewed as a form of conservation law.

Lemma 2. *Let $(\chi^n)_{n\geq 1}$ be a non-increasing sequence of non-decreasing smooth functions matching 0 on some intervals $(-\infty, \Lambda^{-,n})_{n\geq 1}$ and 1 on some intervals $(\Lambda^{+,n}, +\infty)_{n\geq 1}$ and converging towards $\mathbf{1}_{[\Lambda,+\infty)}$, then the associated solutions $(w^n)_{n\geq 1}$, given by Proposition 3 converge towards v constructed in Proposition 5.*

The conclusion remains true if $(\chi^n)_{n\geq 1}$ is a non-decreasing sequence converging towards $\mathbf{1}_{(\Lambda,+\infty)}$.

Proof. Each w^n is a solution of the PDE (32) which is conservative. Considering v^n as in the proof of Proposition 5, we have for any $n, m \geq 1$

$$\int_{\mathbb{R}} (w^n - v^m)(t,e)de = \int_{\mathbb{R}} (\chi^n - \phi^m)(e)de, \quad t \in [0,T).$$

Notice that the integrals are well-defined because of Proposition 4. Since $\phi^m(e) \to \mathbf{1}_{[\Lambda,+\infty)}(e)$ as $m \to +\infty$ for $e \neq \Lambda$, we deduce that

$$\int_{\mathbb{R}} (w^n - v)(t,e)de = \int_{\mathbb{R}} \big[\chi^n(e) - \mathbf{1}_{[\Lambda,+\infty)}(e)\big]de, \quad t \in [0,T).$$

Since the right hand side converges towards 0 as n tends to $+\infty$, so does the left hand side, but since $w^n(t,e) \geq v(t,e)$ by Lemma 1 (choosing $(\bar{E}', Y', Z') = (\bar{E}^{t_0,e}, Y^{t_0,e}, Z^{t_0,e})$), we must also have:

$$\lim_{n\to+\infty} \int_{\mathbb{R}} |w^n(t,e) - v(t,e)|de = 0.$$

Since $(w^n(t,\cdot))_{n\geq 1}$ is equicontinuous (by Proposition 3), we conclude that $w^n(t,e) \to v(t,e)$. The proof is similar if $\chi^n \nearrow \mathbf{1}_{(\Lambda,+\infty)}$. □

To complete the proof of uniqueness, consider a sequence $(\chi^n)_{n\geq 1}$ as in the statement of Lemma 2. For any solution $(\bar{E}'_t, Y'_t, Z'_t)_{t_0\leq t\leq T}$ of (23) with $\bar{E}'_{t_0} = e$, Lemma 1 yields

$$w^n(t, \bar{E}'_t) \geq Y'_t, \quad t \in [t_0, T).$$

Passing to the limit, we deduce that

$$v(t, \bar{E}'_t) \geq Y'_t, \quad t \in [t_0, T).$$

Choosing a non-increasing sequence $(\chi^n)_{n\geq 1}$, instead, we obtain the reverse inequality, and hence, we conclude that $Y'_t = v(t, \bar{E}'_t)$ for $t \in [t_0, T)$. By uniqueness to (38), we deduce that $\bar{E}'_t = \bar{E}^{t_0,e}_t$, so that $Y'_t = Y^{t_0,e}_t$. We easily deduce that $Z'_t = Z^{t_0,e}_t$ as well.

Remark 8. *We conjecture that the analysis performed in this section can be extended to more general conservation laws than Burgers equation. The Burgers case is the simplest one since the corresponding forward - backward stochastic differential equation is purely linear.*

5. Nonlinear PDEs and Option Pricing

In this section, we consider the problem of option pricing in the framework of the first equilibrium model introduced in this paper.

5.1. **PDE Characterization.** Back to the risk neutral dynamics of the (perceived) emissions given by (12), we assume that the emissions of the business as usual scenario are modeled by a geometric Brownian motion, so that $b(t,e) = be$ and $\sigma(t,e) = \sigma e$. As explained in the introduction, this model has been used in most of the early reduced form analyses of emissions allowance forward contracts and option prices (see[5] and [4] for example). The main thrust of this section is to include the impact of the allowance price Y on the dynamics of the cumulative emissions. As we already saw in the previous section, this feedback $f(Y_s)$ is the source of a nonlinearity in the PDE whose solution determines the price of an allowance. Throughout this section, we assume that under the pricing measure (martingale spot measure) the cumulative emissions and the price of a forward contract on an emission allowance satisfy the forward-backward system:

$$\begin{cases} E_t = E_0 + \int_0^t (bE_s - f(Y_s))ds + \int_0^t \sigma E_s d\tilde{W}_s \\ Y_t = \lambda\, \mathbf{1}_{[\Lambda,\infty)}(E_T) - \int_t^T Z_t d\tilde{W}_t. \end{cases} \tag{50}$$

The solution Y_t of the backward equation is constructed as a function $Y_t = v(t, E_t)$ of the solution of the forward equation where the function v is a classical solution of the nonlinear partial differential equation

$$\begin{cases} \partial_t v(t,e) + (be - f(v(t,e)))\partial_e v(t,e) + \frac{1}{2}\sigma^2 e^2 \partial^2_{ee} v(t,e) = 0, \qquad (t,e) \in [0,T) \times \mathbb{R}_+ \\ v(T,.) = \lambda \mathbf{1}_{[\Lambda,\infty)}. \end{cases} \tag{51}$$

The price at time $t < \tau$ of a European call option with maturity $\tau < T$ and strike K on an allowance forward contract maturing at time T is given by the expectation

$$\mathbb{E}_{t,e}\{(Y_\tau - K)^+\} = \mathbb{E}_{t,e}\{(v(\tau, E_\tau) - K)^+\}.$$

which can as before, be written as a function $V(t, E_t)$ of the current value of the cumulative emissions. We use the notation $\mathbb{E}_{t,e}$ for the conditional expectation given that $E_t = e$. Once the function v is known and/or computed, for exactly the same reasons as above, the function V appears as the classical solution of the linear partial differential equation:

$$\begin{cases} \partial_t V(t,e) + (be - f(v(t,e)))\partial_e V(t,e) + \frac{1}{2}\sigma^2 e^2 \partial^2_{ee} V(t,e) = 0, \quad (t,e) \in [0,\tau) \times \mathbb{R}_+ \\ V(\tau,.) = (v(\tau,.) - K)^+, \end{cases} \tag{52}$$

which, given the knowledge of v is a linear partial differential equation. Notice that in the case $f \equiv 0$ of infinite abatement costs, except for the fact that the coefficients of the geometric Brownian motion were assumed to be time dependent, the above option price is the same as the one derived in [4].

5.2. **Small Abatement Asymptotics.** Examining the PDEs (51) and (52), we see that there are two main differences with the classical Black-Scholes framework. First, the underlying contract price is determined by the nonlinear PDE (51). Second, the option pricing PDE (52) involves the nonlinear term $f(v(t,e))$, while still being linear in terms of the unknown function V. Because the function v is determined by the first PDE (51), this nonliearity is inherent to the model, and one cannot simply reduce the PDE to the Black-Scholes equation.

In order to understand the departure of the option prices from those of the Black-Scholes model, we introduce a small parmater $\epsilon \geq 0$, and take the abatement rate to be of the form $f = \epsilon f_0$ for some fixed non-zero increasing continuous function f_0. We denote by v^ϵ and V^ϵ the corresponding prices of the allowance forward contract and the option. Here, what we call *Black-Scholes model* corresponds to the case $f \equiv 0$. Indeed, in this case, both (51) and (52) reduce to the linear Black-Scholes PDE, differing only through their boundary conditions. This model was one of the models used in [4] for the purpose of pricing options on emission allowances based on price data exhibiting no implied volatility smile. Our starting point is the characterization of the emission allowance price and the corresponding option price by the PDEs (51) and (52):

$$\begin{cases} -\partial_t v^\epsilon - (be - \epsilon f_0(v^\epsilon))\partial_e v^\epsilon - \frac{1}{2}\sigma^2 e^2 \partial^2_{ee} v^\epsilon = 0, \quad \text{on } [0,T) \times \mathbb{R}_+ \\ v^\epsilon(T,.) = \lambda \mathbf{1}_{[\Lambda,\infty)}, \end{cases} \tag{53}$$

and

$$\begin{cases} -\partial_t V^\epsilon - (be - \epsilon f(v^\epsilon))\partial_e V^\epsilon - \frac{1}{2}\sigma^2 e^2 \partial^2_{ee} V^\epsilon = 0 \quad \text{on } [0,\tau) \times \mathbb{R}_+, \\ V^\epsilon(\tau,.) = (v^\epsilon(\tau,.) - K)^+. \end{cases} \tag{54}$$

for every $\epsilon \geq 0$. For $\epsilon = 0$, the nonlinear feedback given by the abatement rate disappears and we easily compute that

$$v^0(t,e) = \lambda \mathbb{E}_{t,e}\left[\mathbf{1}_{[\Lambda,\infty)}(E^0_T)\right] = \lambda \Phi\left(\frac{\ln(e/\Lambda e^{-b(T-t)})}{\sigma\sqrt{T-t}} - \frac{\sigma\sqrt{T-t}}{2}\right) \tag{55}$$

$$V^0(t,e) = \mathbb{E}_{t,e}\left[(v^0(\tau,E^0_\tau) - K)^+\right], \qquad 0 \leq t \leq \tau, \tag{56}$$

where E^0 is a geometric Brownian motion:

$$dE^0_t = E^0_t[bdt + \sigma d\tilde{W}_t]. \tag{57}$$

used as a proxy for the cumulative emissions in business as usual, and where we use the notation $\mathbb{E}_{t,e}$ to denote the conditional expectation given that $E^0_t = e$. See for example [4] for details and complements. The main technical result of this section is the following first order Taylor expansion of the option price.

Proposition 10. *As $\epsilon \to 0$, we have*

$$V^\epsilon(t,s) = V^0(t,s) + \epsilon\, \mathbb{E}_{t,e}\left[\mathbf{1}_{[\Lambda,\infty)}(v^0(\tau,E^0_\tau)) \int_t^T f_0(v^0(s,E^0_s))\partial_e v^0(s \vee \tau, E^0_{s\vee\tau}) \frac{E^0_{s\vee\tau}}{E^0_s} ds\right] + \circ(\epsilon),$$

where $\epsilon^{-1} \circ (\epsilon) \longrightarrow 0$ as $\epsilon \to 0$.

In preparation for the proof of this result, we isolate the main steps in separate lemmas.

Lemma 3.

$$\lim_{\epsilon \searrow 0} v^\epsilon = v^0 \quad \textit{and} \quad \lim_{\epsilon \searrow 0} \partial_e v^\epsilon = \partial_e v^0, \tag{58}$$

uniformly on compact subsets of $[0,T) \times \mathbb{R}_+$.

Proof. By definition, the function v^ϵ is a classical solution of

$$-\partial_t v^\epsilon(t,e) + F^\epsilon\left(e, v^\epsilon(t,e), \partial_e v^\epsilon(t,e), \partial^2_{ee} v^\epsilon(t,e)\right) = 0, \tag{59}$$

where

$$F^\epsilon(e,r,p,\gamma) := -(be - \epsilon f_0(r))p - \frac{1}{2}\sigma^2 e^2 \gamma.$$

Let

$$\underline{v} := \liminf_{(\epsilon,t',e')\to(0,t,e)} v^\epsilon(t',e') \text{ and } \quad \overline{v} := \limsup_{(\epsilon,t',e')\to(0,t,e)} v^\epsilon(t',e')$$

be the relaxed semi-limits of v^ϵ, which are finite because v^ϵ is locally bounded in (ϵ, t, e). Since $F^\epsilon(e, r, p, \gamma)$ is jointly continuous in all of its arguments (including ϵ), it follows from the stability of viscosity solutions (see for example [1]), that the functions $\underline{v}$ and $\overline{v}$ are viscosity supersolution and subsolution of the limit equation

$$-\partial_t v(t,e) + F^0\left(e, v(t,e), \partial_e v(t,e), \partial^2_{ee} v(t,e)\right) = 0 \tag{60}$$

which happens to be linear. Moreover, by construction we have $\underline{v}(T,.) \geq \mathbf{1}_{(\Lambda,\infty)}$ and $\overline{v}(T,.) \leq \mathbf{1}_{[\Lambda,\infty)}$. Arguying as in the proof of the Feynman-Kac representation, it follows from Itô's formula (after convenient localization) that $\underline{v} \geq v^0$ and $\overline{v} \leq v^0$. Since $\underline{v} \leq \overline{v}$, by definition, this implies that $\underline{v} = \overline{v} = v^0$. Since v^ϵ decreases to v^0 as $\epsilon \searrow 0$, uniform convergence on compact sets follows from Dini's theorem.

To obtain the convergence of $\partial_e v^\epsilon$ towards $\partial_e v$, we use the smoothness of v^ϵ (implied by classical uniform parabolic regularity) to see that $\partial_e v^\epsilon$ is a classical solution of the equation

$$-\partial_t(\partial_e v^\epsilon) + F^\epsilon\big(e, \partial_e v^\epsilon, D(\partial_e v^\epsilon), D^2(\partial_e v^\epsilon)\big) - \big(b - \epsilon f_0'(v^\epsilon)\partial_e v^\epsilon\big)\partial_e v^\epsilon - \sigma^2 e D(\partial_e v^\epsilon) = 0,$$

and we proceed as above. □

Lemma 4.

$$\lim_{\epsilon\searrow 0} V^\epsilon = V^0 \textit{ and } \quad \lim_{\epsilon\searrow 0} \partial_e V^\epsilon = \partial_e V^0, \tag{61}$$

uniformly on compact subsets of $[0,\tau) \times \mathbb{R}_+$ *and for each* $(t,e) \in [0,\tau) \times \mathbb{R}_+$

$$\partial_e V^0(t,e) = \mathbb{E}_{t,e}\left[\frac{E^0_\tau}{E^0_t}\mathbf{1}_{[K,\infty)}(v^0(\tau, E^0_\tau))\partial_e v^0(\tau, E^0_\tau)\right]$$

Proof. We argue as above by taking the limit $\epsilon \searrow 0$ in the viscosity sense. In the present case V^ϵ is a classical solution of

$$-\partial_t V^\epsilon(t,e) + G^\epsilon\left(t, e, \partial_e V^\epsilon(t,e), \partial^2_{ee} V^\epsilon(t,e)\right) = 0$$

where

$$G^\epsilon(t,e,p,\gamma) := \big[be - \epsilon f_0\big(v^\epsilon(t,e)\big)\big]p - \frac{1}{2}\sigma^2 e^2 \gamma.$$

Since $v^\epsilon \longrightarrow v^0$ uniformly on compact sets, and f_0 is monotone and continuous, the proof of Lemma 3 above applies in the present situation.

The convergence result for $\partial_e V^\epsilon$ is obtained by first differentiating the equation satisfied by V^ϵ, which is justified by classical parabolic regularity, and then using the same argument as above. Notice that the expression of $\partial_e V^0$ is, as expected, obtained by differentiating V^0 in (56) inside the expectation operator. □

In preparation for the next result, for $\epsilon > 0$ we define

$$u^\epsilon(t,e) := \frac{v^\epsilon(t,e) - v^0(t,e)}{\epsilon} \tag{62}$$

Lemma 5.
$$\lim_{\epsilon\searrow 0} u^\epsilon(t,e) = u^0(t,e) := \mathbb{E}_{t,e}\left[\int_t^T f_0(v^0(s,E_s^0))\partial_e v^0(s,E_s^0)ds\right],$$
uniformly on compact subsets of $[0,T)\times\mathbb{R}_+$.

Proof. Since v^ϵ is a classical solution of equation (53), plain computations show that u^ϵ is a classical solution of the equation
$$u^\epsilon(T,.) = 0,\ -\partial_t u^\epsilon + L^\epsilon\left(t,e,Du^\epsilon,D^2u^\epsilon\right) = 0,$$
where L is the linear operator defined by:
$$L^\epsilon(t,e,p,\gamma) := -bep + f_0\big(v^\epsilon(t,e)\big)\partial_e v^\epsilon(t,e) - \frac{1}{2}\sigma^2e^2\gamma.$$
Using the stability result of viscosity solutions together with the Feynman-Kac representation as in the proof of Lemma 3 above, the convergence result of Lemma 3 provides the limit equation:
$$u^0(T,.) = 0,\ \partial_t u^0(t,e) + beDu^0(t,e) + \frac{1}{2}\sigma^2e^2D^2u^0(t,e) = f_0(v^0(t,e))\partial_e v^0(t,e) = 0.$$
The representation of the solution u^0 as the expectation appearing in the statement of the lemma is given by the Feynman-Kac as long as we can show that the expectation makes sense. This is indeed the case since f_0 is nondecreasing and continuous, $v^0 \le \lambda$, so that we have:
$$\mathbb{E}_{t,e}\left[\int_t^T \left|f_0(v^0(s,E_s^0))\partial_e v^0(s,E_s^0)\right|ds\right] \ \le\ f_0(\lambda)\mathbb{E}_{t,e}\left[\int_t^T |\partial_e v^0(s,E_s^0)|\,ds\right]. \tag{63}$$
Now, observe that by Itô's formula:
$$\begin{aligned}\mathbb{E}_{t,e}\left[v^0(T,E_T)^2 - v^0(t,e)^2\right] &= 2\mathbb{E}_{t,e}\left[\int_t^T v^0(s,E_s)\left(\partial_t v^0 + be\partial_e v^0 + \frac{1}{2}\sigma^2e^2\partial^2_{ee}v^0\right)(s,E_s)ds\right]\\ &\qquad + \mathbb{E}_{t,e}\left[\int_t^T |\partial_e v^0(s,E_s)|^2\sigma^2E_s^2ds\right]\\ &= \mathbb{E}_{t,e}\left[\int_t^T |\partial_e v^0(s,E_s)|^2\sigma^2E_s^2ds\right]\end{aligned}$$
because of the PDE satisfied by v^0. Then $\mathbb{E}_{t,e}\left[\int_t^T |\partial_e v(s,E_s)|^2\sigma^2E_s^2ds\right] \le 2\lambda^2$, implying that (63) is finite by the Cauchy-Schwarz inequality. □

Proof of Proposition 10. For each $\epsilon > 0$ we define:
$$U^\epsilon(t,e) := \frac{V^\epsilon(t,e) - V^0(t,e)}{\epsilon}.$$
Since V^ϵ is a classical solution of (54), it follows that U^ϵ is a classical solution of:
$$-\partial_t U^\epsilon - H^\epsilon\left(t,e,DU^\epsilon,D^2U^\epsilon\right) = 0, \quad \text{on } [0,\tau)\times\mathbb{R}_+$$
with
$$H^\epsilon(t,e,p,\gamma) := -bep - \frac{1}{2}\sigma^2e^2\gamma + f_0(v^\epsilon(t,e)\partial_e V^\epsilon(t,e) \quad \text{for } (t,e)\in[0,\tau)\times\mathbb{R}_+$$

satisfying the terminal condition

$$U^\epsilon(\tau,.) = \frac{(v^\epsilon(\tau,.)-K)^+ - (v^0(\tau,.)-K)^+}{\epsilon}.$$

Using the convergence results of lemmas 3, 4 and 5, together with the stability of viscosity solutions and the Feynlan Kac representation for the limiting linear PDE, we see that $U^\epsilon \longrightarrow U^0$ uniformly on compacts, where U^0 is the unique solution of

$$-\partial_t U^0 - be\partial_e U^0 - \frac{1}{2}\sigma^2 e^2 \partial^2_{ee} U^0 - f_0(v^0)\partial_e V^0 = 0 \quad \text{for } (t,e)\in[0,\tau)\times\mathbb{R}_+$$

satisfying the terminal condition

$$U^0(\tau,.) = u^0(\tau,.)\mathbf{1}_{[K,\infty)}(v^0(\tau,.)).$$

By the Feynman-Kac representation of such a solution, we have:

$$U^0(t,e) \;=\; \mathbb{E}_{t,e}\left[\int_t^\tau f_0(v^0(s,E^0_s))\partial_e V^0(s,E^0_s)ds + u^0(\tau,E^0_\tau)\mathbf{1}_{[K,\infty)}(v^0(\tau,E^0_\tau))\right],$$

and the required result is obtained by replacing $\partial_e V^0$ and u^0 by their expressions from Lemmas 4 and 5 respectively, and using the tower property of conditional expectations. □

6. Numerical results

In this section we provide the following numerical evidence of the accuracy of the small abatement asymptotic formula derived above:

(1) We compute numerically v^ϵ with high accuracy, and we then compute values of V^ϵ using the values of v^ϵ so computed. We used an explicit finite difference monotone scheme (see for example [1]. The results are reproduced in Figure 1 We plotted V^ϵ against v^ϵ in order to show how the option price depends upon the value of the underlying allowance.
(2) We compare the previous numerical results with the first order Taylor expansion which can easily be computed as it only involves the Monte Carlo computation of an expectation.

References

[1] G. Barles and P.E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Analysis*, 4:271–283, 1991.

[2] R. Carmona, F. Fehr, J. Hinz, and A. Porchet. Market designs for emissions trading schemes. *SIAM Review*, 52:403–452, 2010.

[3] R. Carmona, M. Fehr, and J. Hinz. Optimal stochastic control and carbon price formation. *SIAM Journal on Control and Optimization*, 48(12):2168–2190, 2009.

[4] R. Carmona and J. Hinz. Risk neutral modeling of emission allowance prices and option valuation. Technical report, Princeton University, 2009.

[5] M Chesney and L. Taschini. The endogenous price dynamics of the emission allowances: An application to co2 option pricing. Technical report, 2008.

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

[7] F. Delarue. Estimates of the solutions of a system of quasi-linear PDEs. A probabilistic scheme. *Séminaire de Probabilités*, XXXVII, 290–332, 2003.

[8] F. Delarue and G. Guatteri. Weak Existence and Uniqueness for FBSDEs. *Stochastic Processes and Applications*, 116:1712–1742, 2006.

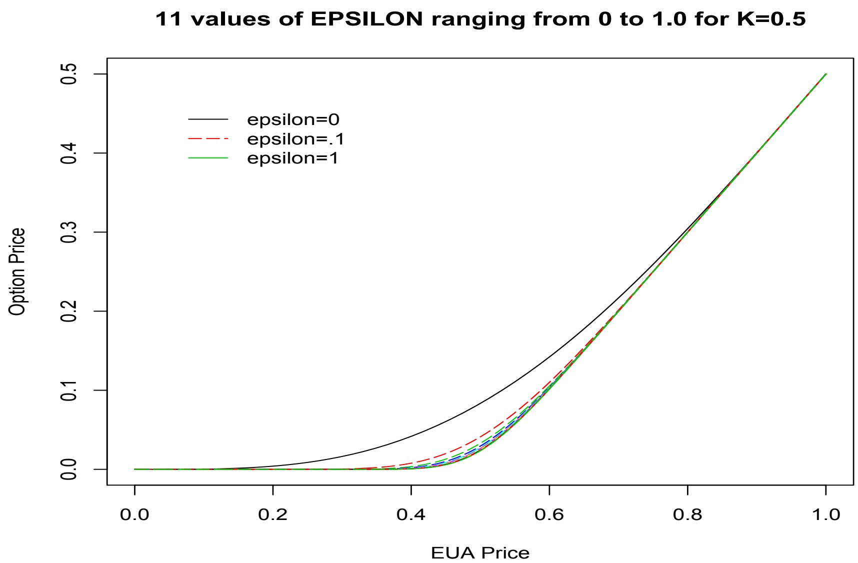

FIGURE 1. European call option prices for $\epsilon = 0, 0.1, 0.2, \cdots, 0.9, 1$.

[9] N. Krylov. *Lectures on Elliptic and Parabolic Equations in Hölder Spaces*, volume 12 of *Graduate Studies in Mathematics*. American Mathematical Society, 1996.

[10] H. Kunita. *Stochastic Flows and Stochastic Differential Equations*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1990.

[11] P. Lax *Hyperbolic Differential Equations*. Courant Lecture Notes. American Mathematical Society, 2006.

[12] J. Ma and J. Yong. *Forward-Backward Stochastic Differential Equations and their Applications*, volume 1702 of *Lecture Notes in Mathematics*. Springer Verlag.

[13] J. Ma, Z. Wu, D. Zhang and J. Zhang. On wellposedness of forward-backward SDEs – a unified approach. *Preprint*.

[14] D. Nualart. *The Malliavin Calculus and Related Topics*. Probability and its Applications. Springer Verlag, 1995.

[15] S. Peng and Z. Wu. Fully Coupled Forward-Backward Stochastic Differential Equations and Applications to Optimal Control. *SIAM Journal on Control and Optimization*, 37:825–843, 1999.

[16] H. Pham. *Continuous-time Stochastic Control and Optimization with Financial Applications*. Stochastic Modelling and Applied Probability. Springer Verlag, 2009.

[17] L.C.G. Rogers and D. Williams *Diffusions, Markov Processes and Martingales. Volume 2. 2nd Edition.* Cambridge University Press, 1994.

[18] J. Seifert, M. Uhrig-Homburg, and M. Wagner. Dynamic behavior of carbon spot prices .theory and empirical evidence. *Journal of Environmental Economics and Management*, 56:180–194, 2008.

[19] Z. Wu and M. Xu. Comparison theorems for forward backward sdes. *Statistics and Probability Letters*, 79:426–435, 2009.

ORFE, BENDHEIM CENTER FOR FINANCE, PRINCETON UNIVERSITY, PRINCETON, NJ 08544, USA.
E-mail address: rcarmona@princeton.edu

LABORATOIRE J.-A. DIEUDONNÉ, UNIVERSITÉ DE NICE SOPHIA-ANTIPOLIS, PARC VALROSE, 06108 CEDEX 02, NICE, FRANCE
E-mail address: francois.delarue@unice.fr

CENTRE DE MATHÉMATIQUES APPLIQUÉES, ECOLE POLYTECHNIQUE, UMR CNRS 7641, 91128 PALAISEAU CEDEX, FRANCE
E-mail address: gilles-edouard.espinosa@polytechnique.org

CENTRE DE MATHÉMATIQUES APPLIQUÉES, ECOLE POLYTECHNIQUE, UMR CNRS 7641, 91128 PALAISEAU CEDEX, FRANCE
E-mail address: nizar.touzi@polytechnique.edu