Variational formulation of American option prices in the Heston Model

Damien Lamberton, Giulia Terenzi

Variational formulation of American option prices in the Heston Model

Damien Lamberton*
Giulia Terenzi†

Abstract

We give an analytical characterization of the price function of an American option in Heston-type models. Our approach is based on variational inequalities and extends recent results of Daskalopoulos and Feehan (2011, 2016) and Feehan and Pop (2015). We study the existence and uniqueness of a weak solution of the associated degenerate parabolic obstacle problem. Then, we use suitable estimates on the joint distribution of the log-price process and the volatility process in order to characterize the analytical weak solution as the solution to the optimal stopping problem. We also rely on semi-group techniques and on the affine property of the model.

Keywords: American options; degenerate parabolic obstacle problem; optimal stopping problem.

1 Introduction

The model introduced by S. Heston in 1993 ([11]) is one of the most widely used stochastic volatility models in the financial world and it was the starting point for several more complex models which extend it. The great success of the Heston model is due to the fact that the dynamics of the underlying asset can take into account the non-lognormal distribution of the asset returns and the observed mean-reverting property of the volatility. Moreover, it remains analytically tractable and provides a closed-form valuation formula for European options using Fourier transform.

These features have called for an extensive literature on numerical methods to price derivatives in Heston-type models. In this framework, besides purely probabilistic methods such as standard Monte Carlo and tree approximations, there is a large class of algorithms which exploit numerical analysis techniques in order to solve the standard PDE (resp. the obstacle problem) formally associated with the European (resp. American) option price function. However, these algorithms have, in general, little mathematical support and in particular, as far as we know, a rigorous and complete study of the analytic characterization of the American price function is not present in the literature.

The main difficulties in this sense come from the degenerate nature of the model. In fact, the infinitesimal generator associated with the two dimensional diffusion given by the log-price process and the volatility process is not uniformly elliptic: it degenerates on the boundary of the domain, that is when the volatility variable vanishes. Moreover, it has unbounded coefficients with linear growth. Therefore, the existence and the uniqueness of the solution to the pricing PDE and obstacle problem do not follow from the classical theory, at least in the case in which the boundary of the state space is reached with positive probability, as happens in many cases of practical importance (see [3]). Moreover, the probabilistic representation of the solution, that is the identification with the price function, is far from trivial in the case of non regular payoffs.

*Université Paris-Est, Laboratoire d'Analyse et de Mathématiques Appliquées (UMR 8050), UPEM, UPEC, CNRS, Projet Mathrisk INRIA, F-77454, Marne-la-Vallée, France - damien.lamberton@u-pem.fr

†Université Paris-Est, Laboratoire d'Analyse et de Mathématiques Appliquées (UMR 8050), UPEM, UPEC, CNRS, Projet Mathrisk INRIA, F-77454, Marne-la-Vallée, France, and Università di Roma Tor Vergata, Dipartimento di Matematica, Italy - terenzi@mat.uniroma2.it

It should be emphasized that a clear analytic characterization of the price function allows not only to formally justify the theoretical convergence of some classical pricing algorithms but also to investigate the regularity properties of the price function (see [13] for the case of the Black and Scholes models).

Concerning the existing literature, E. Ekstrom and J. Tysk in [7] give a rigorous and complete analysis of these issues in the case of European options, proving that, under some regularity assumptions on the payoff functions, the price function is the unique classical solution of the associated PDE with a certain boundary behaviour for vanishing values of the volatility. However, the payoff functions they consider do not include the case of standard put and call options.

Recently, P. Daskalopoulos and P. Feehan studied the existence, the uniqueness, and some regularity properties of the solution of this kind of degenerate PDE and obstacle problems in the elliptic case, introducing suitable weighted Sobolev spaces which clarify the behaviour of the solution near the degenerate boundary (see [5, 6]). In another paper ([8]) P. Feehan and C. Pop addressed the issue of the probabilistic representation of the solution, but we do not know if their assumptions on the solution of the parabolic obstacle problem are satisfied in the case of standard American options. Note that Feehan and Pop did prove regularity results in the elliptic case, see [9]. They also announce results for the parabolic case in [8].

The aim of this paper is to give a precise analytical characterization of the American option price function for a large class of payoffs which includes the standard put and call options. In particular, we give a variational formulation of the American pricing problem using the weighted Sobolev spaces and the bilinear form introduced in [5]. The paper is organized as follows. In Section 2, we introduce our notations and we state our main results. Then, in Section 3, we study the existence and uniqueness of the solution of the associated variational inequality, extending the results obtained in [5] in the elliptic case. The proof relies, as in [5], on the classical penalization technique introduced by Bensoussan and Lions [4] with some technical devices due to the degenerate nature of the problem. We also establish a Comparison Theorem. Finally, in section 4, we prove that the solution of the variational inequality with obstacle function ψ is actually the American option price function with payoff ψ, with conditions on ψ which are satisfied, for example, by the standard call and put options. In order to do this, we use the affine property of the underlying diffusion given by the log price process X and the volatility process Y. Thanks to this property, we first identify the analytic semigroup associated with the bilinear form with a correction term and the transition semigroup of the pair (X, Y) with a killing term. Then, we prove regularity results on the solution of the variational inequality and suitable estimates on the joint law of the process (X, Y) and we deduce from them the analytical characterization of the solution of the optimal stopping problem, that is the American option price.

Contents

2 Notations and main results

2.1 The Heston model

We recall that in the Heston model the dynamics under the pricing measure of the asset price S and the volatility process Y are governed by the stochastic differential equation system

$$\begin{cases} \frac{dS_t}{S_t} = (r-\delta)dt + \sqrt{Y_t}dB_t, & S_0 = s > 0, \\ dY_t = \kappa(\theta - Y_t)dt + \sigma\sqrt{Y_t}dW_t, & Y_0 = y \geq 0, \end{cases}$$

where B and W denote two correlated Brownian motions with

$$d\langle B, W\rangle_t = \rho dt, \qquad \rho \in (-1, 1).$$

We exclude the degenerate case $\rho = \pm 1$, that is the case in which the same Brownian motion drives the dynamics of X and Y. Actually, it can be easily seen that, in this case, S_t reduces to a function of the pair $(Y_t, \int_0^t Y_s ds)$ and the resulting degenerate model cannot be treated with the techniques we develop in this paper. Moreover, this particular situation is not very interesting from a financial point of view.

Here $r \geq 0$ and $\delta \geq 0$ are respectively the risk free rate of interest and the continuous dividend rate. The dynamics of Y follows a CIR process with mean reversion rate $\kappa > 0$ and long run state $\theta > 0$. The parameter $\sigma > 0$ is called the volatility of the volatility. Note that we do not require the Feller condition $2\kappa\theta \geq \sigma^2$: the volatility process Y can hit 0 (see, for example, [2, Section 1.2.4]).

We are interested in studying the price of an American option with payoff function ψ. For technical reasons which will be clarified later on, hereafter we consider the process

$$X_t = \log S_t - \bar{c}t, \qquad \text{with } \bar{c} = r - \delta - \frac{\rho\kappa\theta}{\sigma}, \tag{2.1}$$

which satisfies

$$\begin{cases} dX_t = \left(\frac{\rho\kappa\theta}{\sigma} - \frac{Y_t}{2}\right)dt + \sqrt{Y_t}dB_t, \\ dY_t = \kappa(\theta - Y_t)dt + \sigma\sqrt{Y_t}dW_t. \end{cases} \tag{2.2}$$

Note that, in this framework, we have to consider payoff functions ψ which depend on both the time and the space variables. For example, in the case of a standard put option (resp. a call option) with strike price K we have $\psi(t,x) = (K - e^{x+\bar{c}t})_+$ (resp. $\psi(t,x) = (e^{x+\bar{c}t} - K)_+$). So, the natural price at time t of an American option with a nice enough payoff $(\psi(t, X_t, Y_t))_{0\leq t\leq T}$ is given by $P(t, X_t, Y_t)$, with

$$P(t,x,y) = \sup_{\theta\in\mathcal{T}_{t,T}} \mathbb{E}[e^{-r(\theta-t)}\psi(\theta, X_\theta^{t,x,y}, Y_\theta^{t,y})],$$

where $\mathcal{T}_{t,T}$ is the set of all stopping times with values in $[t,T]$ and $(X_s^{t,x,y}, Y_s^{t,y})_{t\leq s\leq T}$ denotes the solution to (2.2) with the starting condition $(X_t, Y_t) = (x, y)$.

Our aim is to give an analytical characterization of the price function P. We recall that the infinitesimal generator of the two dimensional diffusion (X, Y) is given by

$$\mathcal{L} = \frac{y}{2}\left(\frac{\partial^2}{\partial x^2} + 2\rho\sigma\frac{\partial^2}{\partial y\partial x} + \sigma^2\frac{\partial^2}{\partial y^2}\right) + \left(\frac{\rho\kappa\theta}{\sigma} - \frac{y}{2}\right)\frac{\partial}{\partial x} + \kappa(\theta - y)\frac{\partial}{\partial y},$$

which is defined on the open set $\mathcal{O} := \mathbb{R}\times(0,\infty)$. Note that $\mathcal{L}$ has unbounded coefficients and is not uniformly elliptic: it degenerates on the boundary $\partial\mathcal{O} = \mathbb{R}\times\{0\}$.

2.2 American options and variational inequalities

2.2.1 Heuristics

From the optimal stopping theory, we know that the discounted price process $\tilde{P}(t, X_t, Y_t) = e^{-rt}P(t, X_t, Y_t)$ is a supermartingale and that its finite variation part only decreases on the set $P = \psi$ with respect to the time variable t. We want to have an analytical interpretation of these features on the function $P(t, x, y)$. So, assume that $P \in \mathcal{C}^{1,2}((0, T) \times \mathcal{O})$. Then, by applying Itô's formula, the finite variation part of $\tilde{P}(t, X_t, Y_t)$ is

$$\left(\frac{\partial \tilde{P}}{\partial t} + \mathcal{L}\tilde{P}\right)(t, X_t, Y_t).$$

Since $\tilde{P}$ is a supermartingale, we can deduce the inequality

$$\frac{\partial \tilde{P}}{\partial t} + \mathcal{L}\tilde{P} \leq 0$$

and, since its finite variation part decreases only on the set $P(t, X_t, Y_t) = \psi(t, X_t, Y_t)$, we can write

$$\left(\frac{\partial \tilde{P}}{\partial t} + \mathcal{L}\tilde{P}\right)(\psi - P) = 0.$$

This relation has to be satisfied $dt - a.e.$ along the trajectories of (t, X_t, Y_t). Moreover, we have the two trivial conditions $P(T, x, y) = \psi(T, x, y)$ and $P \geq \psi$.

The previous discussion is only heuristic, since the price function P is not regular enough to apply Itô's formula. However, it suggests the following strategy:

1. Study the obstacle problem

$$\begin{cases} \frac{\partial u}{\partial t} + \mathcal{L}u \leq 0, \qquad u \geq \psi, & in\ [0, T] \times \mathcal{O}, \\ \left(\frac{\partial u}{\partial t} + \mathcal{L}u\right)(\psi - u) = 0, & in\ [0, T] \times \mathcal{O}, \\ u(T, x, y) = \psi(T, x, y). \end{cases} \tag{2.3}$$

2. Show that the discounted price function $\tilde{P}$ is equal to the solution of (2.3) where ψ is replaced by $\tilde{\psi}(t, x, y) = e^{-rt}\psi(t, x, y)$.

We will follow this program providing a variational formulation of system (2.3).

2.2.2 Weighted Sobolev spaces and bilinear form associated with the Heston operator

We consider the measure first introduced in [5]:

$$\mathfrak{m}_{\gamma,\mu}(dx, dy) = y^{\beta-1}e^{-\gamma|x|-\mu y}dxdy,$$

with $\gamma > 0,\ \mu > 0$ and $\beta := \frac{2\kappa\theta}{\sigma^2}$.

It is worth noting that in [5] the authors fix $\mu = \frac{2\kappa}{\sigma^2}$ in the definition of the measure $\mathfrak{m}_{\gamma,\mu}$. This specification will not be necessary in this paper, but it is useful to mention it in order to better understand how this measure arises. In fact, recall that the density of the speed measure of the CIR process is given by $y^{\beta-1}e^{-\frac{2\kappa}{\sigma^2}y}$. Then, the term $y^{\beta-1}e^{-\frac{2\kappa}{\sigma^2}y}$ in the definition of $\mathfrak{m}_{\gamma,\mu}$ has a clear probabilistic interpretation, while the exponential term $e^{-\gamma|x|}$ is classically introduced just to deal with the unbounded domain in the $x-$component.

For $u \in \mathbb{R}^n$ we denote by $|u|$ the standard Euclidean norm of u in $\mathbb{R}^n$. Then, we recall the weighted Sobolev spaces introduced in [5]. The choice of these particular Sobolev spaces will allow us to formulate the obstacle problem (2.3) in a variational framework with respect to the measure $\mathfrak{m}_{\gamma,\mu}$.

Definition 2.1. *For every $p \geq 1$, let $L^p(\mathcal{O}, \mathfrak{m}_{\gamma,\mu})$ be the space of all Borel measurable functions $u : \mathcal{O} \to \mathbb{R}$ for which*

$$\|u\|^p_{L^p(\mathcal{O},\mathfrak{m}_{\gamma,\mu})} := \int_{\mathcal{O}} |u|^p d\mathfrak{m}_{\gamma,\mu} < \infty,$$

and denote $H^0(\mathcal{O}, \mathfrak{m}_{\gamma,\mu}) := L^2(\mathcal{O}, \mathfrak{m}_{\gamma,\mu})$.

1. *If $\nabla u := (u_x, u_y)$ and u_x, u_y are defined in the sense of distributions, we set*

$$H^1(\mathcal{O}, \mathfrak{m}_{\gamma,\mu}) := \{u \in L^2(\mathcal{O}, \mathfrak{m}_{\gamma,\mu}) : \sqrt{1+y}u \text{ and } \sqrt{y}|\nabla u| \in L^2(\mathcal{O}, \mathfrak{m}_{\gamma,\mu})\},$$

and

$$\|u\|^2_{H^1(\mathcal{O},\mathfrak{m}_{\gamma,\mu})} := \int_{\mathcal{O}} \left(y|\nabla u|^2 + (1+y)u^2\right) d\mathfrak{m}_{\gamma,\mu}.$$

2. *If $D^2u := (u_{xx}, u_{xy}, u_{yx}, u_{yy})$ and all derivatives of u are defined in the sense of distributions, we set*

$$H^2(\mathcal{O}, \mathfrak{m}_{\gamma,\mu}) := \{u \in L^2(\mathcal{O}, \mathfrak{m}_{\gamma,\mu}) : \sqrt{1+y}u,\ (1+y)|\nabla u|,\ y|D^2u| \in L^2(\mathcal{O}, \mathfrak{m}_{\gamma,\mu})\}$$

and

$$\|u\|^2_{H^2(\mathcal{O},\mathfrak{m}_{\gamma,\mu})} := \int_{\mathcal{O}} \left(y^2|D^2u|^2 + (1+y)^2|\nabla u|^2 + (1+y)u^2\right) d\mathfrak{m}_{\gamma,\mu}.$$

For brevity and when the context is clear, we shall often denote

$$H := H^0(\mathcal{O}, \mathfrak{m}_{\gamma,\mu}), \qquad V := H^1(\mathcal{O}, \mathfrak{m}_{\gamma,\mu})$$

and

$$\|u\|_H := \|u\|_{L^2(\mathcal{O},\mathfrak{m}_{\gamma,\mu})}, \qquad \|u\|_V := \|u\|_{H^1(\mathcal{O},\mathfrak{m}_{\gamma,\mu})}.$$

Note that we have the inclusion

$$H^2(\mathcal{O}, \mathfrak{m}_{\gamma,\mu}) \subset H^1(\mathcal{O}, \mathfrak{m}_{\gamma,\mu})$$

and that the spaces $H^k(\mathcal{O}, \mathfrak{m}_{\gamma,\mu})$, for $k = 0, 1, 2$ are Hilbert spaces with the inner products

$$(u, v)_H = (u, v)_{L^2(\mathcal{O},\mathfrak{m}_{\gamma,\mu})} = \int_{\mathcal{O}} uv d\mathfrak{m}_{\gamma,\mu},$$

$$(u, v)_V = (u, v)_{H^1(\mathcal{O},\mathfrak{m}_{\gamma,\mu})} = \int_{\mathcal{O}} \left(y\left(\nabla u, \nabla v\right) + (1+y)uv\right) d\mathfrak{m}_{\gamma,\mu}$$

and

$$(u, v)_{H^2(\mathcal{O},\mathfrak{m}_{\gamma,\mu})} := \int_{\mathcal{O}} \left(y^2\left(D^2u, D^2v\right) + (1+y)^2\left(\nabla u, \nabla v\right) + (1+y)uv\right) d\mathfrak{m}_{\gamma,\mu},$$

where $(\cdot,\cdot)$ denotes the standard scalar product in $\mathbb{R}^n$.

Moreover, for every $T > 0$, $p \in [1, +\infty)$ and $i = 0, 1, 2$, we set

$$L^p([0,T]; H^i(\mathcal{O}, \mathfrak{m}_{\gamma,\mu})) = \Big\{u : [0,T] \times \mathcal{O} \to \mathbb{R} \text{ Borel measurable} : u(t,\cdot,\cdot) \in H^i(\mathcal{O}, \mathfrak{m}_{\gamma,\mu}) \text{ for a.e. } t \in [0,T]$$
$$\text{and } \int_0^T \|u(t,\cdot\cdot)\|^p_{H^i(\mathcal{O},\mathfrak{m}_{\gamma,\mu})} dt < \infty\Big\}$$

and

$$\|u\|^p_{L^p([0,T];H^i(\mathcal{O},\mathfrak{m}_{\gamma,\mu}))} = \int_0^T \|u(t,\cdot\cdot)\|^p_{H^i(\mathcal{O},\mathfrak{m}_{\gamma,\mu})} dt.$$

We also define $L^\infty([0,T]; H^i)$ with the usual essential sup norm.

We can now introduce the following bilinear form.

Definition 2.2. *For any $u, v \in H^1(\mathcal{O}, \mathfrak{m}_{\gamma,\mu})$ we define the bilinear form*

$$\begin{aligned} a_{\gamma,\mu}(u, v) =& \frac{1}{2}\int_{\mathcal{O}} y\left(u_x v_x(x,y) + \rho\sigma u_x v_y(x,y) + \rho\sigma u_y v_x(x,y) + \sigma^2 u_y v_y(x,y)\right) d\mathfrak{m}_{\gamma,\mu} \\ &+ \int_{\mathcal{O}} y\left(j_{\gamma,\mu}(x)u_x(x,y) + k_{\gamma,\mu}(x)u_y(x,y)\right) v(x,y) d\mathfrak{m}_{\gamma,\mu}, \end{aligned}$$

where

$$j_{\gamma,\mu}(x) = \frac{1}{2}\left(1 - \gamma sgn(x) - \mu\rho\sigma\right), \qquad k_{\gamma,\mu}(x) = \kappa - \frac{\gamma\rho\sigma}{2} sgn(x) - \frac{\mu\sigma^2}{2}. \tag{2.4}$$

We will prove that $a_{\gamma,\mu}$ is the bilinear form associated with the operator $\mathcal{L}$, in the sense that for every $u \in H^2(\mathcal{O}, \mathfrak{m}_{\gamma,\mu})$ and for every $v \in H^1(\mathcal{O}, \mathfrak{m}_{\gamma,\mu})$, we have

$$(\mathcal{L}u, v)_H = -a_{\gamma,\mu}(u, v).$$

In order to simplify the notation, for the rest of this paper we will write $\mathfrak{m}$ and $a(\cdot,\cdot)$ instead of $\mathfrak{m}_{\gamma,\mu}$ and $a_{\gamma,\mu}(\cdot,\cdot)$ every time the dependence on γ and μ does not play a role in the analysis and computations.

2.3 Variational formulation of the American price

Fix $T > 0$. We consider an assumption on the payoff function ψ which will be crucial in the discussion of the penalized problem.

Assumption $\mathcal{H}^1$. We say that a function ψ satisfies Assumption $\mathcal{H}^1$ if $\psi \in \mathcal{C}([0,T]; H)$, $\sqrt{1+y}\psi \in L^2([0,T]; V)$, $\psi(T) \in V$ and there exists $\Psi \in L^2([0,T]; V)$ such that $\left|\frac{\partial \psi}{\partial t}\right| \leq \Psi$.

We will also need a domination condition on ψ by a function Φ which satisfies the following assumption.

Assumption $\mathcal{H}^2$. We say that a function $\Phi \in L^2([0,T]; H^2(\mathcal{O}, \mathfrak{m}))$ satisfies Assumption $\mathcal{H}^2$ if $(1+y)^{\frac{3}{2}}\Phi \in L^2([0,T]; H)$, $\frac{\partial \Phi}{\partial t} + \mathcal{L}\Phi \leq 0$ and $\sqrt{1+y}\Phi \in L^\infty([0,T]; L^2(\mathcal{O}, \mathfrak{m}_{\gamma,\mu'}))$ for some $0 < \mu' < \mu$.

The domination condition is needed to deal with the lack of coercivity of the bilinear form associated with our problem. Similar conditions are also used in [5].

The first step in the variational formulation of the problem is to introduce the associated variational inequality and to prove the following existence and uniqueness result.

Theorem 2.3. *Assume that ψ satisfies Assumption $\mathcal{H}^1$ together with $0 \leq \psi \leq \Phi$, where Φ satisfies Assumption $\mathcal{H}^2$. Then, there exists a unique function u such that $u \in \mathcal{C}([0,T]; H) \cap L^2([0,T]; V)$, $\frac{\partial u}{\partial t} \in L^2([0,T]; H)$ and*

$$\begin{cases} -\left(\frac{\partial u}{\partial t}, v-u\right)_H + a(u, v-u) \geq 0, \quad \text{a.e. in } [0,T] \quad v \in L^2([0,T]; V),\ v \geq \psi, \\ u \geq \psi \text{ a.e. in } [0,T] \times \mathbb{R} \times (0,\infty), \\ u(T) = \psi(T), \\ 0 \leq u \leq \Phi. \end{cases} \tag{2.5}$$

The proof is presented in Section 3 and essentially relies on the penalization technique introduced by Bensoussan and Lions (see also [10]) with some technical devices due to the degenerate nature of the problem. We extend in the parabolic framework the results obtained in [5] for the elliptic case.

The second step is to identify the unique solution of the variational inequality (2.5) as the solution of the optimal stopping problem, that is the (discounted) American option price. In order to do this, we consider the following assumption on the payoff function.

Assumption $\mathcal{H}^*$. We say that a function $\psi : [0,T] \times \mathbb{R} \times [0,\infty) \to \mathbb{R}$ satisfies Assumption $\mathcal{H}^*$ if ψ is continuous and there exist constants $C > 0$ and $L \in \left[0, \frac{2\kappa}{\sigma^2}\right)$ such that, for all $(t,x,y) \in [0,T] \times \mathbb{R} \times [0,\infty)$,

$$0 \leq \psi(t,x,y) \leq C(e^x + e^{Ly}), \tag{2.6}$$

and

$$\left|\frac{\partial \psi}{\partial t}(t,x,y)\right| + \left|\frac{\partial \psi}{\partial x}(t,x,y)\right| + \left|\frac{\partial \psi}{\partial y}(t,x,y)\right| \leq C(e^{a|x|+by}), \tag{2.7}$$

for some $a, b \in \mathbb{R}$.

Note that the payoff functions of a standard call and put option with strike price K (that is, respectively, $\psi = \psi(t,x) = (K - e^{x+\bar{c}t})_+$ and $\psi = \psi(t,x) = (e^{x+\bar{c}t} - K)_+$) satisfy Assumption $\mathcal{H}^*$. Moreover, it is easy to see that, if ψ satisfies Assumption $\mathcal{H}^*$, then it is possible to choose γ and μ in the definition of the measure $\mathfrak{m}_{\gamma,\mu}$ (see (2.2.2)) such that ψ satisfies the assumptions of Theorem 2.3. Then, for such γ and μ, we get the following identification result.

Theorem 2.4. *Assume that ψ satisfies Assumption $\mathcal{H}^*$ Then, the solution u of the variational inequality* (2.5) *associated with ψ is continuous and coincides with the function u^* defined by*

$$u^*(t,x,y) = \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}\left[\psi(\tau, X_\tau^{t,x,y}, Y_\tau^{t,y})\right].$$

3 Existence and uniqueness of solutions to the variational inequality

3.1 Integration by parts and energy estimates

The following result justifies the definition of the bilinear form a.

Proposition 3.1. *If $u \in H^2(\mathcal{O}, \mathfrak{m})$ and $v \in H^1(\mathcal{O}, \mathfrak{m})$, we have*

$$(\mathcal{L}u, v)_H = -a(u, v). \tag{3.1}$$

This result is proved with the same arguments of [5, Lemma 2.23] or [6, Lemma A.3] but we prefer to repeat here the proof since it clarifies why we have considered the process $X_t = \log S_t - \bar{c}t$ instead of the standard log-price process $\log S_t$.

We first need the following result which justifies the integration by parts formulas with respect to the measure $\mathfrak{m}$. The proof follows standard approximation techniques, so we omit it (see the proof of [5, Lemma 2.23] or [17] for the details).

Lemma 3.2. *Let us consider $u, v : \mathcal{O} \to \mathbb{R}$ locally square-integrable on $\mathcal{O}$. Then, if the derivatives u_x and v_x are locally square-integrable on $\mathcal{O}$ and*

$$\int_{\mathcal{O}} \big(|u_x(x,y)v(x,y)| + |u(x,y)v_x(x,y)| + |u(x,y)v(x,y)|\big) d\mathfrak{m} < \infty,$$

we have

$$\int_{\mathcal{O}} u_x(x,y)v(x,y)d\mathfrak{m} = -\int_{\mathcal{O}} u(x,y)\left(v_x(x,y) - \gamma sgn(x)v\right) d\mathfrak{m}.$$

Similarly, if the derivatives u_y and v_y are locally square-integrable on $\mathcal{O}$ and

$$\int_{\mathcal{O}} y\big(|u_y(x,y)v(x,y)| + |u(x,y)v_xy(x,y)|\big) + |u(x,y)v(x,y)|d\mathfrak{m} < \infty,$$

we have

$$\int_{\mathcal{O}} yu_y(x,y)v(x,y)d\mathfrak{m} = -\int_{\mathcal{O}} yu(x,y)v_y(x,y)d\mathfrak{m} - \int_{\mathcal{O}} (\beta - \mu y)u(x,y)v(x,y)d\mathfrak{m}.$$

We can now prove Proposition 3.1.

Proof of Proposition 3.1. By using Lemma 3.2 we have

$$\int_{\mathcal{O}} y\frac{\partial^2 u}{\partial x^2}vd\mathfrak{m} = -\int_{\mathcal{O}} y\frac{\partial u}{\partial x}\left(\frac{\partial v}{\partial x} - \gamma sgn(x)v\right) d\mathfrak{m},$$

$$\int_{\mathcal{O}} y\frac{\partial^2 u}{\partial y^2}vd\mathfrak{m} = -\int_{\mathcal{O}} y\frac{\partial u}{\partial y}\frac{\partial v}{\partial y}d\mathfrak{m} + \int_{\mathcal{O}} (\mu y - \beta)\frac{\partial u}{\partial y}vd\mathfrak{m},$$

$$\int_{\mathcal{O}} y\frac{\partial^2 u}{\partial x\partial y}vd\mathfrak{m} = -\int_{\mathcal{O}} y\frac{\partial u}{\partial y}\left(\frac{\partial v}{\partial x} - \gamma sgn(x)v\right) d\mathfrak{m}$$

and

$$\int_{\mathcal{O}} y\frac{\partial^2 u}{\partial x\partial y}vd\mathfrak{m} = -\int_{\mathcal{O}} y\frac{\partial u}{\partial x}\frac{\partial v}{\partial y}d\mathfrak{m} + \int_{\mathcal{O}} (\mu y - \beta)\frac{\partial u}{\partial x}vd\mathfrak{m}.$$

Recalling that

$$\mathcal{L} = \frac{y}{2}\left(\frac{\partial^2}{\partial x^2} + 2\rho\sigma\frac{\partial^2}{\partial x\partial y} + \sigma^2\frac{\partial^2}{\partial y^2}\right) + \left(\frac{\rho\kappa\theta}{\sigma} - \frac{y}{2}\right)\frac{\partial}{\partial x} + \kappa(\theta - y)\frac{\partial}{\partial y}$$

and using the equality $\beta = 2\kappa\theta/\sigma^2$, we get

$$\begin{aligned}(\mathcal{L}u, v)_H &= -\int_{\mathcal{O}} \frac{y}{2}\left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + \sigma^2\frac{\partial u}{\partial y}\frac{\partial v}{\partial y} + \rho\sigma\frac{\partial u}{\partial x}\frac{\partial v}{\partial y} + \rho\sigma\frac{\partial u}{\partial y}\frac{\partial v}{\partial x}\right) d\mathfrak{m} + \int_{\mathcal{O}} \frac{1}{2}\frac{\partial u}{\partial x}\left(y\gamma sgn(x) + \rho\sigma(\mu y - \beta)\right) vd\mathfrak{m} \\ &\quad + \int_{\mathcal{O}} \frac{1}{2}\frac{\partial u}{\partial y}\left(\mu\sigma^2 y - \beta\sigma^2 + \rho\sigma y\gamma sgn(x)\right) vd\mathfrak{m} + \int_{\mathcal{O}} \left[\left(\frac{\rho\kappa\theta}{\sigma} - \frac{y}{2}\right)\frac{\partial u}{\partial x} + \kappa(\theta - y)\frac{\partial u}{\partial y}\right] vd\mathfrak{m} \\ &= -a(u, v).\end{aligned}$$

□

Remark 3.3. *By a closer look at the proof of Proposition 3.1 it is clear that the choice of $\bar{c}$ in (2.1) allows to avoid terms of the type $\int (u_x + u_y) v d\mathfrak{m}$ in the associated bilinear form a. This trick will be crucial in order to obtain suitable energy estimates.*

Recall the well-known inequality

$$bc = (\sqrt{\zeta} b)\left(\frac{c}{\sqrt{\zeta}}\right) \leq \frac{\zeta}{2} b^2 + \frac{1}{2\zeta} c^2, \qquad b, c \in \mathbb{R},\ \zeta > 0. \tag{3.2}$$

Hereafter we will often apply (3.2) in the proofs even if it is not explicitly recalled each time.

We have the following energy estimates.

Proposition 3.4. *For every $u, v \in V$, the bilinear form $a(\cdot, \cdot)$ satisfies*

$$|a(u, v)| \leq C_1 \|u\|_V \|v\|_V, \tag{3.3}$$

$$a(u, u) \geq C_2 \|u\|_V^2 - C_3 \|(1 + y)^{\frac{1}{2}} u\|_H^2, \tag{3.4}$$

where

$$C_1 = \delta_0 + K_1, \quad C_2 = \frac{\delta_1}{2}, \quad C_3 = \frac{\delta_1}{2} + \frac{K_1^2}{2\delta_1},$$

with

$$\delta_0 = \sup_{s_1^2 + t_1^2 > 0,\ s_2^2 + t_2^2 > 0} \frac{|s_1 s_2 + \rho\sigma s_1 t_2 + \rho\sigma s_2 t_1 + \sigma^2 t_1 t_2|}{2\sqrt{(s_1^2 + t_1^2)(s_2^2 + t_2^2)}}, \tag{3.5}$$

$$\delta_1 = \inf_{s^2 + t^2 > 0} \frac{s^2 + 2\rho\sigma st + \sigma^2 t^2}{2(s^2 + t^2)}, \tag{3.6}$$

and

$$K_1 = \sup_{x \in \mathbb{R}} \sqrt{j_{\gamma,\mu}^2(x) + k_{\gamma,\mu}^2(x)}. \tag{3.7}$$

It is easy to see that the constants δ_0, δ_1 and K_1 defined in (3.5) and (3.7) are positive and finite (recall that the functions $j_{\gamma,\mu} = j_{\gamma,\mu}(x)$ and $k_{\gamma,\mu} = \kappa_{\gamma,\mu}(x)$ defined in (2.4) are bounded).

These energy estimates were already proved in [5, Lemma 2.40] with a very similar statement. Here we repeat the proof for the sake of completeness, since we will refer to it later on.

Proof of Proposition 3.4. In order to prove (3.4), we note that

$$\frac{1}{2} \int_{\mathcal{O}} y \left(u_x v_x + \rho\sigma u_x v_y + \rho\sigma u_y v_x + \sigma^2 u_y v_y\right) d\mathfrak{m} \geq \delta_1 \int_{\mathcal{O}} y |\nabla u|^2 d\mathfrak{m}.$$

Therefore

$$\begin{aligned}
a(u, u) &\geq \delta_1 \int_{\mathcal{O}} y |\nabla u|^2 d\mathfrak{m} - K_1 \int_{\mathcal{O}} y |\nabla u| |u| d\mathfrak{m} \\
&\geq \delta_1 \int_{\mathcal{O}} y |\nabla u|^2 d\mathfrak{m} - \frac{K_1 \zeta}{2} \int_{\mathcal{O}} y |\nabla u|^2 d\mathfrak{m} - \frac{K_1}{2\zeta} \int_{\mathcal{O}} (1 + y) u^2 d\mathfrak{m} \\
&= \left(\delta_1 - \frac{K_1 \zeta}{2}\right) \int_{\mathcal{O}} \left(y |\nabla u|^2 + (1 + y) u^2\right) d\mathfrak{m} - \left(\delta_1 - \frac{K_1 \zeta}{2} + \frac{K_1}{2\zeta}\right) \int_{\mathcal{O}} (1 + y) u^2 d\mathfrak{m}.
\end{aligned}$$

The assertion then follows by choosing $\zeta = \delta_1 / K_1$. (3.3) can be proved in a similar way. □

3.2 Proof of Theorem 2.3

Among the standard assumptions required in [4] for the penalization procedure, there are the coercivity and the boundedness of the coefficients. In the Heston-type models these assumptions are no longer satisfied and this leads to some technical difficulties. In order to overcome them, we introduce some auxiliary operators.

From now on, we set

$$a(u,v) = \bar{a}(u,v) + \tilde{a}(u,v),$$

where

$$\begin{aligned}
\bar{a}(u,v) &= \int_{\mathcal{O}} \frac{y}{2}\left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + \rho\sigma\frac{\partial u}{\partial x}\frac{\partial v}{\partial y} + \rho\sigma\frac{\partial u}{\partial y}\frac{\partial v}{\partial x} + \sigma^2\frac{\partial u}{\partial y}\frac{\partial v}{\partial y}\right) d\mathfrak{m}, \\
\tilde{a}(u,v) &= \int_{\mathcal{O}} y\frac{\partial u}{\partial x} j_{\gamma,\mu} v d\mathfrak{m} + \int_{\mathcal{O}} y\frac{\partial u}{\partial y} k_{\gamma,\mu} v d\mathfrak{m}.
\end{aligned}$$

Note that $\bar{a}$ is symmetric. As in the proof of Proposition (3.4) we have, for every $u, v \in V$,

$$|\bar{a}(u,v)| \leq \delta_0 \int_{\mathcal{O}} y|\nabla u||\nabla v| d\mathfrak{m},$$

$$\bar{a}(u,u) \geq \delta_1 \int_{\mathcal{O}} y|\nabla u|^2 d\mathfrak{m},$$

and

$$|\tilde{a}(u,v)| \leq K_1 \int_{\mathcal{O}} y|\nabla u||v| d\mathfrak{m},$$

with δ_0, δ_1 and K_1 defined in Proposition 3.4. Moreover, for $\lambda \geq 0$ and $M > 0$ we consider the bilinear forms

$$\begin{aligned}
a_\lambda(u,v) &= a(u,v) + \lambda \int_{\mathcal{O}} (1+y) uv d\mathfrak{m}, \\
\bar{a}_\lambda(u,v) &= \bar{a}(u,v) + \lambda \int_{\mathcal{O}} (1+y) uv d\mathfrak{m}, \\
\tilde{a}^{(M)}(u,v) &= \int_{\mathcal{O}} (y \wedge M)\left(\frac{\partial u}{\partial x} j_{\gamma,\mu} + \frac{\partial u}{\partial y} k_{\gamma,\mu}\right) v d\mathfrak{m}
\end{aligned}$$

and

$$a_\lambda^{(M)}(u,v) = \bar{a}_\lambda(u,v) + \tilde{a}^{(M)}(u,v).$$

The operator a_λ was introduced in [5] to deal with the lack of coercivity of the bilinear form a, while the introduction of the truncated operator $a_\lambda^{(M)}$ with $M > 0$ will be useful in order to overcome the technical difficulty related to the unboundedness of the coefficients.

Lemma 3.5. *Let δ_0, δ_1, K_1 be defined as in (3.5), (3.6) and (3.7) respectively. For any fixed $\lambda \geq \frac{\delta_1}{2} + \frac{K_1^2}{2\delta_1}$ the bilinear forms a_λ and $a_\lambda^{(M)}$ are continuous and coercive. More precisely, we have*

$$|a_\lambda(u,v)| \leq C\|u\|_V \|v\|_V, \qquad u, v \in V, \tag{3.8}$$

$$a_\lambda(u,u) \geq \frac{\delta_1}{2}\|u\|_V^2, \qquad u \in V, \tag{3.9}$$

and

$$|a_\lambda^{(M)}(u,v)| \leq C\|u\|_V \|v\|_V, \qquad u, v \in V, \tag{3.10}$$

$$a_\lambda^{(M)}(u,u) \geq \frac{\delta_1}{2}\|u\|_V^2, \qquad u \in V. \tag{3.11}$$

where $C = \delta_0 + K_1 + \lambda$.

Proof. The proof for the bilinear form a_λ follows as in [5, Lemma 3.2]. We give the details for $a_\lambda^{(M)}$ to check that the constants do not depend on M. Note that, for every $u, v \in V$,

$$|\tilde{a}^{(M)}(u,v)| \le K_1 \int_{\mathcal{O}} y|\nabla u||v| d\mathfrak{m},$$

so that by straightforward computations we get

$$|a_\lambda^{(M)}(u,v)| \le (\delta_0 + \lambda + K_1)\|u\|_V \|v\|_V.$$

On the other hand, for every $\zeta > 0$,

$$\begin{aligned} a_\lambda^{(M)}(u,u) &\ge \delta_1 \int_{\mathcal{O}} y|\nabla u|^2 d\mathfrak{m} + \lambda \int_{\mathcal{O}} (1+y)u^2 d\mathfrak{m} - K_1 \int_{\mathcal{O}} y|\nabla u||u| d\mathfrak{m} \\ &\ge \left(\delta_1 - \frac{K_1\zeta}{2}\right) \int_{\mathcal{O}} y|\nabla u|^2 d\mathfrak{m} + \left(\lambda - \frac{K_1}{2\zeta}\right) \int_{\mathcal{O}} (1+y)u^2 d\mathfrak{m}. \end{aligned}$$

By choosing $\zeta = \delta_1/K_1$, we get

$$a_\lambda^{(M)}(u,u) \ge \frac{\delta_1}{2} \int_{\mathcal{O}} y|\nabla u|^2 dm + \left(\lambda - \frac{K_1^2}{2\delta_1}\right) \int_{\mathcal{O}} (1+y)u^2 d\mathfrak{m} \ge \frac{\delta_1}{2}\|u\|_V^2,$$

for every $\lambda \ge \frac{\delta_1}{2} + \frac{K_1^2}{2\delta_1}$. □

From now on in the rest of this paper we assume $\lambda \ge \frac{\delta_1}{2} + \frac{K_1^2}{2\delta_1}$ as in Lemma 3.5. Moreover, we will denote by $\|b\| = \sup_{u,v\in V, u,v\neq 0} \frac{|b(u,v)|}{\|u\|_V \|v\|_V}$ the norm of a bilinear form $b : V \times V \to \mathbb{R}$.

Remark 3.6. *We stress that Lemma 3.5 gives us*

$$\sup_{M>0} \|a_\lambda^{(M)}\| \le C, \tag{3.12}$$

where $C = \delta_0 + K_1 + \lambda$. This will be crucial in the penalization technique we are going to describe in Section 3.2.1. Roughly speaking, in order to prove the existence of a solution of the penalized coercive problem we will introduce in Theorem 3.8, we proceed as follows. First, we replace the bilinear form a_λ with the operator $a_\lambda^{(M)}$, which has bounded coefficients, and we solve the associated penalized truncated coercive problem (see Proposition 3.9). Then, thanks to (3.12)*, we can deduce estimates on the solution which are uniform in M (see Lemma 3.10) and which will allow us to pass to the limit as M goes to infinity and to find a solution of the original penalized coercive problem.*

Finally, we define

$$\mathcal{L}^\lambda := \mathcal{L} - \lambda(1+y)$$

the differential operator associated with the bilinear form a_λ, that is

$$(\mathcal{L}^\lambda u, v)_H = -a_\lambda(u,v), \qquad u \in H^2(\mathcal{O}, \mathfrak{m}),\ v \in V.$$

3.2.1 Penalized problem

For any fixed $\varepsilon > 0$ we define the penalizing operator

$$\zeta_\varepsilon(t,u) = -\frac{1}{\varepsilon}(\psi(t) - u)_+ = \frac{1}{\varepsilon}\zeta(t,u), \qquad t \in [0,T], u \in V. \tag{3.13}$$

Since for every fixed $t \in [0,T]$ the function $x \mapsto -(\psi(t) - x)_+$ is nondecreasing, we have the following well known monotonicity result (see [4]).

Lemma 3.7. *For any fixed $t \in [0,T]$ the penalizing operator (3.13) is monotone, in the sense that*

$$(\zeta_\varepsilon(t,u) - \zeta_\varepsilon(t,v), u-v)_H \geq 0, \qquad u,v \in V.$$

We now introduce the intermediate penalized coercive problem with a source term g. We consider the following assumption:

Assumption $\mathcal{H}^0$. We say that a function g satisfies Assumption $\mathcal{H}^0$ if $\sqrt{1+y}g \in L^2([0,T];H)$.

Theorem 3.8. *Assume that ψ satisfies Assumption $\mathcal{H}^1$ and g satisfies Assumption $\mathcal{H}^0$. Then, for every fixed $\varepsilon > 0$, there exists a unique function $u_{\varepsilon,\lambda}$ such that $u_{\varepsilon,\lambda} \in L^2([0,T];V)$, $\frac{\partial u_{\varepsilon,\lambda}}{\partial t} \in L^2([0,T];H)$ and, for all $v \in L^2([0,T];V)$,*

$$\begin{cases} -\left(\frac{\partial u_{\varepsilon,\lambda}}{\partial t}(t), v(t)\right)_H + a_\lambda(u_{\varepsilon,\lambda}(t), v(t)) + (\zeta_\varepsilon(t, u_{\varepsilon,\lambda}(t)), v(t))_H = (g(t), v(t))_H, & \text{a.e. in } [0,T], \\ u_{\varepsilon,\lambda}(T) = \psi(T). \end{cases} \tag{3.14}$$

Moreover, the following estimates hold:

$$\|u_{\varepsilon,\lambda}\|_{L^\infty([0,T],V)} \leq K, \tag{3.15}$$

$$\left\| \frac{\partial u_{\varepsilon,\lambda}}{\partial t} \right\|_{L^2([0,T];H)} \leq K, \tag{3.16}$$

$$\frac{1}{\sqrt{\varepsilon}} \left\| (\psi - u_{\varepsilon,\lambda})^+ \right\|_{L^\infty([0,T],H)} \leq K, \tag{3.17}$$

where $K = C\left(\|\Psi\|_{L^2([0,T];V)} + \|\sqrt{1+y}g\|_{L^2([0,T];H)} + \|\sqrt{1+y}\psi\|_{L^2([0,T];V)} + \|\psi(T)\|_V^2\right)$, with $C > 0$ independent of ε, and Ψ is given in Assumption $\mathcal{H}^1$.

The proof of uniqueness of the solution of the penalized coercive problem follows a standard monotonicity argument as in [4], so we omit the proof.

The proof of existence in Theorem 3.8 is quite long and technical, so we split it into two propositions. We first consider the truncated penalized problem, which requires less stringent conditions on ψ and g.

Proposition 3.9. *Let $\psi \in \mathcal{C}([0,T];H) \cap L^2([0,T];V)$ and $g \in L^2([0,T];H)$. Moreover, assume that $\psi(T) \in H^2(\mathcal{O},\mathfrak{m})$, $(1+y)\psi(T) \in H$, $\frac{\partial \psi}{\partial t} \in L^2([0,T];V)$ and $\frac{\partial g}{\partial t} \in L^2([0,T];H)$. Then, there exists a unique function $u_{\varepsilon,\lambda,M}$ such that $u_{\varepsilon,\lambda,M} \in L^2([0,T];V)$, $\frac{\partial u_{\varepsilon,\lambda,M}}{\partial t} \in L^2([0,T];V)$ and for all $v \in L^2([0,T];V)$*

$$\begin{cases} -\left(\frac{\partial u_{\varepsilon,\lambda,M}}{\partial t}(t), v(t)\right)_H + a_\lambda^{(M)}(u_{\varepsilon,\lambda,M}(t), v(t)) + (\zeta_\varepsilon(t, u_{\varepsilon,\lambda,M}(t)), v(t))_H = (g(t), v(t))_H, & \text{a.e. in } [0,T), \\ u_{\varepsilon,\lambda,M}(T) = \psi(T). \end{cases} \tag{3.18}$$

Proof. 1. **Finite dimensional problem** We use the classical Galerkin method of approximation, which consists in introducing a nondecreasing sequence $(V_j)_j$ of subspaces of V such that $dim V_j < \infty$ and, for every $v \in V$, there exists a sequence $(v_j)_{j\in\mathbb{N}}$ such that $v_j \in V_j$ for any $j \in \mathbb{N}$ and $\|v - v_j\|_V \to 0$ as $j \to \infty$. Moreover, we assume that $\psi(T) \in V_j$, for all $j \in \mathbb{N}$. Let P_j be the projection of V onto V_j and $\psi_j(t) = P_j\psi(t)$. We have $\psi_j(t) \to \psi(t)$ strongly in V and $\psi_j(T) = \psi(T)$ for any $j \in \mathbb{N}$. The finite dimensional problem is, therefore, to find $u_j : [0,T] \to V_j$ such that

$$\begin{cases} -\left(\frac{\partial u_j}{\partial t}(t), v\right)_H + a_\lambda^{(M)}(u_j(t), v) - \frac{1}{\varepsilon}((\psi_j(t) - u_j(t))_+, v)_H = (g(t), v)_H, & v \in V_j, \\ u_j(T) = \psi(T). \end{cases} \tag{3.19}$$

This problem can be interpreted as an ordinary differential equation in V_j ($\dim V_j < \infty$), that is

$$\begin{cases} -\frac{\partial u_j}{\partial t}(t) + A_{\lambda,j}^{(M)} u_j(t) - \frac{1}{\varepsilon} Q_j((\psi_j(t) - u_j(t))_+) = Q_j g(t), \\ u_j(T) = \psi(T), \end{cases}$$

where $A_{\lambda,j}^{(M)} : V_j \to V_j$ is a finite dimensional linear operator and Q_j is the projection of H onto V_j. It is not difficult to prove that the function $(t,u) \to Q_j((\psi_j(t) - u(t))_+)$ is continuous with values in V_j and Lipschitz continuous with respect to the u variable and that the term $Q_j g$ belongs to $L^2([0,T];V_j)$ (we refer to [17] for the details). Therefore, by the Cauchy-Lipschitz Theorem, we can deduce the existence and the uniqueness of a solution u_j of (3.19), continuous from $[0,T]$ into V_j, a.e. differentiable and with integrable derivative.

2. **Estimates on the finite dimensional problem** First, we take $v = u_j(t) - \psi_j(t)$ in (3.19). We get

$$-\left(\frac{\partial u_j}{\partial t}(t), u_j(t) - \psi_j(t)\right)_H + a_\lambda^{(M)}(u_j(t), u_j(t) - \psi_j(t)) - \frac{1}{\varepsilon}((\psi_j(t) - u_j(t))_+, u_j(t) - \psi_j(t))_H$$
$$= (g(t), u_j(t) - \psi_j(t))_H,$$

which can be rewritten as

$$-\frac{1}{2}\frac{d}{dt}\|u_j(t) - \psi_j(t)\|_H^2 - \left(\frac{\partial \psi_j}{\partial t}(t), u_j(t) - \psi_j(t)\right)_H + a_\lambda^{(M)}(u_j(t) - \psi_j(t), u_j(t) - \psi_j(t))_H$$
$$+ \frac{1}{\varepsilon}((\psi_j(t) - u_j(t))_+, \psi_j(t) - u_j(t))_H + a_\lambda^{(M)}(\psi_j(t), u_j(t) - \psi_j(t)) = (g(t), u_j(t) - \psi_j(t))_H.$$

We integrate between t and T and we use coercivity and $u_j(T) = \psi_j(T)$ to obtain

$$\frac{1}{2}\|u_j(t) - \psi_j(t)\|_H^2 + \frac{\delta_1}{2}\int_t^T \|u_j(s) - \psi_j(s)\|_V^2 ds + \frac{1}{\varepsilon}\int_t^T \|(\psi_j(s) - u_j(s))_+\|_H^2 ds$$
$$\leq \frac{1}{2\zeta}\int_t^T \left\|\frac{\partial \psi_j(s)}{\partial t}\right\|_H^2 ds + \frac{\zeta}{2}\int_t^T \|u_j(s) - \psi_j(s)\|_H^2 ds + \frac{1}{2\zeta}\int_t^T \|g(s)\|_H^2 ds + \frac{\zeta}{2}\int_t^T \|u_j(s) - \psi_j(s)\|_H^2 ds$$
$$+ \frac{\|a_\lambda^{(M)}\|\zeta}{2}\int_t^T \|u_j(s) - \psi_j(s)\|_V^2 ds + \frac{\|a_\lambda^{(M)}\|}{2\zeta}\int_t^T \|\psi_j(s)\|_V^2 ds,$$

for any $\zeta > 0$. Recall that $\psi_j = P_j\psi$, and so $\|\psi_j(t)\|_V^2 \leq \|\psi(t)\|_V^2$. In the same way $\|\frac{\partial \psi_j(t)}{\partial t}\|_H^2 \leq \|\frac{\partial \psi_j(t)}{\partial t}\|_V^2 \leq \|\frac{\partial \psi(t)}{\partial t}\|_V^2$. Choosing $\zeta = \frac{\delta_1}{4+2\|a_\lambda^{(M)}\|}$ after simple calculations we deduce that there exists $C > 0$ independent of M, ε and j such that

$$\begin{aligned} &\tfrac{1}{4}\|u_j(t)\|_H^2 + \tfrac{\delta_1}{8}\int_t^T \|u_j(s)\|_V^2 ds + \tfrac{1}{\varepsilon}\int_t^T \|(\psi_j(s) - u_j(s))_+\|_H^2 ds \\ &\leq C\left(\left\|\tfrac{\partial \psi}{\partial t}\right\|_{L^2([t,T];V)}^2 + \|g\|_{L^2([t,T];H)}^2 + \|\psi\|_{L^2([t,T];V)}^2 + \|\psi(T)\|_H^2\right). \end{aligned} \tag{3.20}$$

We now go back to (3.19) and we take $v = \frac{\partial u_j}{\partial t}(t)$ so we get

$$-\left\|\frac{\partial u_j}{\partial t}(t)\right\|_H^2 + \bar{a}_\lambda\left(u_j(t), \frac{\partial u_j}{\partial t}(t)\right) + \tilde{a}^{(M)}\left(u_j(t), \frac{\partial u_j}{\partial t}(t)\right) - \frac{1}{\varepsilon}\left((\psi_j(t) - u_j(t))_+, \frac{\partial u_j}{\partial t}(t)\right)_H$$
$$= \left(g(t), \frac{\partial u_j}{\partial t}(t)\right)_H.$$

Note that

$$-\frac{1}{\varepsilon}\left((\psi_j(t) - u_j(t))_+, \frac{\partial u_j}{\partial t}(t)\right)_H = \frac{1}{\varepsilon}\left((\psi_j - u_j)_+, \frac{\partial(\psi_j - u_j)}{\partial t}(t)\right)_H - \frac{1}{\varepsilon}\left((\psi_j(t) - u_j(t))_+, \frac{\partial \psi_j}{\partial t}(t)\right)_H$$
$$= \frac{1}{2\varepsilon}\frac{d}{dt}\|(\psi_j - u_j)_+(t)\|_H^2 - \frac{1}{\varepsilon}\left((\psi_j(t) - u_j(t))_+, \frac{\partial \psi_j}{\partial t}(t)\right)_H.$$

Therefore, using the symmetry of $\bar{a}_\lambda$, we have

$$-\left\|\frac{\partial u_j}{\partial t}(t)\right\|_H^2+\frac{1}{2}\frac{d}{dt}\bar{a}_\lambda(u_j(t),u_j(t))+\tilde{a}^{(M)}\left(u_j(t),\frac{\partial u_j}{\partial t}(t)\right)+\frac{1}{2\varepsilon}\frac{\partial}{\partial t}\|(\psi_j(t)-u_j(t))_+\|_H^2$$
$$-\frac{1}{\varepsilon}\left((\psi_j(t)-u_j(t))_+,\frac{\partial\psi_j}{\partial t}(t)\right)_H=\left(g(t),\frac{\partial u_j}{\partial t}(t)\right)_H.$$

Integrating between t and T, we obtain

$$\begin{aligned}&\int_t^T\left\|\frac{\partial u_j}{\partial t}(s)\right\|_H^2ds+\frac{1}{2}\bar{a}_\lambda(u_j(t),u_j(t))+\frac{1}{2\varepsilon}\|(\psi_j(t)-u_j(t))_+\|_H^2\\&=\int_t^T\tilde{a}^{(M)}\left(u_j(s),\frac{\partial u_j}{\partial s}(s)\right)ds+\frac{1}{2}\bar{a}_\lambda(\psi_j(T),\psi_j(T))-\int_t^T\frac{1}{\varepsilon}\left((\psi_j(s)-u_j(s)_+,\frac{\partial\psi_j}{\partial s}(s)\right)_Hds\\&-\int_t^T\left(g(s),\frac{\partial u_j}{\partial s}(s)\right)_Hds.\end{aligned}$$

Recall that $\bar{a}_\lambda(u_j(t),u_j(t))\geq\frac{\delta_1}{2}\|u_j(t)\|_V^2$, $\bar{a}_\lambda(\psi_j(T),\psi_j(T))=\bar{a}_\lambda(\psi(T),\psi(T))\leq\|\bar{a}_\lambda\|\|\psi(T)\|_V^2$ and $|\tilde{a}^{(M)}(u,v)|\leq K_1\int_{\mathcal{O}}y\wedge M|\nabla u||v|d\mathfrak{m}$, so that, for every $\zeta>0$,

$$\begin{aligned}&\int_t^T\left\|\frac{\partial u_j}{\partial s}(s)\right\|_H^2ds+\frac{\delta_1}{4}\|u_j(t)\|_V^2+\frac{1}{2\varepsilon}\|(\psi_j(t)-u_j(t))_+\|_H^2\\&\leq K_1\int_t^Tds\int_{\mathcal{O}}y\wedge M|\nabla u_j(s,.)|\left|\frac{\partial u_j}{\partial t}(s,.)\right|d\mathfrak{m}+\frac{\|\bar{a}_\lambda\|}{2}\|\psi(T)\|_V^2+\frac{1}{\varepsilon}\int_t^T\|(\psi_j(s)-u_j(s))_+\|_H\left\|\frac{\partial\psi_j}{\partial s}(s)\right\|_Hds\\&\quad+\int_t^T\|g(s)\|_H\left\|\frac{\partial u_j}{\partial s}(s)\right\|_Hds\\&\leq\frac{K_1}{2\zeta}\int_t^T\|u_j(s)\|_V^2ds+\frac{K_1M}{2}\zeta\int_t^T\left\|\frac{\partial u_j}{\partial s}(s)\right\|_H^2ds+\frac{\|\bar{a}_\lambda\|}{2}\|\psi(T)\|_V^2\\&\quad+\frac{\zeta}{2\varepsilon}\int_t^T\|(\psi_j(s)-u_j(s))_+\|_H^2ds+\frac{1}{2\zeta\varepsilon}\int_t^T\left\|\frac{\partial\psi_j}{\partial t}(s)\right\|_H^2ds+\frac{1}{2\zeta}\int_t^T\|g(s)\|_H^2ds+\frac{\zeta}{2}\int_t^T\left\|\frac{\partial u_j}{\partial s}(s)\right\|_H^2ds.\end{aligned}$$

From (3.20), we already know that

$$\int_t^T\|u_j(s)\|_V^2ds+\frac{1}{\varepsilon}\int_t^T\|(\psi_j(s)-u_j(s))_+\|_H^2ds\leq C\left(\left\|\frac{\partial\psi}{\partial t}\right\|_{L^2([t,T];V)}^2+\|g\|_{L^2([t,T];H)}^2+\|\psi\|_{L^2([t,T];V)}^2+\|\psi(T)\|_H^2\right),$$

then we can finally deduce

$$\begin{aligned}&\int_t^T\left\|\frac{\partial u_j}{\partial t}(s)\right\|_H^2ds+\|u_j(t)\|_V^2+\frac{1}{2\varepsilon}\|(\psi_j(t)-u_j(t))_+\|_H^2\\&\leq C_{\varepsilon,M}\left(\left\|\frac{\partial\psi}{\partial t}\right\|_{L^2([t,T];V)}^2+\|g\|_{L^2([t,T];H)}^2+\|\psi\|_{L^2([t,T];V)}^2+\|\psi(T)\|_V^2\right),\end{aligned}\tag{3.21}$$

where $C_{\varepsilon,M}$ is a constant which depends on ε and M but not on j.

We will also need a further estimation. If we denote $\bar{u}_j=\frac{\partial u_j}{\partial t}$ and we differentiate the equation (3.19) with respect to t for a fixed v independent of t, we obtain that $\bar{u}_j$ satisfies

$$-\left(\frac{\partial\bar{u}_j}{\partial t}(t),v\right)_H+a_\lambda^{(M)}(\bar{u}_j(t),v)-\frac{1}{\varepsilon}\left(\left(\frac{\partial\psi_j}{\partial t}(t)-\bar{u}_j(t)\right)\mathbb{1}_{\{\psi_j(t)\geq u_j(t)\}},v\right)_H=\left(\frac{\partial g}{\partial t}(t),v\right)_H,\quad v\in V_j.\tag{3.22}$$

As regards the initial condition, from (3.19) computed in $t=T$, for every $v\in V_j$ we have

$$\begin{aligned}\left(\frac{\partial u_j(T)}{\partial t},v\right)_H &= a_\lambda^{(M)}(\psi(T),v)-(g(T),v)_H.\\ &= -\left(\mathcal{L}\psi(T),v\right)_H+\lambda\left((1+y)\psi(T),v\right)_H+\left((y\wedge M-y)\left(j_{\gamma,\mu}\frac{\psi(T)}{\partial x}+k_{\gamma,\mu}\frac{\psi(T)}{\partial y}\right),v\right)_H+(g(T),v)_H.\end{aligned}$$

Choosing $v=\frac{\partial u_j(T)}{\partial t}$, we deduce that

$$\begin{aligned}\left\|\frac{\partial u_j(T)}{\partial t}\right\|_H &\le C\left(\|\mathcal{L}\psi(T)\|_H+\|(1+y)\psi(T)\|_H+\|(y-M)_+\nabla\psi(T)\|_H+\|g(T)\|_H\right)\\ &\le C\left(\|\psi(T)\|_{H^2(\mathcal{O},\mathfrak{m})}+\|(1+y)\psi(T)\|_H+\|g(T)\|_H\right),\end{aligned}$$

where we have used that$\|\mathcal{L}\psi(T)\|_H+\|(y-M)_+\nabla\psi(T)\|_H\le C\|\psi(T)\|_{H^2(\mathcal{O},\mathfrak{m})}$ for a certain constant $C>0$.
We can take $v=\bar u_j(t)$ in (3.22) and we obtain

$$-\left(\frac{\partial\bar u_j}{\partial t}(t),\bar u_j(t)\right)_H+a_\lambda^{(M)}(\bar u_j(t),\bar u_j(t))-\frac{1}{\varepsilon}\left(\left(\frac{\partial\psi_j}{\partial t}(t)-\bar u_j(t)\right)\mathbb{1}_{\{\psi_j(t)\ge u_j(t)\}},\bar u_j(t)\right)_H=\left(\frac{\partial g}{\partial t}(t),\bar u_j(t)\right)_H,$$

so that

$$\begin{aligned}-\frac12\frac{d}{dt}\|\bar u_j(t)\|_H^2+\frac{\delta_1}{2}\|\bar u_j(t)\|_V^2 &\le \frac{1}{\varepsilon}\left(\left(\frac{\partial\psi_j}{\partial t}(t)-\bar u_j(t)\right)\mathbb{1}_{\{\psi_j(t)\ge u_j\}},\bar u_j(t)\right)_H+\left(\frac{\partial g}{\partial t}(t),\bar u_j(t)\right)_H\\ &\le \frac{1}{\varepsilon}\left(\frac{\partial\psi_j}{\partial t}(t)\mathbb{1}_{\{\psi_j(t)\ge u_j\}},\bar u_j(t)\right)_H+\left(\frac{\partial g}{\partial t}(t),\bar u_j(t)\right)_H.\end{aligned}$$

Integrating between t and T, with the usual calculations, we obtain, in particular, that

$$\begin{aligned}&\|\bar u_j(t)\|_H^2+\frac{\delta_1}{2}\int_t^T\|\bar u_j(s)\|_V^2ds\\ &\le C_\varepsilon\left(\|\psi(T)\|_{H^2(\mathcal{O},\mathfrak{m})}^2+\|(1+y)\psi(T)\|_H^2+\|g(T)\|_H^2+\left\|\frac{\partial\psi}{\partial t}\right\|_{L^2([t,T];H)}^2+\left\|\frac{\partial g}{\partial t}\right\|_{L^2([t,T];H)}^2\right),\end{aligned}\tag{3.23}$$

where C_ε is a constant which depends on ε, but not on j.

3. **Passage to the limit**

Let ε and M be fixed. By passing to a subsequence, from (3.21) we can assume that $\frac{\partial u_j}{\partial t}$ weakly converges to a function $u'_{\varepsilon,\lambda,M}$ in $L^2([0,T];H)$. We deduce that, for any fixed $t\in[0,T]$, $u_j(t)$ weakly converges in H to

$$u_{\varepsilon,\lambda,M}(t)=\psi(T)-\int_t^T u'_{\varepsilon,\lambda,M}(s)ds.$$

Indeed, $u_j(t)$ is bounded in V, so the convergence is weakly in V. Passing to the limit in (3.23) we deduce that $\frac{\partial u_{\varepsilon,\lambda,M}}{\partial t}\in L^2([0,T];V)$. Moreover, from (3.21), we have that $(\psi_j-u_j(t))^+$ weakly converges in H to a certain function $\chi(t)\in H$. Now, for any $v\in V$ we know that there exists a sequence $(v_j)_{j\in\mathbb{N}}$ such that $v_j\in V_j$ for all $j\in\mathbb{N}$ and $\|v-v_j\|_V\to0$. We have

$$-\left(\frac{\partial u_j}{\partial t}(t),v_j\right)_H+a_\lambda^{(M)}(u_j(t),v_j)_H-\frac{1}{\varepsilon}((\psi_j(t)-u_j(t))_+,v_j)_H=(g(t),v_j)_H$$

so, passing to the limit as $j\to\infty$,

$$-\left(\frac{\partial u_{\varepsilon,\lambda,M}}{\partial t}(t),v\right)_H+a_\lambda(u_{\varepsilon,\lambda,M}(t),v)_H-\frac{1}{\varepsilon}(\chi(t),v)_H=(g(t),v)_H.$$

We only have to note that $\chi(t)=(\psi(t)-u_{\varepsilon,\lambda,M}(t))_+$. In fact, $\psi_j(t)\to\psi(t)$ in V and, up to a subsequence, $\mathbb{1}_{\mathcal{U}}u_j(t)\to\mathbb{1}_{\mathcal{U}}u_{\varepsilon,\lambda,M}(t)$ in $L^2(\mathcal{U},\mathfrak{m})$ for every open $\mathcal{U}$ relatively compact in $\mathcal{O}$. Therefore, there exists a subsequence which converges a.e. and this allows to conclude the proof.

□

We now want to get rid of the truncated operator, that is to pass to the limit for $M \to \infty$. In order to do this we need some estimates on the function $u_{\varepsilon,\lambda,M}$ which are uniform in M.

Lemma 3.10. *Assume that, in addition to the assumptions of Proposition 3.9,* $\sqrt{1+y}\psi \in L^2([0,T];V)$, $\left|\frac{\partial \psi}{\partial t}\right| \leq \Psi$ *with* $\Psi \in L^2([0,T];V)$ *and* g *satisfies Assumption* $\mathcal{H}^0$. *Let* $u_{\varepsilon,\lambda,M}$ *be the solution of* (3.18). *Then,*

$$\begin{aligned} &\int_t^T \left\| \tfrac{\partial u_{\varepsilon,\lambda,M}}{\partial s}(s) \right\|_H^2 ds + \|u_{\varepsilon,\lambda,M}(t)\|_V^2 + \tfrac{1}{\varepsilon}\|(\psi(t) - u_{\varepsilon,\lambda,M}(t))_+\|_H^2 \\ &\leq C\left(\|\Psi\|_{L^2([0,T];V)} + \|\sqrt{1+y}g\|_{L^2([0,T];H)} + \|\sqrt{1+y}\psi\|^2_{L^2([0,T];V)} + \|\psi(T)\|_V^2 \right), \end{aligned} \tag{3.24}$$

where C *is a positive constant independent of* M *and* ε.

Proof. To simplify the notation we denote $u_{\varepsilon,\lambda,M}$ by u and $u_{\varepsilon,\lambda,M} - \psi = u - \psi$ by w. For $n \geq 0$, define $\varphi_n(x,y) = 1 + y \wedge n$. Since φ_n and its derivatives are bounded, if $v \in V$, we have $v\varphi_n \in V$. Choosing $v = (u-\psi)\varphi_n = w\varphi_n$ in (3.18), with simple passages we get

$$\begin{aligned} &-\left(\frac{\partial w}{\partial t}(t), w(t)\varphi_n\right)_H + a_\lambda^{(M)}(w(t), w(t)\varphi_n) + (\zeta_\varepsilon(t,u(t)), w(t)\varphi_n)_H \\ &\quad = \left(\frac{\partial \psi}{\partial t}(t) + g(t), w(t)\varphi_n\right)_H - a_\lambda^{(M)}(\psi(t), w(t)\varphi_n). \end{aligned}$$

With the notation $\varphi_n' = \frac{\partial \varphi_n}{\partial y} = \mathbb{1}_{\{y \leq n\}}$, we have

$$\begin{aligned} a_\lambda^{(M)}(w(t), w(t)\varphi_n) &= \int_{\mathcal{O}} \frac{y}{2}\left[\left(\frac{\partial w}{\partial x}(t)\right)^2 + 2\rho\sigma\frac{\partial w}{\partial x}(t)\frac{\partial w}{\partial y}(t) + \sigma^2\left(\frac{\partial w}{\partial y}(t)\right)^2\right]\varphi_n d\mathfrak{m} + \lambda\int_{\mathcal{O}}(1+y)w^2(t)\varphi_n d\mathfrak{m} \\ &+ \int_{\mathcal{O}} \frac{y}{2}\left(\rho\sigma\frac{\partial w}{\partial x}(t) + \sigma^2\frac{\partial w}{\partial y}(t)\right) w(t)\varphi_n' d\mathfrak{m} + \int_{\mathcal{O}} y \wedge M\left(\frac{\partial w}{\partial x}(t) j_{\gamma,\mu} + \frac{\partial w}{\partial y}(t) k_{\gamma,\mu}\right) w(t)\varphi_n d\mathfrak{m} \\ &\geq \delta_1 \int_{\mathcal{O}} y\left|\nabla w(t)\right|^2 \varphi_n d\mathfrak{m} + \lambda\int_{\mathcal{O}}(1+y)w^2(t)\varphi_n d\mathfrak{m} - K_1\int_{\mathcal{O}} y\left|\nabla w(t)\right||w(t)|\varphi_n d\mathfrak{m} \\ &- K_2\int_{\mathcal{O}} y\left|\nabla w(t)\right||w(t)|\mathbb{1}_{\{y\leq n\}} d\mathfrak{m}, \end{aligned}$$

where $K_2 = \frac{\sqrt{\rho^2\sigma^2+\sigma^4}}{2}$. Note that, if $n = 0$, the last term vanishes, and that, for all $n > 0$,

$$\int_{\mathcal{O}} y\left|\nabla w(t)\right||w(t)|\mathbb{1}_{\{y\leq n\}} d\mathfrak{m} \leq \|w(t)\|_V^2.$$

Therefore, for all $\zeta > 0$,

$$\begin{aligned} a_\lambda^{(M)}(w(t), w(t)\varphi_n) &\geq \delta_1 \int_{\mathcal{O}} y\left|\nabla w(t)\right|^2 \varphi_n d\mathfrak{m} + \lambda\int_{\mathcal{O}}(1+y)w^2(t)\varphi_n d\mathfrak{m} \\ &- K_1\int_{\mathcal{O}} y\left(\frac{\zeta}{2}\left|\nabla w(t)\right|^2 + \frac{1}{2\zeta}|w(t)|^2\right)\varphi_n d\mathfrak{m} - K_2\|w(t)\|_V^2 \\ &\geq \left(\delta_1 - \frac{K_1\zeta}{2}\right)\int_{\mathcal{O}} y\left|\nabla w(t)\right|^2\varphi_n d\mathfrak{m} + \left(\lambda - \frac{K_1}{2\zeta}\right)\int_{\mathcal{O}}(1+y)w^2(t)\varphi_n d\mathfrak{m} - K_2\|w(t)\|_V^2 \\ &\geq \frac{\delta_1}{2}\int_{\mathcal{O}}\left(y\left|\nabla w(t)\right|^2 + (1+y)w^2(t)\right)\varphi_n d\mathfrak{m} - K_2\|w(t)\|_V^2, \end{aligned}$$

where, for the last inequality, we have chosen $\zeta = \delta_1/K_1$ and used the inequality $\lambda \geq \frac{\delta_1}{2} + \frac{K_1^2}{2\delta_1}$. Again, in the case $n = 0$ the last term on the righthand side can be omitted.

Hence, we have, with the notation $\|v\|^2_{V,n} = \int_{\mathcal{O}} \left(y\,|\nabla v|^2 + (1+y)v^2\right)\varphi_n d\mathfrak{m}$,

$$\frac{1}{2}\frac{d}{dt}\int_{\mathcal{O}} w^2(t)\varphi_n d\mathfrak{m} + \frac{\delta_1}{2}\|w(t)\|^2_{V,n} + \frac{1}{\varepsilon}\int_{\mathcal{O}}(-w(t))^2_+\varphi_n d\mathfrak{m} \leq \left(g(t) + \frac{\partial\psi}{\partial t}(t), w(t)\varphi_n\right)_H - a^{(M)}_\lambda(\psi(t), w(t)\varphi_n) + K_2\|w(t)\|^2_V.$$

In the case $n = 0$, the inequality reduces to

$$-\frac{1}{2}\frac{d}{dt}\int_{\mathcal{O}} w^2(t)d\mathfrak{m} + \frac{\delta_1}{2}\|w(t)\|^2_V + \frac{1}{\varepsilon}\int_{\mathcal{O}}(\psi - u)^2_+ d\mathfrak{m} \leq \left(g(t) + \frac{\partial\psi}{\partial t}(t), w(t)\right)_H - a^{(M)}_\lambda(\psi(t), w(t)).$$

Now, integrate from t to T and use $u(T) = \psi(T)$ to derive

$$\begin{aligned}&\frac{1}{2}\int_{\mathcal{O}} w(t)^2\varphi_n d\mathfrak{m} + \frac{\delta_1}{2}\int_t^T ds\|w(s)\|^2_{V,n} + \frac{1}{\varepsilon}\int_t^T ds\int_{\mathcal{O}}(-w(s))^2_+\varphi_n d\mathfrak{m}\\ &\leq \int_t^T \left(g(s) + \frac{\partial\psi}{\partial t}(s), w(s)\varphi_n\right)_H ds + \left|\int_t^T a^{(M)}_\lambda(\psi(s), w(s)\varphi_n)ds\right| + K_2\int_t^T \|w(s)\|^2_V ds,\end{aligned} \tag{3.25}$$

and, in the case $n = 0$,

$$\begin{aligned}&\frac{1}{2}\|w(t)\|^2_H + \frac{\delta_1}{2}\int_t^T \|w(s)\|^2_V ds + \frac{1}{\varepsilon}\int_t^T ds\int_{\mathcal{O}}(-w(s))^2_+ d\mathfrak{m}\\ &\leq \int_t^T \left(g(s) + \frac{\partial\psi}{\partial t}(s), w(s)\right)_H ds + \int_t^T \left|a^{(M)}_\lambda(\psi(s), w(s))\right| ds.\end{aligned} \tag{3.26}$$

We have, for all $\zeta_1 > 0$,

$$\begin{aligned}\int_t^T \left(g(s) + \frac{\partial\psi}{\partial t}(s), w(s)\varphi_n\right)_H ds &\leq \frac{\zeta_1}{2}\int_t^T ds\int_{\mathcal{O}}|w(s)|^2\varphi_n d\mathfrak{m} + \frac{1}{2\zeta_1}\int_t^T ds\int_{\mathcal{O}}\left|g(s) + \frac{\partial\psi}{\partial t}(s)\right|^2\varphi_n d\mathfrak{m}\\ &\leq \frac{\zeta_1}{2}\int_t^T ds\int_{\mathcal{O}}|w(s)|^2\varphi_n d\mathfrak{m} + \frac{1}{\zeta_1}\|\sqrt{1+y}g\|^2_{L^2([t,T];H)} + \frac{1}{\zeta_1}\left\|\sqrt{1+y}\frac{\partial\psi}{\partial t}\right\|^2_{L^2([t,T];H)}.\end{aligned}$$

Moreover, it is easy to check that, for all v_1, $v_2 \in V$,

$$|a^{(M)}_\lambda(v_1, v_2\varphi_n)| \leq K_3\|v_1\|_{V,n}\|v_2\|_{V,n}, \qquad \text{with } K_3 = \delta_0 + K_1 + K_2 + \lambda,$$

so that, for any $\zeta_2 > 0$,

$$\int_t^T |a^{(M)}_\lambda(\psi(s), w(s)\varphi_n)|ds \leq K_3\int_t^T ds\|\psi(s)\|_{V,n}\|w(s)\|_{V,n} \leq \frac{K_3\zeta_2}{2}\int_t^T ds\|w(s)\|^2_{V,n} + \frac{K_3}{2\zeta_2}\int_t^T ds\|\psi(s)\|^2_{V,n}.$$

Now, if we chose $\zeta_1 = K_3\zeta_2 = \delta_1/4$ and we go back to (3.25) and (3.26), using $\left|\frac{\partial\psi}{\partial t}\right| \leq \Psi$ we get

$$\begin{aligned}&\frac{1}{2}\int_{\mathcal{O}} w^2(t)\varphi_n d\mathfrak{m} + \frac{\delta_1}{4}\int_t^T \|w(s)\|^2_{V,n} ds + \frac{1}{\varepsilon}\int_t^T ds\int_{\mathcal{O}}(-w(s))^2_+\varphi_n d\mathfrak{m}\\ &\leq \frac{4}{\delta_1}\left(\|\sqrt{1+y}g\|^2_{L^2([t,T];H)} + \|\sqrt{1+y}\Psi\|^2_{L^2([t,T];H)}\right) + \frac{2K_3^2}{\delta_1}\int_t^T \|\psi(s)\|^2_{V,n} ds + K_2\|w\|^2_{L^2([t,T];H)} ds,\\ &\leq \frac{4}{\delta_1}\left(\|\sqrt{1+y}g\|^2_{L^2([t,T];H)} + \|\sqrt{1+y}\Psi\|^2_{L^2([t,T];H)}\right) + \frac{4K_3^2}{\delta_1}\left\|\sqrt{1+y}\psi\right\|^2_{L^2([t,T];V)} + K_2\|w\|^2_{L^2([t,T];H)},\end{aligned} \tag{3.27}$$

where the last inequality follows from the estimate $\|v\|^2_{V,n} \leq 2\|\sqrt{1+y}v\|^2_V$, and, in the case $n = 0$,

$$\frac{1}{2}\|w(t)\|^2_H + \frac{\delta_1}{4}\int_t^T \|w(s)\|^2_V ds + \frac{1}{\varepsilon}\int_t^T ds\int_{\mathcal{O}}(-w(s))^2_+ d\mathfrak{m} \leq \frac{4}{\delta_1}\left(\|g\|^2_{L^2([t,T];H)} + \|\Psi\|^2_{L^2([t,T];H)}\right) + \frac{2K_3^2}{\delta_1}\|\psi\|^2_{L^2([t,T];V)}. \tag{3.28}$$

From (3.28) recalling that $w = u - \psi$ we deduce

$$\int_t^T \|u(s)\|_V^2 ds \leq \int_t^T 2(\|w(s)\|_V^2 + \|\psi(s)\|_V^2) ds \leq \frac{32}{\delta_1^2}\left(\|g\|^2_{L^2([t,T];H)} + \|\Psi\|^2_{L^2([t,T];H)}\right) + \left(\frac{16K_3^2}{\delta_1^2} + 2\right)\|\psi\|^2_{L^2([t,T];V)}. \tag{3.29}$$

Moreover, combining (3.27) and (3.28), we have

$$\begin{aligned}
&\frac{1}{2}\int_{\mathcal{O}} w^2(t)\varphi_n d\mathfrak{m} + \frac{\delta_1}{4}\int_t^T \|w(s)\|^2_{V,n} ds + \frac{1}{\varepsilon}\int_t^T ds \int_{\mathcal{O}} (-w(s))_+^2 \varphi_n d\mathfrak{m} \\
&\leq \left(\frac{4}{\delta_1} + \frac{16K_2}{\delta_1^2}\right)\left(\|\sqrt{1+y}g\|^2_{L^2([t,T];H)} + \|\sqrt{1+y}\Psi\|^2_{L^2([t,T];H)}\right) + \frac{4K_3^2}{\delta_1}\left(1 + \frac{2K_2}{\delta_1}\right)\|\sqrt{1+y}\psi\|^2_{L^2([t,T];V)}.
\end{aligned}$$

In particular,

$$\begin{aligned}
&\int_t^T ds \int_{\mathcal{O}} y|\nabla u(s)|^2 \varphi_n d\mathfrak{m} \leq \int_t^T \|u(s)\|^2_{V,n} ds \leq 2\int_t^T \|w(s)\|^2_{V,n} ds + 2\int_t^T ds \|\psi(s)\|^2_{V,n} ds \\
&\leq \frac{8}{\delta_1}\left(\frac{4}{\delta_1} + \frac{16K_2}{\delta_1^2}\right)\left(\|\sqrt{1+y}g\|^2_{L^2([t,T];H)} + \|\sqrt{1+y}\Psi\|^2_{L^2([t,T];H)}\right) \\
&\quad + \left(\frac{32K_3^2}{\delta_1^2}\left(1 + \frac{2K_2}{\delta_1}\right) + 4\right)\|\sqrt{1+y}\psi\|^2_{L^2([t,T];V)}
\end{aligned}$$

and, by using the Monotone convergence theorem, we deduce

$$\int_t^T \||y|\nabla u(s)|\|_H^2 ds \leq K_4\left(\|\sqrt{1+y}g\|^2_{L^2([t,T];H)} + \|\sqrt{1+y}\Psi\|^2_{L^2([t,T];H)} + \|\sqrt{1+y}\psi\|^2_{L^2([t,T];V)}\right), \tag{3.30}$$

where $K_4 = \frac{8}{\delta_1}\left(\frac{4}{\delta_1} + \frac{16K_2}{\delta_1^2}\right) \vee \left(\frac{32K_3^2}{\delta_1^2}\left(1 + \frac{2K_2}{\delta_1}\right) + 4\right)$.

We are now in a position to prove (3.24). Taking $v = \frac{\partial u}{\partial t}$ in (3.18), we have

$$-\left\|\frac{\partial u}{\partial t}(t)\right\|_H^2 + \bar{a}_\lambda\left(u(t), \frac{\partial u}{\partial t}(t)\right) + \tilde{a}^{(M)}\left(u(t), \frac{\partial u}{\partial t}(t)\right) - \frac{1}{\varepsilon}\left((\psi(t) - u(t))_+, \frac{\partial u}{\partial t}(t)\right)_H = \left(g(t), \frac{\partial u}{\partial t}(t)\right)_H.$$

Note that, since $\bar{a}_\lambda$ is symmetric, $\frac{d}{dt}\bar{a}_\lambda(u(t), u(t)) = 2\bar{a}_\lambda\left(u(t), \frac{\partial u}{\partial t}(t)\right)$. On the other hand,

$$\left((\psi(t) - u(t))_+, \frac{\partial u}{\partial t}\right)_H = -\frac{1}{2}\frac{d}{dt}\|(\psi(t) - u(t))_+\|_H^2 + \left((\psi(t) - u(t))_+, \frac{\partial \psi}{\partial t}(t)\right)_H,$$

so that

$$\begin{aligned}
&\left\|\frac{\partial u}{\partial t}(t)\right\|_H^2 - \frac{1}{2}\frac{d}{dt}\bar{a}_\lambda(u(t), u(t)) - \frac{1}{2\varepsilon}\frac{d}{dt}\|(\psi(t) - u(t))_+\|_H^2 \\
&\quad = \tilde{a}^{(M)}\left(u(t), \frac{\partial u}{\partial t}(t)\right) - \left(g(t), \frac{\partial u}{\partial t}(t)\right)_H - \frac{1}{\varepsilon}\left((\psi(t) - u(t))_+, \frac{\partial \psi}{\partial t}(t)\right)_H \\
&\quad \leq \left|\tilde{a}^{(M)}\left(u(t), \frac{\partial u}{\partial t}(t)\right)\right| + \|g(t)\|_H \left\|\frac{\partial u}{\partial t}(t)\right\|_H + \frac{1}{\varepsilon}\left((\psi(t) - u(t)_+, \Psi(t)\right)_H \\
&\quad \leq \left(K_1\||y|\nabla u(t)|\|_H + \|g(t)\|_H\right)\left\|\frac{\partial u}{\partial t}(t)\right\|_H + \frac{1}{\varepsilon}\left((\psi(t) - u(t))_+, \Psi(t)\right)_H.
\end{aligned}$$

Moreover, if we take $v = \Psi(t)$ in (3.18), we get

$$-\left(\frac{\partial u}{\partial t}(t), \Psi(t)\right)_H + a_\lambda^{(M)}(u(t), \Psi(t)) - \frac{1}{\epsilon}\left((\psi(t) - u(t))_+, \Psi(t)\right)_H = \left(g(t), \Psi(t)\right)_H,$$

so that

$$\frac{1}{\varepsilon}\left((\psi(t)-u(t))_+,\Psi(t)\right)_H \le \left\|\frac{\partial u}{\partial t}(t)\right\|_H \|\Psi(t)\|_H + \|a_\lambda^{(M)}\|\|u(t)\|_V\|\Psi(t)\|_V + \|g(t)\|_H\|\Psi(t)\|_H. \tag{3.31}$$

Therefore,

$$\begin{aligned}
&\left\|\frac{\partial u}{\partial t}(t)\right\|_H^2 - \frac{1}{2}\frac{d}{dt}\bar{a}_\lambda\left(u(t),u(t)\right) - \frac{1}{2\varepsilon}\frac{d}{dt}\|(\psi(t)-u(t))_+\|_H^2 \\
&\qquad \le \left(K_1\left\||y|\nabla u(t)|\right\|_H + \|g(t)\|_H + \|\Psi(t)\|_H\right)\left\|\frac{\partial u}{\partial t}(t)\right\|_H + \|a_\lambda^{(M)}\|\|u(t)\|_V\|\Psi(t)\|_V + \|g(t)\|_H\|\Psi(t)\|_H,
\end{aligned}$$

hence

$$\begin{aligned}
&\frac{1}{2}\left\|\frac{\partial u}{\partial t}(t)\right\|_H^2 - \frac{1}{2}\frac{d}{dt}\bar{a}_\lambda\left(u(t),u(t)\right) - \frac{1}{2\varepsilon}\frac{d}{dt}\|(\psi(t)-u(t))_+\|_H^2 \\
&\qquad \le \frac{1}{2}\left(K_1\left\||y|\nabla u(t)|\right\|_H + \|g(t)\|_H + \|\Psi(t)\|_H\right)^2 + \|a_\lambda^{(M)}\|\|u(t)\|_V^2\|\Psi(t)\|_V^2 + \|g(t)\|_H\|\Psi(t)\|_H.
\end{aligned}$$

Integrating between t and T, we get,

$$\begin{aligned}
&\frac{1}{2}\left\|\frac{\partial u}{\partial s}\right\|_{L^2([t,T];H)}^2 + \frac{1}{2}\bar{a}_\lambda\left(u(t),u(t)\right) + \frac{1}{2\varepsilon}\|(\psi(t)-u(t))_+\|_H^2 \le \frac{1}{2}\bar{a}_\lambda(\psi(T),\psi(T)) + 2\|g\|_{L^2([t,T];H)}^2 \\
&\qquad + 2\|\Psi\|_{L^2([t,T];H)}^2 + \frac{3K_1^2}{2}\left\||y|\nabla u|\right\|_{L^2([t,T];H)}^2 + \frac{\|a_\lambda^{(M)}\|}{2}\|u\|_{L^2([t,T];V)} + \frac{\|a_\lambda^{(M)}\|}{2}\|\Psi\|_{L^2([t,T];V)},
\end{aligned}$$

so, recalling that $\bar{a}_\lambda(u(t),u(t) \ge \delta_1\int_{\mathcal{O}} y|\nabla u(t)|^2 d\mathfrak{m} + \lambda\int_{\mathcal{O}}(1+y)u^2 d\mathfrak{m} \ge (\delta_1\wedge\lambda)\|u(t)\|_V^2$,

$$\begin{aligned}
&\frac{1}{2}\left\|\frac{\partial u}{\partial s}\right\|_{L^2([t,T];H)}^2 + \frac{\delta_1\wedge\lambda}{2}\|u(t)\|_V^2 + \frac{1}{2\varepsilon}\|(\psi(t)-u(t))_+\|_H^2 \\
&\le \frac{\|\bar{a}_\lambda\|}{2}\|\psi(T)\|_V^2 + 2\|g\|_{L^2([t,T];H)}^2 + 2\|\Psi\|_{L^2([t,T];H)}^2 \\
&\quad + \frac{3K_1^2}{2}\left\||y|\nabla u|\right\|_{L^2([t,T];H)}^2 + \frac{\|a_\lambda^{(M)}\|}{2}\|u\|_{L^2([t,T];V)} + \frac{\|a_\lambda^{(M)}\|}{2}\|\Psi\|_{L^2([t,T];V)}^2 \\
&\le \frac{\|\bar{a}_\lambda\|}{2}\|\psi(T)\|_V^2 + 2\|g\|_{L^2([t,T];H)}^2 + 2\|\Psi\|_{L^2([t,T];H)}^2 \\
&\quad + \frac{3K_1^2}{2}K_4\left(\|\sqrt{1+y}g\|_{L^2([t,T];H)}^2 + \|\sqrt{1+y}\Psi\|_{L^2([t,T];H)}^2 + \|\sqrt{1+y}\psi\|_{L^2([t,T];V)}^2\right) \\
&\quad + \frac{\|a_\lambda^{(M)}\|}{2}\left(\frac{32}{\delta_1^2}\left(\|g\|_{L^2([t,T];H)}^2 + \|\Psi\|_{L^2([t,T];H)}^2\right) + \left(\frac{16K_3^2}{\delta_1^2}+2\right)\|\psi\|_{L^2([t,T];V)}^2\right) + \frac{\|a_\lambda^{(M)}\|}{2}\|\Psi\|_{L^2([t,T];V)}^2,
\end{aligned}$$

where the last inequality follows from (3.29) and (3.30). Rearranging the terms, we deduce that there exists a constant $C>0$ independent of M and ε such that

$$\begin{aligned}
&\frac{1}{2}\left\|\frac{\partial u}{\partial s}\right\|_{L^2([t,T];H)}^2 + \frac{\delta_1\wedge\lambda}{4}\|u(t)\|_V^2 + \frac{1}{2\varepsilon}\|(\psi(t)-u(t))_+\|_H^2 \\
&\qquad\qquad \le C\left(\|\sqrt{1+y}g\|_{L^2([t,T];H)}^2 + \|\Psi\|_{L^2([t,T];V)}^2 + \left\|\sqrt{1+y}\psi\right\|_{L^2([t,T];V)}^2 + \|\psi(T)\|_V^2\right),
\end{aligned}$$

which concludes the proof. □

Proof of Theorem 3.8: existence. Assume for a first moment that we have the further assumptions $\psi(T)\in H^2(\mathcal{O},\mathfrak{m})$, $(1+y)\psi(T)\in H$, $\frac{\partial\psi}{\partial t}\in L^2([0,T];V)$ and $\frac{\partial g}{\partial t}\in L^2([0,T];H)$. Thanks to (3.24) we can repeat the

same arguments as in the proof of Proposition 3.9 in order to pass to the limit in j, but this time as $M \to \infty$. Therefore, we deduce the existence of a function $u_{\varepsilon,\lambda} \in L^2([0,T];V)$ with $\frac{\partial u_{\varepsilon,\lambda}}{\partial t} \in L^2([0,T];H)$ and such that

$$-\left(\frac{\partial u_{\varepsilon,\lambda}}{\partial t}(t), v\right)_H + a_\lambda(u_{\varepsilon,\lambda}(t), v)_H - \frac{1}{\varepsilon}((\psi(t) - u_{\varepsilon,\lambda}(t))_+, v)_H = (g(t), v)_H.$$

The estimates (3.15), (3.16) and (3.17) directly follow from (3.24) as $M \to \infty$.

We have now to weaken the assumptions on g and ψ. We can do this by a regularization procedure. In fact, let us assume that ψ satisfies Assumption $\mathcal{H}^1$ (so, in particular, $\left|\frac{\partial \psi}{\partial t}\right| \le \Psi$ for a certain $\Psi \in L^2([0,T];V)$) and g satisfies Assumption $\mathcal{H}^0$. Then, by standard regularization techniques (see for example [5, Corollary A.12]), we can find sequences of functions $(g_n)_n$, $(\psi_n)_n$ and $(\Psi_n)_n$ of class C^∞ with compact support such that, for any $n \in \mathbb{N}$, $n \in \mathbb{N}$, $|\frac{\partial \psi_n}{\partial t}| \le \Psi_n$ and all the regularity assumptions required in the first part of the proof are satisfied. Moreover, $\|\sqrt{1+y}g_n - \sqrt{1+y}g\|_{L^2([0,T];H)} \to 0$, $\|\sqrt{1+y}\psi_n - \sqrt{1+y}\psi\|_{L^2([0,T];V)} \to 0$, $\|\Psi_n - \Psi\|_{L^2([0,T];V)} \to 0$, $\|\psi_n(T) - \psi(T)\|_V \to 0$ as $n \to \infty$ (we refer to [17] for the details). Therefore, the solution $u^n_{\varepsilon,\lambda,M}$ of the equation (3.14) with source function g_n and obstacle function ψ_n satisfies

$$\begin{aligned} &\int_t^T \left\|\tfrac{\partial u^n_{\varepsilon,\lambda,M}}{\partial s}(s)\right\|_H^2 ds + \|u^n_{\varepsilon,\lambda,M}(t)\|_V^2 + \tfrac{1}{\varepsilon}\|(\psi_n(t) - u^n_{\varepsilon,\lambda,M}(t))_+\|_H^2 \\ &\le C\left(\|\sqrt{1+y}g_n\|_{L^2([0,T];H)} + \|\sqrt{1+y}\psi_n\|^2_{L^2([0,T];V)} + \|\Psi_n\|^2_{L^2([0,T];V)} + \|\psi_n(T)\|_V^2\right). \end{aligned} \tag{3.32}$$

Then, we can take the limit for $n \to \infty$ in (3.32) and the assertion follows as in the first part of the proof. □

Moreover, we have the following Comparison principle for the coercive penalized problem.

Proposition 3.11. *1. Assume that ψ_i satisfies Assumption $\mathcal{H}^1$ for $i = 1, 2$ and g satisfies Assumption $\mathcal{H}^0$. Let $u^i_{\varepsilon,\lambda}$ be the unique solution of (3.14) with obstacle function ψ_i and source function g. If $\psi_1 \le \psi_2$, then $u^1_{\varepsilon,\lambda} \le u^2_{\varepsilon,\lambda}$.*

2. Assume that ψ satisfies Assumption $\mathcal{H}^1$ and g_i satisfy Assumption $\mathcal{H}^0$ for $i = 1, 2$. Let $u^i_{\varepsilon,\lambda}$ be the unique solution of (3.14) with obstacle function ψ and source function g_i. If $g_1 \le g_2$, then $u^1_{\varepsilon,\lambda} \le u^2_{\varepsilon,\lambda}$.

3. Assume that ψ_i satisfies Assumption $\mathcal{H}^1$ for $i = 1, 2$ and g satisfies Assumption $\mathcal{H}^0$. Let $u^i_{\varepsilon,\lambda}$ be the unique solution of (3.14) with obstacle function ψ_i and source function g. If $\psi_1 - \psi_2 \in L^\infty$, then $u^1_{\varepsilon,\lambda} - u^2_{\varepsilon,\lambda} \in L^\infty$ and $\|u^1_{\varepsilon,\lambda} - u^2_{\varepsilon,\lambda}\|_\infty \le \|\psi_1 - \psi_2\|_\infty$.

Proposition 3.11 can be proved with standard techniques introduced in [4, Chapter 3] so we omit the proof.

3.2.2 Coercive variational inequality

Proposition 3.12. *Assume that ψ satisfies Assumption $\mathcal{H}^1$ and g satisfies Assumption $\mathcal{H}^0$. Moreover, assume that $0 \le \psi \le \Phi$ with $\Phi \in L^2([0,T];H^2(\mathcal{O},\mathfrak{m}))$ such that $\frac{\partial \Phi}{\partial t} + \mathcal{L}\Phi \le 0$ and $0 \le g \le -\frac{\partial \Phi}{\partial t} - \mathcal{L}^\lambda \Phi$. Then, there exists a unique function u_λ such that $u_\lambda \in L^2([0,T];V)$, $\frac{\partial u_\lambda}{\partial t} \in L^2([0,T];H)$ and*

$$\begin{cases} -\left(\frac{\partial u_\lambda}{\partial t}, v - u_\lambda\right)_H + a_\lambda(u_\lambda, v - u_\lambda) \ge (g, v - u_\lambda)_H, & \text{a.e. in } [0,T] \qquad v \in L^2([0,T];V),\ v \ge \psi, \\ u_\lambda(T) = \psi(T), \\ u_\lambda \ge \psi \text{ a.e. in } [0,T] \times \mathbb{R} \times (0,\infty). \end{cases} \tag{3.33}$$

Moreover, $0 \le u_\lambda \le \Phi$.

Proof. The uniqueness of the solution of (3.33) follows by a standard monotonicity argument introduced in [4, Chapter 3] (see [17]). As regards the existence of a solution, we follow the lines of the proof of [4, Theorem 2.1] but we repeat here the details since we use a compactness argument which is not present in the classical theory.

For each fixed $\varepsilon > 0$ we have the estimates (3.15) and (3.16), so, for every $t \in [0,T]$, we can extract a subsequence $u_{\varepsilon,\lambda}$ such that $u_{\varepsilon,\lambda}(t) \rightharpoonup u_\lambda(t)$ in V as $\varepsilon \to 0$ and $u'_\varepsilon(t) \rightharpoonup u'_\lambda(t)$ in H for some function $u_\lambda \in V$.

Note that $u = 0$ is the unique solution of (3.14) when $\psi = g = 0$, while $u = \Phi$ is the unique solution of (3.14) when $\psi = \Phi$ and $g = -\frac{\partial \Phi}{\partial t} - \mathcal{L}^\lambda \Phi = -\frac{\partial \Phi}{\partial t} - \mathcal{L}\Phi + \lambda(1+y)\Phi$. Therefore, Proposition 3.11 implies that $0 \le u_{\varepsilon,\lambda} \le \Phi$. Recall that $u_{\varepsilon,\lambda}(t) \to u_\lambda(t)$ in $L^2(\mathcal{U}, \mathfrak{m})$ for every relatively compact open $\mathcal{U} \subset \mathcal{O}$. This, together with the fact that $d\mathfrak{m}$ is a finite measure, allows to conclude that we have strong convergence of $u_{\varepsilon,\lambda}$ to u_λ in H. In fact, if $\delta > 0$ and $\mathcal{O}_\delta := (-\frac{1}{\delta}, \frac{1}{\delta}) \times (\delta, \frac{1}{\delta})$,

$$\begin{aligned}\int_0^T ds \int_{\mathcal{O}} |u_{\varepsilon,\lambda}(s) - u_\lambda(s)|^2 d\mathfrak{m} &\le \int_0^T ds \int_{\mathcal{O}_\delta} |u_{\varepsilon,\lambda}(s) - u_\lambda(s)|^2 d\mathfrak{m} + \int_0^T ds \int_{\mathcal{O}_\delta^c} |u_{\varepsilon,\lambda}(s) - u_\lambda(s)|^2 d\mathfrak{m} \\ &\le \int_0^T ds \int_{\mathcal{O}_\delta} |u_{\varepsilon,\lambda}(s) - u_\lambda(s)|^2 d\mathfrak{m} + \int_0^T ds \int_{\mathcal{O}_\delta^c} 4\Phi^2(s) d\mathfrak{m}\end{aligned}$$

and it is enough to let δ goes to 0.

From (3.17) we also have that $(\psi(t) - u_{\varepsilon,\lambda}(t))^+ \to 0$ strongly in H as $\varepsilon \to 0$. On the other hand $(\psi(t) - u_{\varepsilon,\lambda}(t))_+ \rightharpoonup \chi(t)$ weakly in H and $\chi = (\psi - u_\lambda)_+$ since there exists a subsequence of $u_{\varepsilon,\lambda}(t)$ which converges pointwise to $u_\lambda(t)$. Therefore, $(\psi(t) - u_\lambda(t))^+ = 0$, which means $u_\lambda(t) \ge \psi(t)$.

Then we consider the penalized coercive equation in (3.14) replacing v by $v - u_{\varepsilon,\lambda}(t)$, with $v \ge \psi(t)$. Since $\zeta_\varepsilon(t, v) = 0$ and $(\zeta_\varepsilon(t, v) - \zeta_\varepsilon(t, u_{\varepsilon,\lambda}(t)), v - u_{\varepsilon,\lambda}(t))_H \ge 0$ we easily deduce that

$$-\left(\frac{\partial u_{\varepsilon,\lambda}}{\partial t}(t), v - u_{\varepsilon,\lambda}(t)\right)_H + a_\lambda(u_{\varepsilon,\lambda}(t), v - u_{\varepsilon,\lambda}(t)) \ge (g(t), v - u_{\varepsilon,\lambda}(t))_H$$

so that, letting ε goes to 0, we have

$$\begin{aligned}-\left(\frac{\partial u_\lambda}{\partial t}(t), v - u_\lambda(t)\right)_H + a_\lambda(u_\lambda(t), v) &\ge (g(t), v - u_\lambda(t))_H + \liminf_{\varepsilon \to 0} a_\lambda(u_{\varepsilon,\lambda}(t), u_{\varepsilon,\lambda}(t)) \\ &\ge (g(t), v - u_\lambda(t))_H + a_\lambda(u_\lambda(t), u_\lambda(t)).\end{aligned}$$

Moreover, since $0 \le u_{\varepsilon,\lambda} \le \Phi$ for every $\varepsilon > 0$ and $u_\lambda = \lim_{\varepsilon \to 0} u_{\varepsilon,\lambda}$, we have $0 \le u_\lambda \le \Phi$ and the assertion follows. □

The following Comparison Principle is a direct consequence of Proposition 3.11,.

Proposition 3.13. *1. For $i = 1, 2$, assume that ψ_i satisfies Assumption $\mathcal{H}^1$, g satisfies Assumption $\mathcal{H}^0$ and $0 \le \psi_i \le \Phi$ with $\Phi \in L^2([0,T]; H^2(\mathcal{O}, \mathfrak{m}))$ such that $\frac{\partial \Phi}{\partial t} + \mathcal{L}\Phi \le 0$ and $0 \le g \le -\frac{\partial \Phi}{\partial t} - \mathcal{L}^\lambda \Phi$. Let u_λ^i be the unique solution of (3.33) with obstacle function ψ_i and source function g. If $\psi_1 \le \psi_2$, then $u_\lambda^1 \le u_\lambda^2$.*

2. For $i = 1, 2$, assume that ψ satisfies Assumption $\mathcal{H}^1$, g_i satisfy Assumption $\mathcal{H}^0$ and $0 \le \psi \le \Phi$ with $\Phi \in L^2([0,T]; H^2(\mathcal{O}, \mathfrak{m}))$ such that $\frac{\partial \Phi}{\partial t} + \mathcal{L}\Phi \le 0$ and $0 \le g_i \le -\frac{\partial \Phi}{\partial t} - \mathcal{L}^\lambda \Phi$. Let u_λ^i be the unique solution of (3.33) with obstacle function ψ and source function g_i. If $g_1 \le g_2$, then $u_\lambda^1 \le u_\lambda^2$.

3. For $i = 1, 2$, assume that ψ_i satisfies Assumption $\mathcal{H}^1$, g satisfies Assumption $\mathcal{H}^0$ and $0 \le \psi_i \le \Phi$ with $\Phi \in L^2([0,T]; H^2(\mathcal{O}, \mathfrak{m}))$ such that $\frac{\partial \Phi}{\partial t} + \mathcal{L}\Phi \le 0$ and $0 \le g \le -\frac{\partial \Phi}{\partial t} - \mathcal{L}^\lambda \Phi$. Let u_λ^i be the unique solution of (3.33) with obstacle function ψ_i and source function g. If $\psi_1 - \psi_2 \in L^\infty$, then $u_\lambda^1 - u_\lambda^2 \in L^\infty$ and $\|u_\lambda^1 - u_\lambda^2\|_\infty \le \|\psi_1 - \psi_2\|_\infty$.

3.2.3 Non-coercive variational inequality

We can finally prove Theorem 2.3. Again, we first study the uniqueness of the solution and then we deal with the existence.

Proof of uniqueness in Theorem 2.3. Suppose that there are two functions u_1 and u_2 which satisfy (2.5). As usual, we take $v = u_2$ in the equation satisfied by u_1 and $v = u_1$ in the one satisfied by u_2 and we add the resulting equations. Setting $w := u_2 - u_1$, we get that, a.e. in $[0, T]$,

$$\left(\frac{\partial w}{\partial t}(t), w(t)\right)_H - a(w(t), w(t)) \ge 0.$$

From the energy estimate (3.4), we know that

$$a(u(t),u(t)) \geq C_1\|u(t)\|_V^2 - C_2\|(1+y)^{\frac{1}{2}}u(t)\|_H^2,$$

so that

$$\frac{1}{2}\frac{d}{dt}\|w(t)\|_H^2 + C_2\|(1+y)^{\frac{1}{2}}w(t)\|_H^2 \geq 0.$$

By integrating from t to T, since $w(T)=0$, we have

$$\begin{aligned}
\|w(t)\|_H^2 &\leq C_2\int_t^T \|(1+y)^{\frac{1}{2}}w(s)\|_H^2 ds \\
&\leq C_2\bigg(\int_t^T ds\int_{\mathcal{O}} \mathbb{1}_{\{y\leq\lambda\}}(1+y)w^2(s)d\mathfrak{m} + \int_t^T ds\int_{\mathcal{O}} \mathbb{1}_{\{y>\lambda\}}(1+y)w^2(s)d\mathfrak{m}\bigg) \\
&\leq C\bigg(\int_t^T ds\int_{\mathcal{O}} (1+\lambda)w^2(s)y^{\beta-1}e^{-\gamma|x|}e^{-\mu y}dxdy + \int_t^T ds\int_{\mathcal{O}} \mathbb{1}_{\{y>\lambda\}}(1+y)w^2(s)y^{\beta-1}e^{-\gamma|x|}e^{-(\mu-\mu')y}e^{-\mu' y}dxdy\bigg) \\
&\leq C\bigg(\int_t^T ds\int_{\mathcal{O}} dxdy(1+\lambda)w^2(s)y^{\beta-1}e^{-\gamma|x|}e^{-\mu y} + e^{-(\mu-\mu')\lambda}\int_t^T ds\int_{\mathcal{O}} dxdy(1+y)\Phi^2(s)y^{\beta-1}e^{-\gamma|x|}e^{-\mu' y}\bigg),
\end{aligned}$$

where $\mu' < \mu$ and $\lambda > 0$. Since $C_2 = \int_{\mathcal{O}} dxdy(1+y)\Phi^2(s)y^{\beta-1}e^{-\gamma|x|}e^{-\mu' y} < \infty$, we have

$$\|w(t)\|_H^2 \leq C(1+\lambda)\int_t^T \|w(s)\|_H^2 ds + C_2(T-t)e^{-(\mu-\mu')\lambda},$$

so, by using the Gronwall Lemma,

$$\|w(t)\|_H^2 \leq C_2 T e^{-(\mu-\mu')\lambda + C(T-t)(1+\lambda)}.$$

Sending $\lambda\to\infty$, we deduce that $w(t)=0$ in $[T,t]$ for t such that $T-t < \frac{\mu-\mu'}{C}$. Then, we iterate the same argument: we integrate between t' and t with $t-t' < \frac{\mu-\mu'}{C}$ and we have $w(t)=0$ in $[T,t']$ and so on. We deduce that $w(t)=0$ for all $t\in[0,T]$ so the assertion follows. □

Proof of existence in Theorem 2.3. Given $u_0=\Phi$, we can construct a sequence $(u_n)_n \subset V$ such that

$$u_n \geq \psi \text{ a.e. in } [0,T]\times\mathcal{O}, \qquad n\geq 1, \tag{3.34}$$

$$\begin{aligned}
-\left(\frac{\partial u_n}{\partial t}, v-u_n\right)_H + a(u_n, v-u_n) + \lambda((1+y)u_n, v-u_n)_H \geq \lambda((1+y)u_{n-1}, v-u_n)_H, \\
v\in V, \quad v\geq\psi, \quad \text{a.e. on } [0,T]\times\mathcal{O}, \qquad n\geq 1,
\end{aligned} \tag{3.35}$$

$$u_n(T) = \psi(T), \qquad \text{in } \mathcal{O}, \tag{3.36}$$

$$\Phi \geq u_1 \geq u_2 \geq \cdots \geq u_{n-1} \geq u_n \geq \cdots \geq 0, \qquad \text{a.e. on } [0,T]\times\mathcal{O}. \tag{3.37}$$

In fact, if we have $0\leq u_{n-1}\leq\Phi$ for all $n\in\mathbb{N}$, then the assumptions of Proposition 3.12 are satisfied with

$$g_n = \lambda(1+y)u_{n-1}.$$

Indeed, since $(1+y)^{\frac{3}{2}}\Phi \in L^2([0,T];H)$, we have that g_n and $\sqrt{1+y}g_n$ belong to $L^2([0,T];H)$ and, moreover, $0\leq g_n \leq \lambda(1+y)\Phi \leq -\frac{\partial\Phi}{\partial t} - \mathcal{L}_\lambda\Phi$. Therefore, step by step, we can deduce the existence and the uniqueness of a solution u_n to (3.35) such that $0\leq u_n\leq\Phi$. (3.37) is a simple consequence of Proposition 3.13. In fact, proceeding by induction, at each step we have

$$g_n = \lambda(1+y)u_{n-1} \leq \lambda(1+y)u_{n-2} = g_{n-1}$$

so that $u_n \leq u_{n-1}$. Now, recall that

$$\|u_n\|_{L^\infty([0,T],V)} \leq K,$$

$$\left\| \frac{\partial u_n}{\partial t} \right\|_{L^2([0,T];H)} \leq K,$$

where $K = C\left(\|\Psi\|_{L^2([0,T];V)} + \|\sqrt{1+y}g_n\|_{L^2([0,T];H)} + \|\sqrt{1+y}\psi\|_{L^2([0,T];V)} + \|\psi(T)\|_V\right)$. Note that the constant K is independent of n since $|g_n| = |\lambda(1+y)u_{n-1},| \leq \lambda(1+y)\Phi$, for every $n \in \mathbb{N}$. Therefore, by passing to a subsequence, we can assume that there exists a function u such that $u \in L^2([0,T];V)$, $\frac{\partial u}{\partial t} \in L^2([0,T];H)$ and for every $t \in [0,T]$, $u_n'(t) \rightharpoonup u'(t)$ in H and $u_n(t) \rightharpoonup u(t)$ in V. Indeed, again thanks to the fact that $0 \leq u_n \leq \Phi$, we can deduce that $u_n(t) \to u(t)$ in H. Therefore we can pass to the limit in

$$-\left(\frac{\partial u_n}{\partial t}, u_n - v\right)_H + a(u_n, v-u_n) + \lambda((1+y)u_n, v-u_n)_H \geq \lambda((1+y)u_{n-1}, v-u_n)_H$$

and the assertion follows. □

Remark 3.14. *Keeping in mind our purpose of identifying the solution of the variational inequality* (2.5) *with the American option price we have considered the case without source term ($g = 0$) in the variational inequality* (2.5)*. However, under the same assumptions of Theorem 2.3, we can prove in the same way the existence and the uniqueness of a solution of*

$$\begin{cases} -\left(\frac{\partial u}{\partial t}, v-u\right)_H + a(u, v-u) \geq (g, v-u)_H, \quad \text{a.e. in } [0,T] \quad v \in L^2([0,T];V),\ v \geq \psi, \\ u \geq \psi \text{ a.e. in } [0,T] \times \mathbb{R} \times (0,\infty), \\ u(T) = \psi(T), \\ 0 \leq u \leq \Phi, \end{cases}$$

where g satisfies Assumption $\mathcal{H}^0$ and $0 \leq g \leq -\frac{\partial \Phi}{\partial t} - \mathcal{L}\Phi$.

We conclude stating the following Comparison Principle, whose proof is a direct consequence of Proposition 3.13 and the proof of Proposition 2.3.

Proposition 3.15. *For $i = 1, 2$, assume that ψ_i satisfies Assumption $\mathcal{H}^1$ and $0 \leq \psi_i \leq \Phi$ with Φ satisfying Assumption $\mathcal{H}^2$. Let u_λ^i be the unique solution of* (3.33) *with obstacle function ψ_i. Then:*

1. *If $\psi_1 \leq \psi_2$, then $u_\lambda^1 \leq u_\lambda^2$.*
2. *If $\psi_1 - \psi_2 \in L^\infty$, then $u_\lambda^1 - u_\lambda^2 \in L^\infty$ and $\|u_\lambda^1 - u_\lambda^2\|_\infty \leq \|\psi_1 - \psi_2\|_\infty$.*

4 Connection with the optimal stopping problem

Once we have the existence and the uniqueness of a solution u of the variational inequality (2.3), our aim is to prove that it matches the solution of the optimal stopping problem, that is

$$u(t,x,y) = u^*(t,x,y), \qquad \text{on } [0,T] \times \bar{\mathcal{O}},$$

where u^* is defined by

$$u^*(t,x,y) = \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}\left[\psi(\tau, X_\tau^{t,x,y}, Y_\tau^{t,y})\right],$$

$\mathcal{T}_{t,T}$ being the set of the stopping times with values in $[t,T]$. Since the function u is not regular enough to apply Itô's Lemma, we use another strategy in order to prove the above identification. So, we first show, by using the affine character of the underlying diffusion, that the semigroup associated with the bilinear form a_λ coincides with the transition semigroup of the two dimensional diffusion (X,Y) with a killing term. Then, we prove suitable estimates on the joint law of (X,Y) and L^p-regularity results on the solution of the variational inequality and we deduce from them the probabilistic interpretation.

4.1 Semigroup associated with the bilinear form

We introduce now the semigroup associated with the coercive bilinear form a_λ. With a natural notation, we define the following spaces

$$L^2_{loc}(\mathbb{R}^+;H)=\{f:\mathbb{R}^+\to H:\forall t\geq 0\int_0^t\|f(s)\|_H^2ds<\infty\},$$

$$L^2_{loc}(\mathbb{R}^+;V)=\{f:\mathbb{R}^+\to V:\forall t\geq 0\int_0^t\|f(s)\|_V^2ds<\infty\}.$$

First of all, we state the following result:

Proposition 4.1. *For every $\psi\in V$, $f\in L^2_{loc}(\mathbb{R}^+;H)$ with $\sqrt{y}f\in L^2_{loc}(\mathbb{R}^+;H)$, there exists a unique function $u\in L^2_{loc}(\mathbb{R}^+;V)$ such that $\frac{\partial u}{\partial t}\in L^2_{loc}(\mathbb{R}^+;H)$, $u(0)=\psi$ and*

$$\left(\frac{\partial u}{\partial t},v\right)_H+a_\lambda(u,v)=(f,v)_H,\quad v\in V. \tag{4.1}$$

Moreover we have, for every $t\geq 0$,

$$\|u(t)\|_H^2+\frac{\delta_1}{2}\int_0^t\|u(s)\|_V^2ds\leq\|\psi\|_H^2+\frac{2}{\delta_1}\int_0^t\|f(s)\|_H^2ds \tag{4.2}$$

and

$$||u(t)||_V^2+\int_0^t||u_t(s)||_H^2ds\leq C\left(||\psi||_V^2+\frac{1}{2}\int_0^t||\sqrt{1+y}f(s)||_H^2ds\right),$$

with $C>0$.

The proof follows the same lines as the proof of Proposition 3.12 so we omit it. Moreover, we can prove a Comparison Principle for the equation (4.1) as we have done for the variational inequality.

We denote $u(t)=\bar{P}_t^\lambda\psi$ the solution of (4.1) corresponding to $u(0)=\psi$ and $f=0$. From (4.2) we deduce that the operator $\bar{P}_t^\lambda$ is a linear contraction on H and, from uniqueness, we have the semigroup property.

Proposition 4.2. *Let us consider $f:\mathbb{R}^+\to H$ such that $\sqrt{1+y}f\in L^2_{loc}(\mathbb{R}^+,H)$. Then, the solution of*

$$\begin{cases}\left(\frac{\partial u}{\partial t},v\right)_H+a_\lambda(u,v)=(f,v)_H,\quad v\in V,\\ u(0)=0,\end{cases}$$

is given by $u(t)=\int_0^t\bar{P}_s^\lambda f(t-s)ds=\int_0^t\bar{P}_{t-s}^\lambda f(s)ds$.

Proof. Note that V is dense in H and recall the estimate (4.2), so it is enough to prove the assertion for $f=\mathbb{1}_{(t_1,t_2]}\psi$, with $0\leq t_1<t_2$ and $\psi\in V$. If we set $u(t)=\int_0^t\bar{P}_{t-s}^\lambda f(s)ds$, we have

$$\begin{aligned}u(t)&=\mathbb{1}_{\{t\geq t_1\}}\int_{t_1}^{t\wedge t_2}\bar{P}_{t-s}^\lambda\psi ds\\&=\begin{cases}\int_{t_1}^{t_2}\bar{P}_{t-s}^\lambda\psi ds=\int_{t-t_2}^{t-t_1}\bar{P}_s^\lambda\psi ds & \text{if } t\geq t_2\\ \int_{t_1}^{t}\bar{P}_{t-s}^\lambda\psi ds=\int_{0}^{t-t_1}\bar{P}_s^\lambda\psi ds & \text{if } t\in[t_1,t_2)\end{cases}.\end{aligned}$$

Therefore, for every $v\in V$, we have $(u_t,v)_H+a_\lambda(u,v)=0$ if $t\leq t_1$ and, if $t\geq t_1$,

$$\left(\frac{\partial u}{\partial t},v\right)_H+a_\lambda(u(t),v)=\begin{cases}\left(\bar{P}_{t-t_1}^\lambda\psi-\bar{P}_{t-t_2}^\lambda\psi,v\right)_H+a_\lambda\left(\int_{t-t_2}^{t-t_1}\bar{P}_s^\lambda\psi ds,v\right) & \text{if } t\geq t_2\\ \left(\bar{P}_{t-t_1}^\lambda\psi,v\right)_H+a_\lambda\left(\int_0^{t-t_1}\bar{P}_s^\lambda\psi ds,v\right) & \text{if } t\in[t_1,t_2)\end{cases}.$$

The assertion follows from $(\bar{P}_t^\lambda\psi,v)_H+\int_0^t a_\lambda(\bar{P}_s\psi,v)ds=(\psi,v)_H$. □

Remark 4.3. *It is not difficult to prove that $\bar{P}_t^\lambda:L^p(\mathcal{O},\mathfrak{m})\to L^p(\mathcal{O},\mathfrak{m})$ is a contraction for every $p\geq 2$, and it is an analytic semigroup. This is not useful to our purposes so we omit the proof.*

4.2 Transition semigroup

We define $\mathbb{E}_{x_0,y_0}(\quad) = \mathbb{E}(\quad | X_0 = x_0, Y_0 = y_0)$. Fix $\lambda > 0$. For every measurable positive function f defined on $\mathbb{R} \times [0, +\infty)$, we define

$$P_t^\lambda f(x_0, y_0) = \mathbb{E}_{x_0,y_0}\left(e^{-\lambda \int_0^t (1+Y_s)ds} f(X_t, Y_t)\right).$$

The operator P_t^λ is the transition semigroup of the two dimensional diffusion (X, Y) with the killing term $e^{-\lambda \int_0^t (1+Y_s)ds}$.

Set $\mathbb{E}_{y_0}(\quad) = \mathbb{E}(\quad | Y_0 = y_0)$. We first prove some useful results about the Laplace transform of the pair $(Y_t, \int_0^t Y_s ds)$. These results rely on the affine structure of the model and have already appeared in slightly different forms in the literature (see, for example, [2, Section 4.2.1]). We include a proof for convenience.

Proposition 4.4. *Let z and w be two complex numbers with nonpositive real parts. The equation*

$$\psi'(t) = \frac{\sigma^2}{2}\psi^2(t) - \kappa\psi(t) + w \tag{4.3}$$

has a unique solution $\psi_{z,w}$ defined on $[0, +\infty)$, such that $\psi_{z,w}(0) = z$. Moreover, for every $t \geq 0$,

$$\mathbb{E}_{y_0}\left(e^{zY_t + w\int_0^t Y_s ds}\right) = e^{y_0\psi_{z,w}(t) + \theta\kappa\phi_{z,w}(t)},$$

with $\phi_{z,w}(t) = \int_0^t \psi_{z,w}(s)ds$.

Proof. Let ψ be the solution of (4.3). We define ψ_1 (resp. w_1) and ψ_2 (resp. w_2) the real and the imaginary part of ψ (resp. w). We have

$$\begin{cases} \psi_1'(t) = \frac{\sigma^2}{2}\left(\psi_1^2(t) - \psi_2^2(t)\right) - \kappa\psi_1(t) + w_1, \\ \psi_2'(t) = \sigma^2\psi_1(t)\psi_2(t) - \kappa\psi_2(t) + w_2. \end{cases}$$

From the first equation we deduce that $\psi_1'(t) \leq \frac{\sigma^2}{2}\left(\psi_1(t) - \frac{2\kappa}{\sigma^2}\right)\psi_1(t) + w_1$ and, since $w_1 \leq 0$, the function $t \mapsto \psi_1(t)e^{-\frac{\sigma^2}{2}\int_0^t (\psi_1(s) - \frac{2\kappa}{\sigma^2})ds}$ is nonincreasing. Therefore $\psi_1(t) \leq 0$ if $\psi_1(0) \leq 0$. Multiplying the first equation by $\psi_1(t)$ and the second one by $\psi_2(t)$ and adding we get

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\left(|\psi(t)|^2\right) &= \left(\frac{\sigma^2}{2}\psi_1(t) - \kappa\right)|\psi(t)|^2 + w_1\psi_1(t) + w_2\psi_2(t) \\ &\leq \left(\frac{\sigma^2}{2}\psi_1(t) - \kappa\right)|\psi(t)|^2 + |w||\psi(t)| \\ &\leq \left(\frac{\sigma^2}{2}\psi_1(t) - \kappa\right)|\psi(t)|^2 + \epsilon|\psi(t)|^2 + \frac{|w|^2}{4\epsilon}. \end{aligned}$$

We deduce that $|\psi(t)|$ cannot explode in finite time and, therefore, $\psi_{z,w}$ actually exists on $[0, +\infty)$.

Now, let us define the function $F_{z,w}(t, y) = e^{y\psi_{z,w}(t) + \theta\kappa\phi_{z,w}(t)}$. $F_{z,w}$ is $C^{1,2}$ on $[0, +\infty) \times \mathbb{R}$ and it satisfies by construction the following equation

$$\frac{\partial F_{z,w}}{\partial t} = \frac{\sigma^2}{2}y\frac{\partial^2 F_{z,w}}{\partial y^2} + \kappa(\theta - y)\frac{\partial F_{z,w}}{\partial y} + wyF_{z,w}.$$

Therefore, for every $T > 0$, the process $(M_t)_{0\leq t\leq T}$ defined by

$$M_t = e^{w\int_0^t Y_s ds} F_{z,w}(T - t, Y_t) \tag{4.4}$$

is a local martingale. On the other hand, note that

$$|M_t| = \left|e^{w\int_0^t Y_s ds}\right| \left|e^{Y_t\psi_{z,w}(T-t) + \theta\kappa\phi_{z,w}(T-t)}\right| \leq 1$$

since w, $\psi_{z,w}(t)$ and $\phi_{z,w}(t) = \int_0^t \psi_{z,w}(s)ds$ all have nonpositive real parts. Therefore the process $(M_t)_t$ is a true martingale indeed. We deduce that $F_{z,w}(T, y_0) = \mathbb{E}_{y_0}\left(e^{w\int_0^T Y_s ds} e^{zY_T}\right)$ and the assertion follows. □

We also have the following result which specifies the behaviour of the Laplace transform of $(Y_t, \int_0^t Y_s ds)$ when evaluated in two real numbers, not necessarily nonpositive.

Proposition 4.5. *Let λ_1 and λ_2 be two real numbers such that*

$$\frac{\sigma^2}{2}\lambda_1^2 - \kappa\lambda_1 + \lambda_2 \leq 0.$$

Then, the equation

$$\psi'(t) = \frac{\sigma^2}{2}\psi^2(t) - \kappa\psi(t) + \lambda_2 \tag{4.5}$$

has a unique solution $\psi_{\lambda_1,\lambda_2}$ defined on $[0, +\infty)$ such that $\psi_{\lambda_1,\lambda_2}(0) = \lambda_1$. Moreover, for every $t \geq 0$, we have

$$\mathbb{E}_{y_0}\left(e^{\lambda_1 Y_t + \lambda_2 \int_0^t Y_s ds}\right) \leq e^{y_0\psi_{\lambda_1,\lambda_2}(t) + \theta\kappa\phi_{\lambda_1,\lambda_2}(t)},$$

with $\phi_{\lambda_1,\lambda_2}(t) = \int_0^t \psi_{\lambda_1,\lambda_2}(s)ds$.

Proof. Let ψ be the solution of (4.5) with $\psi(0) = \lambda_1$. We have

$$\psi''(t) = (\sigma^2\psi(t) - \kappa)\psi'(t).$$

Therefore, the function $t \mapsto \psi'(t)e^{-\int_0^t(\sigma^2\psi(s)-\kappa)ds}$is a constant, hence $\psi'(t)$ has constant sign. Moreover, the assumption on λ_1 and λ_2 ensures that $\psi'(0) \leq 0$. We deduce that $\psi'(t) \leq 0$ and $\psi(t)$ remains between the solutions of the equation

$$\frac{\sigma^2}{2}\lambda^2 - \kappa\lambda + \lambda_2 = 0.$$

This proves that the solution is defined on the whole interval $[0, +\infty)$. Now the assertion follows as in the proof of Proposition 4.4: just note that the process $(M_t)_t$ defined as in (4.4) is no more uniformly bounded, so we cannot directly deduce that it is a martingale. However, it remains a positive local martingale, hence a supermartingale. □

Remark 4.6. *Let us now consider two real numbers λ_1 and λ_2 such that*

$$\frac{\sigma^2}{2}\lambda_1^2 - \kappa\lambda_1 + \lambda_2 < 0.$$

From the proof of Proposition 4.5, by using the optional sampling theorem we have

$$\sup_{\tau\in\mathcal{T}_{0,T}} \mathbb{E}_y\left(e^{\lambda_2\int_0^\tau Y_s ds} e^{\psi_{\lambda_1,\lambda_2}(T-\tau)Y_\tau + \theta\kappa\phi_{\lambda_1,\lambda_2}(T-\tau)}\right) \quad \leq \quad e^{y\psi_{\lambda_1,\lambda_2}(T) + \theta\kappa\phi_{\lambda_1,\lambda_2}(T)}.$$

Consider now $\epsilon > 0$ and let $\lambda_1^\epsilon = (1+\epsilon)\lambda_1$ and $\lambda_2^\epsilon = (1+\epsilon)\lambda_2$. For ϵ small enough, we have $\frac{\sigma^2}{2}(\lambda_1^\epsilon)^2 - \kappa\lambda_1^\epsilon + \lambda_2^\epsilon < 0$. Therefore

$$\sup_{\tau\in\mathcal{T}_{0,T}} \mathbb{E}_y\left(e^{\lambda_2^\epsilon\int_0^\tau Y_s ds} e^{\psi_{\lambda_1^\epsilon,\lambda_2^\epsilon}(T-\tau)Y_\tau + \theta\kappa\phi_{\lambda_1^\epsilon,\lambda_2^\epsilon}(T-\tau)}\right) \quad \leq \quad e^{y\psi_{\lambda_1^\epsilon,\lambda_2^\epsilon}(T) + \theta\kappa\phi_{\lambda_1^\epsilon,\lambda_2^\epsilon}(T)}.$$

If we have $\psi_{\lambda_1^\epsilon,\lambda_2^\epsilon} \geq (1+\epsilon)\psi_{\lambda_1,\lambda_2}$, we can deduce that

$$\sup_{\tau\in\mathcal{T}_{0,T}} \mathbb{E}_y\left(e^{\lambda_2(1+\epsilon)\int_0^\tau Y_s ds} e^{(1+\epsilon)\left(\psi_{\lambda_1,\lambda_2}(T-\tau)Y_\tau + \theta\kappa\phi_{\lambda_1,\lambda_2}(T-\tau)\right)}\right) \quad \leq \quad e^{y\psi_{\lambda_1^\epsilon,\lambda_2^\epsilon}(T) + \theta\kappa\phi_{\lambda_1^\epsilon,\lambda_2^\epsilon}(T)},$$

and, therefore, that the family $\left(e^{\lambda_2\int_0^\tau Y_s ds}e^{\psi_{\lambda_1,\lambda_2}(T-\tau)Y_\tau+\theta\kappa\phi_{\lambda_1,\lambda_2}(T-\tau)}\right)_{\tau\in\mathcal{T}_{0,T}}$ *is uniformly integrable. As a consequence, the process* $(M_t)_t$ *is a true martingale and we have*

$$\mathbb{E}_y\left(e^{\lambda_1 Y_t+\lambda_2\int_0^t Y_s ds}\right)=e^{y\psi_{\lambda_1,\lambda_2}(t)+\theta\kappa\phi_{\lambda_1,\lambda_2}(t)}.$$

So, it remains to show that $\psi_{\lambda_1^\epsilon,\lambda_2^\epsilon}\geq(1+\epsilon)\psi_{\lambda_1,\lambda_2}$*. In order to do this we set* $g_\epsilon(t)=\psi_{\lambda_1^\epsilon,\lambda_2^\epsilon}(t)-(1+\epsilon)\psi_{\lambda_1,\lambda_2}(t)$*. From the equations satisfied by* $\psi_{\lambda_1^\epsilon,\lambda_2^\epsilon}$ *and* $\psi_{\lambda_1,\lambda_2}$ *we deduce that*

$$\begin{aligned}
g_\epsilon'(t) &= \frac{\sigma^2}{2}\left(\psi^2_{\lambda_1^\epsilon,\lambda_2^\epsilon}(t)-(1+\epsilon)\psi^2_{\lambda_1,\lambda_2}(t)\right)-\kappa\left(\psi_{\lambda_1^\epsilon,\lambda_2^\epsilon}(t)-(1+\epsilon)\psi_{\lambda_1,\lambda_2}(t)\right)\\
&= \frac{\sigma^2}{2}\left(\psi^2_{\lambda_1^\epsilon,\lambda_2^\epsilon}(t)-(1+\epsilon)^2\psi^2_{\lambda_1,\lambda_2}(t)\right)-\kappa g_\epsilon(t)+\frac{\sigma^2}{2}\left((1+\epsilon)^2-(1+\epsilon)\right)\psi^2_{\lambda_1,\lambda_2}(t)\\
&= \frac{\sigma^2}{2}\left(\psi_{\lambda_1^\epsilon,\lambda_2^\epsilon}(t)+(1+\epsilon)\psi_{\lambda_1,\lambda_2}(t)\right)g_\epsilon(t)-\kappa g_\epsilon(t)+\frac{\sigma^2}{2}\epsilon(1+\epsilon)\psi^2_{\lambda_1,\lambda_2}(t)\\
&= f_\epsilon(t)g_\epsilon(t)+\frac{\sigma^2}{2}\epsilon(1+\epsilon)\psi^2_{\lambda_1,\lambda_2}(t),
\end{aligned}$$

where

$$f_\epsilon(t)=\frac{\sigma^2}{2}\left(\psi_{\lambda_1^\epsilon,\lambda_2^\epsilon}(t)+(1+\epsilon)\psi_{\lambda_1,\lambda_2}(t)\right)-\kappa.$$

Therefore, the function $g_\epsilon(t)e^{-\int_0^t f_\epsilon(s)ds}$ *is nondecreasing and, since* $g_\epsilon(0)=0$*, we have* $g_\epsilon(t)\geq 0$*.*

We can now prove the following Lemma, which will be useful in Section 4.4 to prove suitable estimates on the joint law of the process (X,Y).

Lemma 4.7. *For every* $q>0$ *there exists* $C>0$ *such that for all* $y_0\geq 0$*,*

$$\mathbb{E}_{y_0}\left(\int_0^t Y_v dv\right)^{-q}\leq\frac{C}{t^{2q}}. \tag{4.6}$$

Proof. If we take $\lambda_1=0$ and $\lambda_2=-s$ with $s>0$ in Proposition 4.5, we get

$$\mathbb{E}_{y_0}\left(e^{-s\int_0^t Y_v dv}\right)=e^{y_0\psi_{0,-s}(t)+\theta\kappa\phi_{0,-s}(t)}.$$

Since $\psi'_{0,-s}(0)=-s<0$, we can deduce by the proof of Proposition 4.5 that $\psi'_{0,-s}(t)=-se^{\int_0^t(\sigma^2\psi(u)-\kappa)du}$. Therefore, since $\psi_{0,-s}=0$, we have

$$\psi_{0,-s}(t)=-s\int_0^t e^{\int_0^u(\sigma^2\psi(v)-\kappa)dv}du. \tag{4.7}$$

Again from the proof of Proposition 4.5,

$$\psi_{0,-s}(t)\geq\frac{\kappa}{\sigma^2}-\sqrt{\left(\frac{\kappa}{\sigma^2}\right)^2+2\frac{s}{\sigma^2}}\geq-\sqrt{2s/\sigma^2},$$

so, by using (4.7), we deduce that

$$\psi_{0,-s}(t)\leq-s\int_0^t e^{\int_0^u-(\sigma\sqrt{2s}+\kappa)dv}du=-s\int_0^t e^{-\lambda_s u}du=-\frac{s}{\lambda_s}(1-e^{-t\lambda_s}).$$

where $\lambda_s=\sigma\sqrt{2s}+\kappa$. Since $\phi_{0,-s}(t)=\int_0^t\psi_{0,-s}(u)du$, we have

$$\phi_{0,-s}(t)\leq-\frac{s}{\lambda_s^2}\left(t\lambda_s-1+e^{-t\lambda_s}\right).$$

Therefore, since $\psi_{0,-s}(t) \leq 0$, for any $y_0 \geq 0$ we get

$$\mathbb{E}_{y_0}\left(e^{-s\int_0^t Y_v dv}\right) \leq e^{\kappa\theta\phi_{0,-s}(t)} \leq e^{-\frac{\kappa\theta s}{\lambda_s^2}(t\lambda_s - 1 + e^{-t\lambda_s})}.$$

Now, recall that for every $q > 0$ we can write

$$\frac{1}{y^q} = \frac{1}{\Gamma(q)}\int_0^\infty s^{q-1}e^{-sy}ds.$$

Therefore

$$\begin{aligned}\mathbb{E}_{y_0}\left(\int_0^t Y_v dv\right)^{-q} &= \mathbb{E}_{y_0}\left(\frac{1}{\Gamma(q)}\int_0^\infty s^{q-1}e^{-s\int_0^t Y_v dv}ds\right)\\ &\leq \frac{1}{\Gamma(q)}\int_0^1 s^{q-1}e^{-\frac{\kappa\theta s}{\lambda_s^2}(t\lambda_s - 1 + e^{-t\lambda_s})}ds + \frac{1}{\Gamma(q)}\int_1^\infty s^{q-1}e^{-\frac{\kappa\theta s}{\lambda_s^2}(t\lambda_s - 1 + e^{-t\lambda_s})}ds.\end{aligned}$$

Recall that $\lambda_s = \sigma\sqrt{2s} + \kappa$, so the first terms in the right hand side is finite. Moreover, for $s > 1$, we have $\frac{\kappa\theta s}{\lambda_s^2} \leq C$. Then, by noting that the function $u \mapsto tu - 1 + e^{-tu}$ is nondecreasing, we have

$$\begin{aligned}\mathbb{E}_{y_0}\left(\int_0^t Y_v dv\right)^{-q} &\leq C + \frac{1}{\Gamma(q)}\int_1^\infty s^{q-1}e^{-C(t\sigma\sqrt{2s} - 1 + e^{-t\sigma\sqrt{2s}})}ds\\ &\leq C + \frac{1}{t^{2q}\Gamma(q)}\int_0^\infty v^{q-1}e^{-C(\sigma\sqrt{2v} - 1 + e^{-\sigma\sqrt{2v}})}dv\\ &\leq \frac{C}{t^{2q}},\end{aligned}$$

which concludes the proof. □

Now recall that the diffusion (X, Y) evolves according to the following stochastic differential system

$$\begin{cases} dX_t = \left(\frac{\rho\kappa\theta}{\sigma} - \frac{Y_t}{2}\right)dt + \sqrt{Y_t}dB_t,\\ dY_t = \kappa(\theta - Y_t)dt + \sigma\sqrt{Y_t}dW_t.\end{cases}$$

If we set $\tilde{X}_t = X_t - \frac{\rho}{\sigma}Y_t$, we have

$$\begin{cases} d\tilde{X}_t = \left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right)Y_t dt + \sqrt{1-\rho^2}\sqrt{Y_t}d\tilde{B}_t,\\ dY_t = \kappa(\theta - Y_t)dt + \sigma\sqrt{Y_t}dW_t.\end{cases} \tag{4.8}$$

where $\tilde{B}_t = (1-\rho^2)^{-1/2}(B_t - \rho W_t)$. Note that $\tilde{B}$ is a standard Brownian motion with $\langle\tilde{B}, W\rangle_t = 0$.

Proposition 4.8. *For all $u, v \in \mathbb{R}$, for all $\lambda \geq 0$ and for all $(x_0, y_0) \in \mathbb{R} \times [0, +\infty)$ we have*

$$\mathbb{E}_{x_0,y_0}\left(e^{iuX_t + ivY_t}e^{-\lambda\int_0^t Y_s ds}\right) = e^{iux_0 + y_0(\psi_{\lambda_1,\mu}(t) - iu\frac{\rho}{\sigma}) + \theta\kappa\phi_{\lambda_1,\mu}(t)},$$

where $\lambda_1 = i(u\frac{\rho}{\sigma} + v)$, $\mu = iu\left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right) - \frac{u^2}{2}(1-\rho^2) - \lambda$ and the function $\psi_{\lambda_1,\mu}$ and $\phi_{\lambda_1,\mu}$ are defined in Proposition 4.4.

Proof. We have

$$\mathbb{E}_{x_0,y_0}\left(e^{iuX_t + ivY_t - \lambda\int_0^t Y_s ds}\right) = \mathbb{E}_{x_0,y_0}\left(e^{iu(\tilde{X}_t + \frac{\rho}{\sigma}Y_t) + ivY_t - \lambda\int_0^t Y_s ds}\right)$$

and

$$\tilde{X}_t = x_0 - \frac{\rho}{\sigma}y_0 + \int_0^t\left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right)Y_s ds + \int_0^t\sqrt{(1-\rho^2)Y_s}d\tilde{B}_s.$$

Since $\tilde{B}$ and W are independent,

$$\mathbb{E}\left(e^{iu\tilde{X}_t} \mid W\right) = e^{iu\left(x_0-\frac{\rho}{\sigma}y_0+\int_0^t\left(\frac{\rho\kappa}{\sigma}-\frac{1}{2}\right)Y_s ds\right)-\frac{u^2}{2}(1-\rho^2)\int_0^t Y_s ds}$$

and

$$\mathbb{E}_{x_0,y_0}\left(e^{iuX_t+ivY_t-\lambda\int_0^t Y_s ds}\right) = e^{iu\left(x_0-\frac{\rho}{\sigma}y_0\right)}\mathbb{E}_{y_0}\left(e^{i\left(u\frac{\rho}{\sigma}+v\right)Y_t+\left(iu(\frac{\rho\kappa}{\sigma}-\frac{1}{2})-\frac{u^2}{2}(1-\rho^2)-\lambda\right)\int_0^t Y_s ds}\right).$$

Then the assertion follows by using Proposition 4.4. □

4.3 Identification of the semigroups

We now show that the semigroup $\bar{P}_t^\lambda$ associated with the coercive bilinear form can be actually identified with the transition semigroup P_t^λ. Recall the Sobolev spaces $L^p(\mathcal{O},\mathfrak{m}_{\gamma,\mu})$ introduced in Definition 2.1 for $p \geq 1$. In order to prove the identification of the semigroups, we need the following property of the transition semigroup.

Theorem 4.9. *For all $p > 1$, $\gamma > 0$ and $\mu > 0$ there exists $\lambda > 0$ such that, for every compact $K \subseteq \mathbb{R}\times[0,+\infty)$ and for every $T > 0$, there is $C_{p,K,T} > 0$ such that*

$$P_t^\lambda f(x_0,y_0) \leq \frac{C_{p,K,T}}{t^{\frac{\beta}{p}+\frac{3}{2p}}}||f||_{L^p(\mathcal{O},\mathfrak{m}_{\gamma,\mu})}, \qquad (x_0,y_0)\in K.$$

for every measurable positive function f on $\mathbb{R}\times[0,+\infty)$ and for every $t \in (0,T]$.

Theorem 4.9 will also play a crucial role in order to prove Theorem 2.4. Its proof relies on suitable estimates on the joint law of the diffusion (X,Y) and we postpone it to the following section. Then, we can prove the following result.

Proposition 4.10. *There exists $\lambda > 0$ such that, for every function $f \in H$ and for every $t \geq 0$,*

$$\bar{P}_t^\lambda f(x,y) = P_t^\lambda f(x,y), \quad dxdy \ a.e.$$

Proof. We can easily deduce from Theorem 4.9 with $p = 2$ that, for λ large enough, if $(f_n)_n$ is a sequence of functions which converges to f in H, then the sequence $(P_t^\lambda f_n)_n$ converges uniformly to $P_t^\lambda f$ on the compact sets. On the other hand, recall that $\bar{P}_t^\lambda$ is a contraction semigroup on H so that the function $f \mapsto \bar{P}_t^\lambda f$ is continuous and we have $\bar{P}_t^\lambda f_n \to \bar{P}_t^\lambda f$ in H.

Therefore, by density arguments, it is enough to prove the equality for $f(x,y) = e^{iux+ivy}$ with u, $v \in \mathbb{R}$. We have, by using Proposition 4.8,

$$\begin{aligned} P_t^\lambda f(x,y) &= \mathbb{E}_{x,y}\left(e^{-\lambda\int_0^t(1+Y_s)ds}e^{iuX_t+ivY_t}\right) \\ &= e^{-\lambda t}e^{iux+y\left(\psi_{\lambda_1,\mu}(t)-iu\frac{\rho}{\sigma}\right)+\theta\kappa\phi_{\lambda_1,\mu}(t)}, \end{aligned}$$

with $\lambda_1 = i(u\frac{\rho}{\sigma}+v)$, $\mu = iu\left(\frac{\rho\kappa}{\sigma}-\frac{1}{2}\right)-\frac{u^2}{2}(1-\rho^2)-\lambda$. The function $F(t,x,y)$ defined by $F(t,x,y) = e^{-\lambda t}e^{iux+y\left(\psi_{\lambda_1,\mu}(t)-iu\frac{\rho}{\sigma}\right)+\theta\kappa\phi_{\lambda_1,\mu}(t)}$ satisfies $F(0,x,y) = e^{iux+ivy}$ and

$$\frac{\partial F}{\partial t} = (\mathcal{L}-\lambda(1+y))\,F.$$

Moreover, since the real parts of λ_1 and μ are nonnegative, we can deduce from the proof of Proposition 4.4 that the real part of the function $t \to \psi(t)$ is nonnegative. Then, it is straightforward to see that, for every $t \geq 0$, we have $F(t,\cdot,\cdot) \in H^2(\mathcal{O},\mathfrak{m})$ and $t \mapsto F(t,\cdot,\cdot)$ is continuous, so that, for every $v \in V$, $(\mathcal{L}F(t,.,.),v)_H = -a(F(t,.,.),v)$. Therefore

$$\left(\frac{\partial F}{\partial t},v\right)_H + a_\lambda(F(t,.,.),v) = 0 \quad v \in V,$$

and $F(t,.,.) = \bar{P}_t^\lambda f$. □

4.4 Estimates on the joint law

In this section we prove Theorem 4.9. We first recall some results about the density of the process Y.

With the notations

$$\nu = \beta - 1 = \frac{2\kappa\theta}{\sigma^2} - 1, \quad y_t = y_0 e^{-\kappa t}, \quad L_t = \frac{\sigma^2}{4\kappa}\left(1 - e^{-\kappa t}\right),$$

it is well known (see, for example, [15, Section 6.2.2]) that the transition density of the process Y is given by

$$p_t(y_0, y) = \frac{e^{-\frac{y_t}{2L_t}}}{2y_t^{\nu/2} L_t} e^{-\frac{y}{2L_t}} y^{\nu/2} I_\nu\left(\frac{\sqrt{yy_t}}{L_t}\right),$$

where I_ν is the first-order modified Bessel function with index ν, defined by

$$I_\nu(y) = \left(\frac{y}{2}\right)^\nu \sum_{n=0}^{\infty} \frac{(y/2)^{2n}}{n!\Gamma(n+\nu+1)}.$$

It is clear that near $y = 0$ we have $I_\nu(y) \sim \frac{1}{\Gamma(\nu+1)}\left(\frac{y}{2}\right)^\nu$ while, for $y \to \infty$, we have the asymptotic behaviour $I_\nu(y) \sim e^y/\sqrt{2\pi y}$ (see [1, page 377]).

Proposition 4.11. *There exists a constant $C_\beta > 0$ (which depends only on β) such that, for every $t > 0$,*

$$p_t(y_0, y) \le \frac{C_\beta}{L_t^{\beta+\frac{1}{2}}} e^{-\frac{(\sqrt{y}-\sqrt{y_t})^2}{2L_t}} y^{\beta-1}\left(L_t^{1/2} + (yy_t)^{1/4}\right), \qquad (y_0, y) \in [0, +\infty)\times]0, +\infty).$$

Proof. From the asymptotic behaviour of I_ν near 0 and ∞ we deduce the existence of a constant $C_\nu > 0$ such that

$$I_\nu(x) \le C_\nu\left(x^\nu \mathbb{1}_{\{x\le 1\}} + \frac{e^x}{\sqrt{x}}\mathbb{1}_{\{x>1\}}\right).$$

Therefore

$$\begin{aligned}
p_t(y_0, y) &= \frac{e^{-\frac{y_t+y}{2L_t}}}{2y_t^{\nu/2}L_t} y^{\nu/2} I_\nu\left(\frac{\sqrt{yy_t}}{L_t}\right) \\
&\le \frac{e^{-\frac{y_t+y}{2L_t}}}{2y_t^{\nu/2}L_t} y^{\nu/2} C_\nu\left(\frac{(yy_t)^{\nu/2}}{L_t^\nu}\mathbb{1}_{\{yy_t\le L_t^2\}} + \frac{e^{\frac{\sqrt{yy_t}}{L_t}}}{(yy_t)^{1/4}/L_t^{1/2}}\mathbb{1}_{\{yy_t>L_t^2\}}\right) \\
&= \frac{C_\nu}{2} e^{-\frac{y_t+y}{2L_t}}\left(\frac{y^\nu}{L_t^{\nu+1}}\mathbb{1}_{\{yy_t\le L_t^2\}} + \frac{y^{\frac{\nu}{2}-\frac{1}{4}} e^{\frac{\sqrt{yy_t}}{L_t}}}{(y_t)^{\frac{\nu}{2}+\frac{1}{4}} L_t^{1/2}}\mathbb{1}_{\{yy_t>L_t^2\}}\right).
\end{aligned}$$

On $\{yy_t > L_t^2\}$, we have $y_t^{-1} \le y/L_t^2$ and, since $\nu + 1 > 0$,

$$\frac{y^{\frac{\nu}{2}-\frac{1}{4}}}{(y_t)^{\frac{\nu}{2}+\frac{1}{4}}} = y_t^{1/4}\frac{y^{\frac{\nu}{2}-\frac{1}{4}}}{(y_t)^{\frac{\nu}{2}+\frac{1}{2}}} \le y_t^{1/4}\frac{y^{\nu+\frac{1}{4}}}{L_t^{\nu+1}}.$$

So

$$\begin{aligned}
p_t(y_0, y) &\le \frac{C_\nu}{2} e^{-\frac{y_t+y}{2L_t}}\left(\frac{y^\nu}{L_t^{\nu+1}}\mathbb{1}_{\{yy_t\le L_t^2\}} + \frac{(yy_t)^{1/4} y^\nu e^{\frac{\sqrt{yy_t}}{L_t}}}{L_t^{\nu+\frac{3}{2}}}\mathbb{1}_{\{yy_t>L_t^2\}}\right) \\
&\le \frac{C_\nu}{2L_t^{\nu+\frac{3}{2}}} e^{-\frac{y_t+y}{2L_t}} y^\nu e^{\frac{\sqrt{yy_t}}{L_t}}\left(L_t^{1/2}\mathbb{1}_{\{yy_t\le L_t^2\}} + (yy_t)^{1/4}\mathbb{1}_{\{yy_t>L_t^2\}}\right) \\
&= \frac{C_\nu}{2L_t^{\nu+\frac{3}{2}}} e^{-\frac{(\sqrt{y}-\sqrt{y_t})^2}{2L_t}} y^\nu\left(L_t^{1/2}\mathbb{1}_{\{yy_t\le L_t^2\}} + (yy_t)^{1/4}\mathbb{1}_{\{yy_t>L_t^2\}}\right),
\end{aligned}$$

and the assertion follows. □

We are now ready to prove Theorem 4.9, which we have used in order to prove the identification of the semigroups in Proposition 4.10 and which we will use again later on in this paper.

Proof of Theorem 4.9. Note that

$$P_t^\lambda f(x_0,y_0) \quad = \quad \mathbb{E}_{x_0,y_0}\left(e^{-\lambda\int_0^t(1+Y_s)ds}\tilde{f}(\tilde{X}_t,Y_t)\right),$$

where

$$\tilde{f}(x,y)=f\left(x+\frac{\rho}{\sigma}y,y\right) \quad \text{and} \quad \tilde{X}_t=X_t-\frac{\rho}{\sigma}Y_t.$$

Recall that the dynamics of $\tilde{X}$ is given by (4.8) so we have

$$\tilde{X}_t=\tilde{x}_0+\bar{\kappa}\int_0^t Y_s ds+\bar{\rho}\int_0^t\sqrt{Y_s}d\tilde{B}_s,$$

with

$$\tilde{x}_0=x_0-\frac{\rho}{\sigma}y_0, \quad \bar{\kappa}=\frac{\rho\kappa}{\sigma}-\frac{1}{2}, \quad \bar{\rho}=\sqrt{1-\rho^2}.$$

Recall that the Brownian motion $\tilde{B}$ is independent of the process Y. We set $\Sigma_t=\sqrt{\int_0^t Y_s ds}$ and $n(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$. Therefore

$$\begin{aligned}
P_t^\lambda f(x_0,y_0) &= \mathbb{E}_{y_0}\left(e^{-\lambda t-\lambda\Sigma_t^2}\int\tilde{f}\left(\tilde{x}_0+\bar{\kappa}\Sigma_t^2+\bar{\rho}\Sigma_t z,Y_t\right)n(z)dz\right)\\
&\leq \mathbb{E}_{y_0}\left(e^{-\lambda\Sigma_t^2}\int\tilde{f}\left(\tilde{x}_0+\bar{\kappa}\Sigma_t^2+\bar{\rho}\Sigma_t z,Y_t\right)n(z)dz\right)\\
&= \mathbb{E}_{y_0}\left(e^{-\lambda\Sigma_t^2}\int\tilde{f}\left(\tilde{x}_0+z,Y_t\right)n\left(\frac{z-\bar{\kappa}\Sigma_t^2}{\bar{\rho}\Sigma_t}\right)\frac{dz}{\bar{\rho}\Sigma_t}\right).
\end{aligned}$$

Hölder's inequality with respect to the measure $e^{-\gamma|z|-\bar{\mu}Y_t}dzd\mathbb{P}_{y_0}$, where $\gamma>0$ and $\bar{\mu}>0$ will be chosen later on, gives, for every $p>1$

$$P_t^\lambda f(x_0,y_0) \quad \leq \quad \left[\mathbb{E}_{y_0}\left(\int e^{-\gamma|z|-\bar{\mu}Y_t}\tilde{f}^p\left(\tilde{x}_0+z,Y_t\right)dz\right)\right]^{1/p}J_q, \tag{4.9}$$

with $q=p/(p-1)$ and

$$(J_q)^q=\mathbb{E}_{y_0}\left(\int e^{(q-1)\gamma|z|+(q-1)\bar{\mu}Y_t-q\lambda\Sigma_t^2}n^q\left(\frac{z-\bar{\kappa}\Sigma_t^2}{\bar{\rho}\Sigma_t}\right)\frac{dz}{(\bar{\rho}\Sigma_t)^q}\right).$$

Using Proposition 4.11 we can write, for every $z\in\mathbb{R}$,

$$\begin{aligned}
\mathbb{E}_{y_0}\left(e^{-\bar{\mu}Y_t}\tilde{f}^p\left(\tilde{x}_0+z,Y_t\right)\right) &= \int_0^\infty dy p_t(y_0,y)e^{-\bar{\mu}y}\tilde{f}^p\left(\tilde{x}_0+z,y\right)\\
&\leq \frac{C_\beta\left(\sqrt{\frac{\sigma^2}{4\kappa}}+y_0^{1/4}\right)}{L_t^{\beta+\frac{1}{2}}}\int_0^\infty dy e^{-\frac{(\sqrt{y}-\sqrt{y_t})^2}{2L_t}-\bar{\mu}y}y^{\beta-1}\left(1+y^{1/4}\right)\tilde{f}^p\left(\tilde{x}_0+z,y\right).
\end{aligned}$$

If we set $L_\infty=\sigma^2/(4\kappa)$, for every $\epsilon\in(0,1)$ we have

$$\begin{aligned}
e^{-\frac{(\sqrt{y}-\sqrt{y_t})^2}{2L_t}} &\leq e^{-\frac{(\sqrt{y}-\sqrt{y_t})^2}{2L_\infty}}\\
&= e^{-\frac{y}{2L_\infty}}e^{\frac{\sqrt{yy_t}}{L_\infty}-\frac{y_t}{2L_\infty}}\\
&\leq e^{-\frac{y}{2L_\infty}}e^{\epsilon\frac{y}{2L_\infty}}e^{\frac{y_t}{2\epsilon L_\infty}}e^{-\frac{y_t}{2L_\infty}}\\
&= e^{-(1-\epsilon)\frac{y}{2L_\infty}}e^{\frac{y_t}{2\epsilon L_\infty}(1-\epsilon)}\\
&\leq e^{-(1-\epsilon)\frac{y}{2L_\infty}}e^{\frac{y_0}{2\epsilon L_\infty}(1-\epsilon)}.
\end{aligned}$$

It is easy to see that $e^{-y\left(\bar{\mu}+\frac{1-\epsilon}{2L_\infty}\right)}(1+y^{1/4}) \leq C_{\epsilon,\sigma,\kappa}e^{-y\left(\bar{\mu}+\frac{1-2\epsilon}{2L_\infty}\right)}$. Therefore, we can write

$$\begin{aligned}\mathbb{E}_{y_0}\left(e^{-\bar{\mu}Y_t}\tilde{f}^p\left(\tilde{x}_0+z,Y_t\right)\right) &\leq \frac{C_\beta e^{\frac{y_0(1-\epsilon)}{2\epsilon L_\infty}}\left(\sqrt{\frac{\sigma^2}{4\kappa}}+y_0^{1/4}\right)}{L_t^{\beta+\frac{1}{2}}}\int_0^\infty dy e^{-y\left(\bar{\mu}+\frac{1-\epsilon}{2L_\infty}\right)}y^{\beta-1}\left(1+y^{1/4}\right)\tilde{f}^p\left(\tilde{x}_0+z,y\right)\\ &\leq \frac{C_{\beta,\sigma,\kappa,\epsilon}e^{\frac{y_0(1-\epsilon)}{\epsilon L_\infty}}}{L_t^{\beta+\frac{1}{2}}}\int_0^\infty dy e^{-y\left(\bar{\mu}+\frac{1-2\epsilon}{2L_\infty}\right)}y^{\beta-1}\tilde{f}^p\left(\tilde{x}_0+z,y\right).\end{aligned}$$

As regards J_q, setting $z'=\frac{z-\bar{\kappa}\Sigma_t^2}{\bar{\rho}\Sigma_t}$, we have

$$\begin{aligned}(J_q)^q &= \mathbb{E}_{y_0}\left(\int e^{(q-1)\gamma|z'\bar{\rho}\Sigma_t+\bar{\kappa}\Sigma_t^2|+(q-1)\bar{\mu}Y_t-q\lambda\Sigma_t^2}n^q\left(z'\right)\frac{dz'}{(\bar{\rho}\Sigma_t)^{q-1}}\right)\\ &\leq \mathbb{E}_{y_0}\left(\int e^{(q-1)\gamma\bar{\rho}\Sigma_t|z|+(q-1)\bar{\mu}Y_t+((q-1)|\bar{\kappa}|\gamma-q\lambda)\Sigma_t^2}n^q\left(z\right)\frac{dz}{(\bar{\rho}\Sigma_t)^{q-1}}\right).\end{aligned}$$

Note that

$$\begin{aligned}\int e^{(q-1)\gamma\bar{\rho}\Sigma_t|z|}n^q\left(z\right)dz &= \frac{1}{(\sqrt{2\pi})^q}\int e^{(q-1)\gamma\bar{\rho}\Sigma_t|z|}e^{-qz^2/2}dz\\ &\leq \frac{2}{\sqrt{2\pi}}\int e^{(q-1)\gamma\bar{\rho}\Sigma_t z}e^{-qz^2/2}dz\\ &= \frac{2}{\sqrt{2\pi}}e^{\frac{(q-1)^2}{2q}\gamma^2\bar{\rho}^2\Sigma_t^2}\int e^{-\frac{1}{2}\left(\sqrt{q}z-\frac{(q-1)\gamma\bar{\rho}\Sigma_t}{\sqrt{q}}\right)^2}dz\\ &= \frac{2}{\sqrt{q}}e^{\frac{(q-1)^2}{2q}\gamma^2\bar{\rho}^2\Sigma_t^2},\end{aligned}$$

so that

$$(J_q)^q \leq \frac{2}{\sqrt{q}}\mathbb{E}_{y_0}\left(e^{(q-1)\bar{\mu}Y_t+\bar{\lambda}_q\Sigma_t^2}\frac{1}{(\bar{\rho}\Sigma_t)^{q-1}}\right),$$

with

$$\bar{\lambda}_q=(q-1)|\bar{\kappa}|\gamma+\frac{(q-1)^2}{2q}\gamma^2\bar{\rho}^2-q\lambda=\frac{1}{p-1}\left(|\bar{\kappa}|\gamma+\frac{1}{2p}\gamma^2\bar{\rho}^2-p\lambda\right).$$

Using Hölder's inequality again we get, for every $p_1>1$ and $q_1=p_1/(p_1-1)$,

$$\begin{aligned}(J_q)^q &\leq \sqrt{\frac{2}{q}}\left(\mathbb{E}_{y_0}\left(e^{p_1(q-1)\bar{\mu}Y_t+p_1\bar{\lambda}_q\Sigma_t^2}\right)\right)^{1/p_1}\left(\mathbb{E}_{y_0}\left(\frac{1}{(\bar{\rho}\Sigma_t)^{q_1(q-1)}}\right)\right)^{1/q_1}\\ &\leq \frac{C_{q,q_1}}{t^{q-1}}\left(\mathbb{E}_{y_0}\left(e^{p_1(q-1)\bar{\mu}Y_t+p_1\bar{\lambda}_q\Sigma_t^2}\right)\right)^{1/p_1},\end{aligned}$$

where the last inequality follows from Lemma 4.7.

We now apply Proposition 4.5 with $\lambda_1=p_1(q-1)\bar{\mu}$ and $\lambda_2=p_1\bar{\lambda}_q$. The assumption on λ_1 and λ_2 becomes

$$\frac{\sigma^2}{2}p_1(q-1)\bar{\mu}^2-\kappa\bar{\mu}+|\bar{\kappa}|\gamma+\frac{1}{2p}\gamma^2\bar{\rho}^2-p\lambda\leq 0$$

or, equivalently,

$$\lambda\geq\frac{\sigma^2}{2p(p-1)}p_1\bar{\mu}^2-\kappa\frac{\bar{\mu}}{p}+|\bar{\kappa}|\frac{\gamma}{p}+\frac{1}{2p^2}\gamma^2\bar{\rho}^2.$$

Note that the last inequality is satisfied for at least a $p_1>1$ if and only if

$$\lambda>\frac{\sigma^2}{2p(p-1)}\bar{\mu}^2-\kappa\frac{\bar{\mu}}{p}+|\bar{\kappa}|\frac{\gamma}{p}+\frac{1}{2p^2}\gamma^2\bar{\rho}^2. \tag{4.10}$$

Going back to (4.9) under the condition (4.10), we have

$$\begin{aligned}
P_t^\lambda f(x_0,y_0) &\leq \frac{C_{p,\epsilon}}{L_t^{\frac{\beta}{p}+\frac{1}{2p}}t^{1/p}}e^{A_{p,\epsilon}y_0}\left(\int dz e^{-\gamma|z|}\int_0^\infty dy e^{-y\left(\bar\mu+\frac{1-2\epsilon}{2L_\infty}\right)}y^{\beta-1}\tilde f^p\left(\tilde x_0+z,y\right)\right)^{1/p}\\
&\leq \frac{C_{p,\epsilon}e^{A_{p,\epsilon}y_0}}{t^{\frac{\beta}{p}+\frac{3}{2p}}}\left(\int dz e^{-\gamma|z|}\int_0^\infty dy e^{-y\left(\bar\mu+\frac{1-2\epsilon}{2L_\infty}\right)}y^{\beta-1} f^p\left(\tilde x_0+z+\frac{\rho}{\sigma}y,y\right)\right)^{1/p}\\
&= \frac{C_{p,\epsilon}e^{A_{p,\epsilon}y_0}}{t^{\frac{\beta}{p}+\frac{3}{2p}}}\left(\int dz e^{-\gamma|z-\tilde x_0-\frac{\rho}{\sigma}y|}\int_0^\infty dy e^{-y\left(\bar\mu+\frac{1-2\epsilon}{2L_\infty}\right)}y^{\beta-1} f^p\left(z,y\right)\right)^{1/p}\\
&\leq \frac{C_{p,\epsilon}e^{A_{p,\epsilon}y_0+\gamma|\tilde x_0|}}{t^{\frac{\beta}{p}+\frac{3}{2p}}}\left(\int dz e^{-\gamma|z|}\int_0^\infty dy e^{-y\left(\bar\mu-\gamma\frac{|\rho|}{\sigma}+\frac{1-2\epsilon}{2L_\infty}\right)}y^{\beta-1} f^p\left(z,y\right)\right)^{1/p}.
\end{aligned}$$

If we choose $\epsilon = 1/2$ and $\bar\mu = \mu + \gamma\frac{|\rho|}{\sigma}$, the assertion follows provided λ satisfies

$$\lambda > \frac{\sigma^2}{2p(p-1)}\left(\mu+\gamma\frac{|\rho|}{\sigma}\right)^2 - \kappa\frac{\mu+\gamma\frac{|\rho|}{\sigma}}{p} + |\bar\kappa|\frac{\gamma}{p} + \frac{1}{p^2}\gamma^2\bar\rho^2.$$

□

4.5 Proof of Theorem 2.4

We are finally ready to prove the identification Theorem 2.4. We first prove the result under further regularity assumptions on the payoff function ψ, then we deduce the general statement by an approximation technique.

4.5.1 Case with a regular function ψ

The following regularity result paves the way for the identification theorem in the case of a regular payoff function.

Proposition 4.12. *Assume that ψ satisfies Assumption $\mathcal{H}^1$ and $0 \leq \psi \leq \Phi$ with Φ satisfying Assumption $\mathcal{H}^2$. If moreover we assume $\psi \in L^2([0,T];H^2(\mathcal{O},\mathfrak{m}))$ and $\frac{\partial\psi}{\partial t} + \mathcal{L}\psi$, $(1+y)\Phi \in L^p([0,T];L^p(\mathcal{O},\mathfrak{m}))$ for some $p \geq 2$, then there exist $\lambda_0 > 0$ and $F \in L^p([0,T];L^p(\mathcal{O},\mathfrak{m}))$ such that for all $\lambda \geq \lambda_0$ the solution u of (2.5) satisfies*

$$-\left(\frac{\partial u}{\partial t}, v\right)_H + a_\lambda(u,v) = (F,v)_H, \qquad \text{a.e. in } [0,T], \quad v \in V. \tag{4.11}$$

Proof. Note that, for λ large enough, u can be seen as the solution u_λ of an equivalent coercive variational inequality, that is

$$-\left(\frac{\partial u_\lambda}{\partial t}, v-u_\lambda\right)_H + a_\lambda(u_\lambda, v-u_\lambda) \geq (g, v-u_\lambda)_H,$$

where $g = \lambda(1+y)u$ satisfies the assumptions of Proposition 3.12. Therefore, there exists a sequence $(u_{\varepsilon,\lambda})_\varepsilon$ of non negative functions such that $\lim_{\varepsilon\to 0} u_{\varepsilon,\lambda} = u_\lambda$ and

$$-\left(\frac{\partial u_{\varepsilon,\lambda}}{\partial t}, v\right)_H + a_\lambda(u_{\varepsilon,\lambda}, v) - \left(\frac{1}{\varepsilon}(\psi - u_{\varepsilon,\lambda})_+, v\right)_H = (g,v)_H, \qquad v \in V.$$

Since both $u_{\varepsilon,\lambda}$ and ψ are positive and ψ belongs to $L^p([0,T];L^p(\mathcal{O},\mathfrak{m}))$, we have $(\psi - u_{\varepsilon,\lambda})_+ \in L^p([0,T];L^p(\mathcal{O},\mathfrak{m}))$. In order to simplify the notation, we set $w = (\psi - u_{\varepsilon,\lambda})_+$. Taking $v = w^{p-1}$ and assuming that ψ is bounded we observe that $v \in L^2([0,T];V)$ and we can write

$$-\left(\frac{\partial u_{\varepsilon,\lambda}}{\partial t}, w^{p-1}\right)_H + a_\lambda(u_{\varepsilon,\lambda}, w^{p-1}) - \frac{1}{\varepsilon}\|w\|^p_{L^p(\mathcal{O},\mathfrak{m})} = \left(g, w^{p-1}\right)_H,$$

so that

$$\frac{1}{p}\frac{d}{dt}\|w\|^p_{L^p(\mathcal{O},\mathfrak{m})} - a_\lambda(\psi - u_{\varepsilon,\lambda}, w^{p-1}) - \frac{1}{\varepsilon}\|w\|^p_{L^p(\mathcal{O},\mathfrak{m})} = \left(g, w^{p-1}\right)_H - \left(\frac{\partial\psi}{\partial t}, w^{p-1}\right)_H + a_\lambda(\psi, w^{p-1}).$$

Integrating from 0 to T we get

$$\begin{aligned} &-\frac{1}{p}\|w(0)\|^p_{L^p(\mathcal{O},\mathfrak{m})} - \int_0^T a_\lambda((\psi - u_{\varepsilon,\lambda})(t), w^{p-1}(t))dt - \frac{1}{\varepsilon}\int_0^T \|w(t)\|^p_{L^p(\mathcal{O},\mathfrak{m})}dt \\ &= \int_0^T \left(g(t), w^{p-1}(t)\right)_H dt - \int_0^T \left(\frac{\partial\psi}{\partial t}(t), w_+^{p-1}(t)\right)_H dt + \int_0^T a_\lambda(\psi(t), w^{p-1}(t))dt. \end{aligned} \tag{4.12}$$

Now, with the usual integration by parts,

$$\begin{aligned} a_\lambda(w, w^{p-1}) &= \int_{\mathcal{O}} \frac{y}{2}(p-1)w^{p-2}\left[\left(\frac{\partial w}{\partial x}\right)^2 + 2\rho\sigma\frac{\partial w}{\partial x}\frac{\partial w}{\partial y} + \sigma^2\left(\frac{\partial w}{\partial y}\right)^2\right]d\mathfrak{m} \\ &\quad + \int_{\mathcal{O}} y\left(j_{\gamma,\mu}(x)\frac{\partial w}{\partial x} + k_{\gamma,\mu}(x)\frac{\partial w}{\partial y}\right)w^{p-1}d\mathfrak{m} + \lambda\int_{\mathcal{O}}(1+y)w^p d\mathfrak{m} \\ &\geq \delta_1(p-1)\int_{\mathcal{O}} yw^{p-2}\left[\left(\frac{\partial w}{\partial x}\right)^2 + \left(\frac{\partial w}{\partial y}\right)^2\right]d\mathfrak{m} + \int_{\mathcal{O}} y\left(j_{\gamma,\mu}(x)\frac{\partial w}{\partial x} + k_{\gamma,\mu}(x)\frac{\partial w}{\partial y}\right)w^{p-1}d\mathfrak{m} + \lambda\int_{\mathcal{O}} yw^p d\mathfrak{m} \\ &= \int_{\mathcal{O}} yw^{p-2}\left[\delta_1(p-1)\left(\frac{\partial w}{\partial x}\right)^2 + j_{\gamma,\mu}(x)\frac{\partial w}{\partial x}w + \frac{\lambda}{2}w^2\right]d\mathfrak{m} \\ &\quad + \int_{\mathcal{O}} yw^{p-2}\left[\delta_1(p-1)\left(\frac{\partial w}{\partial y}\right)^2 + k_{\gamma,\mu}(x)\frac{\partial w}{\partial y}w + \frac{\lambda}{2}w^2\right]d\mathfrak{m} \geq 0, \end{aligned}$$

since, for λ large enough, the quadratic forms $(a,b) \to \delta_1(p-1)a^2 + j_{\gamma.\mu}ab + \frac{\lambda}{2}b^2$ and $(a,b) \to \delta_1(p-1)a^2 + k_{\gamma.\mu}ab + \frac{\lambda}{2}b^2$ are both positive definite.

Recall that $\psi \in L^2([0,T]; H^2(\mathcal{O},\mathfrak{m}))$, $\frac{\partial\psi}{\partial_t} + \mathcal{L}\psi \in L^p([0,T], L^p(\mathcal{O},\mathfrak{m}))$, $(1+y)\psi \leq (1+y)\Phi \in L^p([0,T], L^p(\mathcal{O},\mathfrak{m}))$ and $g = (1+y)u \leq (1+y)\Phi \in L^p([0,T]; L^p(\mathcal{O},\mathfrak{m}))$. Therefore, going back to (4.12) and using Holder's inequality,

$$\frac{1}{\varepsilon}\int_0^T \|w(t)\|^p_{L^p(\mathcal{O},\mathfrak{m})}dt \leq \left[\left(\int_0^T \|g(t)\|^p_{L^p(\mathcal{O},\mathfrak{m})}dt\right)^{\frac{1}{p}} + \left(\int_0^T \left\|\frac{\partial\psi}{\partial t}(t) + \mathcal{L}^\lambda\psi(t)\right\|^p_{L^p(\mathcal{O},\mathfrak{m})}dt\right)^{\frac{1}{p}}\right]\left(\int_0^T \|w\|^p_{L^p(\mathcal{O},\mathfrak{m})}dt\right)^{\frac{p-1}{p}}.$$

Recalling that $w = (\psi - u_{\varepsilon,\lambda})_+$, we deduce that

$$\left\|\frac{1}{\varepsilon}(\psi - u_{\varepsilon,\lambda})_+\right\|_{L^p([0,T];L^p(\mathcal{O},\mathfrak{m}))} \leq C, \tag{4.13}$$

for a positive constant C independent of ε. Note that the estimate does not involve the L^∞-norm of ψ (which we assumed to be bounded for the payoff) so that by a standard approximation argument, it remains valid for unbounded ψ. The assertion then follows passing to the limit for $\varepsilon \to 0$ in

$$-\left(\frac{\partial u_{\varepsilon,\lambda}}{\partial t}, v\right)_H + a_\lambda(u_{\varepsilon,\lambda}, v) = \left(\frac{1}{\varepsilon}(\psi - u_{\varepsilon,\lambda})_+, v\right)_H + (g,v)_H, \qquad v \in V.$$

□

Now, note that we can easily prove the continuous dependence of the process X with respect to the initial state.

Lemma 4.13. *Fix $(x,y) \in \mathbb{R} \times [0,+\infty)$. Denote by $(X_t^{x,y}, Y_t^y)_{t\geq 0}$ the solution of the system*

$$\begin{cases} dX_t = \left(\frac{\rho\kappa\theta}{\sigma} - \frac{Y_t}{2}\right)dt + \sqrt{Y_t}dB_t, \\ dY_t = \kappa(\theta - Y_t)dt + \sigma\sqrt{Y_t}dW_t, \end{cases}$$

with $X_0 = x$, $Y_0 = y$ and $\langle B, W\rangle_t = \rho t$. We have, for every $t \geq 0$ and for every (x,y), $(x',y') \in \mathbb{R} \times [0,+\infty)$, $\mathbb{E}\left|Y_t^{y'} - Y_t^y\right| \leq |y'-y|$ and

$$\mathbb{E}\left|X_t^{x',y'} - X_t^{x,y}\right| \leq |x'-x| + \frac{t}{2}|y'-y| + \sqrt{t|y'-y|}.$$

The proof of Lemma 4.13 is straightforward so we omit the details: the inequality $\mathbb{E}\left|Y_t^{y'} - Y_t^y\right| \leq |y'-y|$ can be proved by using standard techniques introduced in [12] (see the proof of Theorem 3.2 and its Corollary in Section IV.3) and the other inequality easily follows.

Then, we can prove the following result.

Proposition 4.14. *Let $\psi : \mathbb{R} \times [0,\infty) \to \mathbb{R}$ be continuous and such that there exist $C > 0$ and $a, b \geq 0$ with $|\psi(x,y)| \leq Ce^{a|x|+by}$ for every $(x,y) \in \mathbb{R} \times [0,+\infty)$. Then, if*

$$\lambda > ab|\rho|\sigma + \frac{b^2\sigma^2}{2} - \kappa b + \frac{a^2 - a}{2},$$

we have $P_t^\lambda|\psi|(x,y) < \infty$ for every $t \geq 0$, $(x,y) \in \mathbb{R} \times [0,+\infty)$ and the function $(t,x,y) \mapsto P_t^\lambda\psi(x,y)$ is continuous on $[0,\infty) \times \mathbb{R} \times [0,\infty)$.

Proof. We can prove, as in the proof of Proposition 4.8, that

$$\mathbb{E}_{x,y}\left(e^{aX_t + bY_t - \lambda\int_0^t Y_s ds}\right) = e^{a\left(x - \frac{\rho}{\sigma}y\right)}\mathbb{E}_y\left(e^{\left(a\frac{\rho}{\sigma}+b\right)Y_t + \left(a(\frac{\rho\kappa}{\sigma} - \frac{1}{2}) + \frac{a^2}{2}(1-\rho^2) - \lambda\right)\int_0^t Y_s ds}\right).$$

Thanks to Proposition 4.5, if

$$\frac{\sigma^2}{2}\left(a\frac{\rho}{\sigma} + b\right)^2 - \kappa\left(a\frac{\rho}{\sigma} + b\right) + \left(a(\frac{\rho\kappa}{\sigma} - \frac{1}{2}) + \frac{a^2}{2}(1-\rho^2) - \lambda\right) < 0, \tag{4.14}$$

we have, for any $T > 0$ and for any compact $K \subseteq \mathbb{R} \times [0,+\infty[$,

$$\sup_{(t,x,y)\in[0,T]\times K} \mathbb{E}_{x,y}\left(e^{aX_t + bY_t - \lambda\int_0^t Y_s ds}\right) < \infty.$$

Note that (4.14) is equivalent to

$$\lambda > ab\rho\sigma + \frac{b^2\sigma^2}{2} - \kappa b + \frac{a^2 - a}{2}.$$

Therefore, under the assumptions of the Proposition, we have, for any $T > 0$ and for any compact set $K \subseteq \mathbb{R} \times [0,+\infty[$,

$$\sup_{(t,x,y)\in[0,T]\times K} \mathbb{E}_{x,y}\left(e^{a|X_t| + bY_t - \lambda\int_0^t Y_s ds}\right) < \infty.$$

Moreover, for ϵ small enough,

$$\sup_{(t,x,y)\in[0,T]\times K} \mathbb{E}_{x,y}\left(e^{a(1+\epsilon)|X_t| + b(1+\epsilon)Y_t - \lambda(1+\epsilon)\int_0^t Y_s ds}\right) < \infty. \tag{4.15}$$

Then, let ψ be a continuous function on $\mathbb{R} \times [0,+\infty[$ such that $|\psi(x,y)| \leq Ce^{a|x|+by}$. It is evident that $P_t^\lambda|\psi|(x,y) < \infty$ and we have

$$P_t^\lambda\psi(x,y) = \mathbb{E}\left(e^{-\lambda\int_0^t (1+Y_s^y)ds}\psi(X_t^{x,y}, Y_t^y)\right).$$

If $((t_n,x_n,y_n))_n$ converges to (t,x,y), we deduce from Lemma 4.13 that $X_{t_n}^{x_n,y_n}\to X_t^{x,y}$, $Y_{t_n}^{y_n}\to Y_t^y$ and $\int_0^{t_n}Y_s^{y_n}ds\to\int_0^t Y_s^y ds$ in probability. Therefore $e^{-\lambda\int_0^{t_n}(1+Y_s)ds}\psi(X_{t_n}^{x_n,y_n},Y_{t_n}^{y_n})$ converges to $e^{-\lambda\int_0^{t}(1+Y_s)ds}\psi(X_{t}^{x,y},Y_{t}^{y})$ in probability. The estimate (4.15) ensures the uniformly integrability of $e^{-\lambda\int_0^{t_n}(1+Y_s)ds}\psi(X_{t_n}^{x_n,y_n},Y_{t_n}^{y_n})$ so that $\lim_{n\to\infty}P_{t_n}^\lambda\psi(x_n,y_n)=P_t^\lambda\psi(x,y)$ which concludes the proof. □

Proposition 4.15. *Fix $p>\beta+\frac{5}{2}$ and λ as in Theorem 4.9. Let us consider $u\in C([0,T];H)\cap L^2([0,T];V)$, with $\frac{\partial u}{\partial t}\in L^2([0,T];H)$ such that*

$$\begin{cases}\left(\frac{\partial u}{\partial t},v\right)_H+a_\lambda(u(t),v)=(f(t),v)_H, & v\in V,\\ u(0)=\psi,\end{cases}$$

with ψ continuous, $\psi\in V$, $\sqrt{1+y}f\in L^2([0,T];H)$ and $f\in L^p([0,T];L^p(\mathcal{O},\mathfrak{m}))$. Then, if ψ and λ satisfy the assumptions of Proposition 4.14, we have

1. *For every $t\in[0,T]$, $u(t)=P_t^\lambda\psi+\int_0^t P_s^\lambda f(t-s)ds$.*
2. *The function $(t,x,y)\mapsto u(t,x,y)$ is continuous on $[0,T]\times\mathbb{R}\times[0,+\infty)$.*
3. *If $\Lambda_t=\lambda\int_0^t(1+Y_s)ds$, the process $(M_t)_{0\le t\le T}$, defined by*

$$M_t=e^{-\Lambda_t}u(T-t,X_t,Y_t)+\int_0^t e^{-\Lambda_s}f(T-s,X_s,Y_s)ds,$$

with $X_0=x$, $Y_0=y$ is a martingale for every $(x,y)\in\mathbb{R}\times[0,+\infty)$.

Proof. The first assertion follows from Proposition 4.2.

The continuity of $(t,x,y)\mapsto P_t^\lambda\psi(x,y)$ is given by Proposition 4.14. The continuity of $(t,x,y)\mapsto\int_0^t P_s^\lambda f(t-s,.)(x,y)ds$ is trivial if $(t,x,y)\mapsto f(t,x,y)$ is bounded continuous. If $f\in L^p([0,T];L^p(\mathcal{O},\mathfrak{m}))$, f is the limit in L^p of a sequence of bounded continuous functions and we have $\int_0^t P_s^\lambda f_n(t-s,\cdot)ds\to\int_0^t P_s^\lambda f(t-s,\cdot)ds$ uniformly in $[0,T]\times K$ for every compact K of $\mathbb{R}\times[0,+\infty)$). In fact, thanks to Theorem 4.9, we can write for $t\in[0,T]$ and $(x,y)\in K$

$$\begin{aligned}\int_0^t P_s^\lambda|f_n-f|(t-s,\cdot,\cdot)(x,y)ds&\le\int_0^t\frac{C_{p,K,T}}{s^{\frac{2\beta+3}{2p}}}ds||(f_n-f)(t-s,\cdot,\cdot)||_{L^p(\mathcal{O},\mathfrak{m})}\\&\le C_{p,K,T}\left(\int_0^t||(f_n-f)(t-s,\cdot,\cdot)||^p_{L^p(\mathcal{O},\mathfrak{m})}ds\right)^{1/p}\left(\int_0^t\frac{ds}{s^{\frac{2\beta+3}{2(p-1)}}}\right)^{1-\frac{1}{p}}\\&\le C_{p,K,T}\left(\int_0^T||(f_n-f)(s,\cdot,\cdot)||^p_{L^p(\mathcal{O},\mathfrak{m})}ds\right)^{1/p}\left(\int_0^T\frac{ds}{s^{\frac{2\beta+3}{2(p-1)}}}\right)^{1-\frac{1}{p}}.\end{aligned}\tag{4.16}$$

The assumption $p>\beta+\frac{5}{2}$ ensures the convergence of the integral in the right hand side.

For the last assertion, note that $M_T=e^{-\Lambda_T}\psi(X_T,Y_T)+\int_0^T e^{-\Lambda_s}f(T-s,X_s,Y_s)ds$. Then, we can prove that M_t is integrable with the same arguments that we used to show the continuity of $(t,x,y)\mapsto u(t,x,y)$. Moreover, by using the Markov property,

$$\begin{aligned}\mathbb{E}_{x,y}\left(M_T\,|\,\mathcal{F}_t\right)&=e^{-\Lambda_t}P_{T-t}^\lambda\psi(X_t,Y_t)+\int_0^t e^{-\Lambda_s}f(T-s,X_s,Y_s)ds+e^{-\Lambda_t}\int_t^T P_{s-t}^\lambda f(T-s,.,.)(X_t,Y_t)ds\\&=e^{-\Lambda_t}\left(P_{T-t}^\lambda\psi(X_t,Y_t)+\int_0^{T-t}P_s^\lambda f(T-t-s,.,.)(X_t,Y_t)ds\right)+\int_0^t e^{-\Lambda_s}f(T-s,X_s,Y_s)ds\\&=e^{-\Lambda_t}u(T-t,X_t,Y_t)+\int_0^t e^{-\Lambda_s}f(T-s,X_s,Y_s)ds=M_t.\end{aligned}$$

□

We are now ready to prove the following proposition.

Proposition 4.16. *Assume that ψ satisfies Assumption $\mathcal{H}^*$. Moreover, fix $p > \beta + \frac{5}{2}$ and assume that $\psi \in L^2([0,T]; H^2(\mathcal{O}, \mathfrak{m}))$ and $\frac{\partial \psi}{\partial t} + \mathcal{L}\psi \in L^p([0,T]; L^p(\mathcal{O}, \mathfrak{m}))$. Then, the solution u of the variational inequality (2.5) satisfies*

$$u(t,x,y) = u^*(t,x,y), \qquad \text{on } [0,T] \times \bar{\mathcal{O}}, \tag{4.17}$$

where u^* is defined by

$$u^*(t,x,y) = \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}\left[\psi(\tau, X_\tau^{t,x,y}, Y_\tau^{t,y})\right].$$

Proof. We first check that ψ satisfies the assumptions of Proposition 4.12. Note that, thanks to the growth condition (2.6), it is possible to write $0 \le \psi(t,x,y) \le \Phi(t,x,y)$ with $\Phi(t,x,y) = C_T(e^{x - \frac{\rho\kappa\theta}{\sigma}t} + e^{Ly - \kappa\theta Lt})$, where $L \in \left[0, \frac{2\kappa}{\sigma^2}\right)$ and C_T is a positive constant which depends on T. Moreover, recall the growth condition on the derivatives (2.7). Then, it is easy to see that we can choose γ and μ in the definition of the measure $\mathfrak{m}$ (see (2.2.2)) such that ψ satisfies Assumption $\mathcal{H}^1$, Φ satisfies Assumption $\mathcal{H}^2$ (note that $\frac{\partial \Phi}{\partial t} + \mathcal{L}\Phi \le 0$) and $(1+y)\Phi$, $\frac{\partial \psi}{\partial t} + \mathcal{L}\psi \in L^p([0,T]; L^p(\mathcal{O}, \mathfrak{m}))$. Therefore we can apply Proposition 4.12 and we get that, for λ large enough, there exists $F \in L^p([0,T]; L^p(\mathcal{O}, \mathfrak{m}))$ such that u satisfies

$$-\left(\frac{\partial u}{\partial t}, v\right)_H + a_\lambda(u,v) = (F,v)_H, \qquad v \in V,$$

that is

$$-\left(\frac{\partial u}{\partial t}, v\right)_H + a(u,v) = (F - \lambda(1+y)u, v)_H, \qquad v \in V.$$

On the other hand we know that

$$\begin{cases} -\left(\frac{\partial u}{\partial t}, v-u\right)_H + a(u, v-u) \ge 0, & \text{a.e. in } [0,T] \qquad v \in V,\ v \ge \psi, \\ u(T) = \psi(T), \\ u \ge \psi \text{ a.e. in } [0,T] \times \mathbb{R} \times (0,\infty). \end{cases}$$

From the previous relations we easily deduce that $F - \lambda(1+y)u \ge 0$ a.e. and, taking $v = \psi$, that $(F - \lambda(1+y)u, \psi - u)_H = 0$. Moreover, note that the assumptions of Proposition 4.15 are satisfied, so the process $(M_t)_{0 \le t \le T}$ defined by

$$M_t = e^{-\Lambda_t} u(t, X_t, Y_t) + \int_0^t e^{-\Lambda_s} F(s, X_s, Y_s) ds, \tag{4.18}$$

with $X_0 = x$, $Y_0 = y$ is a martingale for every $(x,y) \in \mathbb{R} \times [0, +\infty)$. Then, we deduce that the process

$$\tilde{M}_t = u(t, X_t, Y_t) + \int_0^t \left(F(s, X_s, Y_s) - \lambda(1+Y_s)u(s, X_s, Y_s)\right) ds$$

is a local martingale. In fact, from (4.18) we can write

$$\begin{aligned} d\tilde{M}_t &= d\left[e^{\Lambda_t} M_t - e^{\Lambda_t} \int_0^t e^{-\Lambda_s} F(s, X_s, Y_s) ds\right] + F(t, X_t, Y_t) dt - \lambda(1+Y_t) u(t, X_t, Y_t) dt \\ &= e^{\Lambda_t} dM_t + \Big[\lambda(1+Y_t) e^{\Lambda_t} M_t - \lambda(1+Y_t) e^{\Lambda_t} \int_0^t e^{-\Lambda_s} F(s, X_s, Y_s) ds \\ &\quad - e^{\Lambda_t} e^{-\Lambda_t} F(t, X_t, Y_t) + F(t, X_t, Y_t) - \lambda(1+Y_t) u(t, X_t, Y_t)\Big] dt \\ &= e^{\Lambda_t} dM_t. \end{aligned}$$

So, for any stopping time τ there exists an increasing sequence of stopping times $(\tau_n)_n$ such that $\lim_n \tau_n = \infty$ and

$$\mathbb{E}_{x,y}[u(\tau \wedge \tau_n, X_{\tau \wedge \tau_n}, Y_{\tau \wedge \tau_n})] = u(0,x,y) - \mathbb{E}_{x,y}\left[\int_0^{\tau \wedge \tau_n} (F(s, X_s, Y_s) - \lambda(1+Y_s) u(s, X_s, Y_s)) ds\right]. \tag{4.19}$$

Since $F - \lambda(1+y)u \geq 0$ we can pass to the limit in the right hand side of (4.19) thanks to the monotone convergence theorem. Recall now that an adapted right continuous process $(Z_t)_{t\geq 0}$ is said to be of class $\mathcal{D}$ if the family $(Z_\tau)_{\tau\in\mathcal{T}_{0,\infty}}$, where $\mathcal{T}_{0,\infty}$ is the set of all stopping times with values in $[0,\infty)$, is uniformly integrable. Moreover, recall that $0 \leq u(t,x,y) \leq \Phi(x,y) = C_T(e^{x-\frac{\rho\kappa\theta}{\sigma}t} + e^{Ly-\kappa\theta Lt})$. The discounted and dividend adjusted price process $(e^{-(r-\delta)t}S_t)_t = (e^{X_t - \frac{\rho\kappa\theta}{\sigma}t})_t$ is a martingale (we refer to [14] for an analysis of the martingale property in general affine stochastic volatility models), so we deduce that it is of class $\mathcal{D}$. On the other hand, we can prove that the process $(e^{LY_t-\kappa\theta t})_t$ is of class $\mathcal{D}$ following the same arguments used in Remark 4.6. Therefore, the process $(\Phi(t+s, X_s^{t,x,y}))_{s\in[t,T]}$ is of class $\mathcal{D}$ for every $(t,x,y) \in [0,T]\times\mathbb{R}\times[0,\infty)$. So we can pass to the limit in the left hand side of (4.19) and we get that $\lim_{n\to\infty}\mathbb{E}_{x,y}[u(\tau\wedge\tau_n, X_{\tau\wedge\tau_n}, Y_{\tau\wedge\tau_n})] = \mathbb{E}_{x,y}[u(\tau, X_\tau, Y_\tau)]$. Therefore, passing to the limit as $n\to\infty$, we get

$$\mathbb{E}_{x,y}[u(\tau, X_\tau, Y_\tau)] = u(0,x,y) - \mathbb{E}_{x,y}\left[\int_0^\tau (F(s,X_s,Y_s) - \lambda(1+Y_s)u(s,X_s,Y_s))ds\right],$$

for every $\tau \in \mathcal{T}_{[0,T]}$. Recall that $F - \lambda(1+y)u \geq 0$, so the process $u(t,X_t,Y_t)$ is actually a supermartingale. Since $u \geq \psi$, we deduce directly from the definition of Snell envelope that $u(t,X_t,Y_t) \geq u^*(t,X_t,Y_t)$ a.e. for $t\in[0,T]$.

In order to show the opposite inequality, we consider the so called continuation region

$$\mathcal{C} = \{(t,x,y) \in [0,T)\times\mathbb{R}\times[0,\infty) : u(t,x,y) > \psi(t,x,y)\},$$

its t-sections

$$\mathcal{C}_t = \{(x,y)\in\mathbb{R}\times[0,\infty) : (t,x,y)\in\mathcal{C}\}, \qquad t\in[0,T),$$

and the stopping time

$$\tau_t = \inf\{s\geq t : (s,X_s,Y_s)\notin\mathcal{C}\} = \inf\{s\geq t : u(s,X_s,Y_s) = \psi(s,X_s,Y_s)\}.$$

Note that $u(x,X_s,Y_s) > \psi(s,X_s,Y_s)$ for $t\leq s<\tau_t$. Moreover, recall that $(F-\lambda(1+y)u, \psi-u) = 0$ a.e., so $Leb\{(x,y)\in\mathcal{C}_t : F-\lambda(1+y)u \neq 0\} = 0\,dt$ a.e.. Since the two dimensional diffusion (X,Y) has a density, we deduce that $\mathbb{E}\left[F(s,X_s,Y_s) - \lambda(1+Y_s)u(s,X_s,Y_s)\mathbf{1}_{\{(X_s,Y_s)\in\mathcal{C}_s\}}\right] = 0$, and so $F(s,X_s,Y_s) - \lambda(1+Y_s)u(s,X_s,Y_s) = 0\ ds,\ d\mathbb{P}-a.e.$ on $\{s<\tau_t\}$. Therefore,

$$\mathbb{E}\left[u(\tau_t, X_{\tau_t}, Y_{\tau_t})\right] = \mathbb{E}\left[u(t,X_t,Y_t)\right],$$

and, since $u(\tau_t, X_{\tau_t}, Y_{\tau_t}) = \psi(\tau_t, X_{\tau_t}, Y_{\tau_t})$ thanks to the continuity of u and ψ,

$$E\left[u(t,X_t,Y_t)\right] = \mathbb{E}\left[\psi(\tau_t, X_{\tau_t}, Y_{\tau_t})\right] \leq \mathbb{E}\left[u^*(t,X_t,Y_t)\right],$$

so that $u(t,X_t,Y_t) = u^*(t,X_t,Y_t)$ a.e.. With the same arguments we can prove that $u(t,x,y) = u^*(t,x,y)$ and this concludes the proof. □

4.5.2 Weaker assumptions on ψ

The last step is to establish the equality $u = u^*$ under weaker assumptions on ψ, so proving Theorem 2.4.

Proof of Theorem 2.4. First assume that there exists a sequence $(\psi_n)_{n\in\mathbb{N}}$ of continuous functions on $[0,T]\times\mathbb{R}\times[0,\infty)$ which converges uniformly to ψ and such that, for each $n\in\mathbb{N}$, ψ_n satisfies the assumptions of Proposition 4.16. For every $n\in\mathbb{N}$, we set $u_n = u_n(t,x,y)$ the unique solution of the variational inequality (2.3) with final condition $u_n(T,x,y) = \psi_n(T,x,y)$ and $u_n^*(t,x,y) = \sup_{\tau\in\mathcal{T}_{t,T}}\mathbb{E}[\psi_n(\tau, X_\tau^{t,x,y}, Y_\tau^{t,y})]$. Then, thanks to Proposition 4.16, for every $n\in\mathbb{N}$ we have

$$u_n(t,x,y) = u_n^*(t,x,y) \qquad \text{on } [0,T]\times\bar{\mathcal{O}}.$$

Now, the left hand side converges to $u(t,x,y)$ thanks to the Comparison Principle. As regards the right hand side,

$$\sup_{\tau\in\mathcal{T}_{t,T}}\mathbb{E}\left[\psi_n(\tau, X_\tau^{t,x,y}, Y_\tau^{t,x,y})\right] \to \sup_{\tau\in\mathcal{T}_{t,T}}\mathbb{E}\left[e^{-r(\tau-t)}\psi(\tau, X_\tau^{t,x,y}, Y_\tau^{t,x,y})\right]$$

thanks to the uniform convergence of ψ_n to ψ.

Therefore, it is enough to prove that, if ψ satisfies Assumption $\mathcal{H}^*$, then it is the uniform limit of a sequence of functions ψ_n which satisfy the assumptions of Proposition 4.16. This can be done following the very same arguments of [13, Lemma 3.3] so we omit the technical details (see [17]). □

References

[1] M. Abramowitz and I. A. Stegun (1992): *Handbook of mathematical functions with formulas, graphs and mathematical tables.* Dover publications, Inc., New York.

[2] A. Alfonsi (2015): *Affine diffusions and related processes: simulation, theory and applications.* volume 6 of *Bocconi & Springer Series*. Springer, Cham; Bocconi University Press, Milan.

[3] L. Andersen (2008): Simple and efficient simulation of the Heston stochastic volatility model. *J. Comput. Financ.,* **11**(3), 1-42.

[4] A. Bensoussan, J.L. Lions (1982): *Applications of aariational inequalities in Stochastic Control*, volume 12 of *Studies in Mathematics and its Applications.* North-Holland Publishing Co., Amsterdam-New York.

[5] P. Daskalopoulos and P. Feehan (2011): Existence, uniqueness and global regularity for degenerate elliptic obstacle problems in mathematical finance. *arxiv:1109.1075.*

[6] P. Daskalopoulos and P. Feehan (2016): $C^{1,1}$ regularity for degenerate elliptic obstacle problems in mathematical finance. *J. Differential Equations*, **26**(6), 5043-5074.

[7] E. Ekstrom, J. Tysk (2010): The Black–Scholes equation in stochastic volatility models. *J. Math. Anal. Appl.*, **368**(2), 498-507.

[8] P. Feehan, C. A. Pop (2015): Stochastic representation of solutions to degenerate elliptic and parabolic boundary value and obstacle problems with Dirichlet boundary conditions. *Trans. Amer. Math. Soc.*, **367**(2), 981-1031.

[9] P. Feehan, C. A. Pop (2015): Higher-order regularity for solutions to degenerate elliptic variational equations in mathematical finance. *Adv. Differential Equations* **20**, 361–432.

[10] A. Friedman (2010): *Variational principles and free-boundary problems.*. Courier Corporation, 2010.

[11] S. L. Heston (1993): A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies*, **6**, 327-343.

[12] N. Ikeda, S. Watanabe (1989): *Stochastic differential equations and diffusion processes.* Volume 24 of *North-Holland Mathematical Library.* North-Holland Publishing Co., Amsterdam;Kpdansha, Ltd., Tokio,second edition.

[13] P. Jaillet, D. Lamberton, and B. Lapeyre (1990): Variational inequalities and the pricing of American options, *Acta Appl. Math.*, **21**(3), 263-289.

[14] M. Keller Ressel (2011): Moment explosions and long-term behaviour of affine stochastic volatility models. *Math. Finance,* **21**(1), 73-98.

[15] D. Lamberton, B. Lapeyre (2008): *Introduction to stochastic calculus applied to finance.* Chapman&Hall/CRC Financial Mathematics Series. Chapman&Hall/CRC, Boca Raton, FL, second edition.

[16] J. L. Lions(1969): *Quelques methodes de résolution des problémes aux limites non-linéaires.* Dunod, Gauthier-Villars, Paris.

[17] G. Terenzi (2018): American options in stochastic volatility models. Ph.D. thesis. In preparation.